

# Roots of Autocracy\*

Oded Galor

Marc Klemp

February 5, 2018

## Abstract

This research advances the hypothesis and establishes empirically that variations in population diversity across human societies, as determined in the course of the exodus of humans from Africa tens of thousands of years ago, had contributed to the differential formation of pre-colonial autocratic institutions across ethnic groups and the emergence and persistence of contemporary autocratic institutions across countries. Exploiting a novel geo-referenced data set of population diversity across ethnic groups, the study demonstrates that while diversity has amplified the importance of institutions in mitigating the adverse effects of social non-cohesiveness on productivity, it has contributed to inequality and the scope for domination, leading to the formation and persistence of institutions of the autocratic type.

**Keywords** Autocracy, Economic Growth, Diversity, Institutions, The Out-of-Africa Hypothesis, Comparative Development

**JEL Classification Codes** O1, O43, Z10

---

\*The authors are grateful for valuable comments from Alberto Alesina, James Fenske, Martin Fiszbein, Ed Glaeser, Elhanan Helpman, Stelios Michalopoulos, Nathan Nunn, Nico Voigtlaender, participants in the conferences *CEPR Growth Conference 2016*, London, *Deep Determinants of International Comparative Development*, Brown University, *Towards Sustained Economic Growth*, Barcelona, *MEHR*, Copenhagen, *NBER Political Economy*, *Association for the Study of Religion, Economics, and Culture*, *ASSA*, as well as seminar participants at Bonn, Columbia, Copenhagen, Hamburg, Harvard (history), Harvard-MIT (political economy), Kiel, Louvain, Oxford, Tel Aviv, and Warwick. The research of Galor is supported by NSF grant SES-1338426. The research of Klemp has been funded partially by the Carlsberg Foundation, the Danish Research Council (ref. no. 1329-00093 and ref. no. 1327-00245), and the European Commission (grant no. 753615). Oded Galor (oded.galor@brown.edu): Department of Economics and Population Studies and Training Center, Brown University, 64 Waterman St., Providence, RI 02912, USA. Marc Klemp (marc.klemp@brown.edu): Department of Economics and Population Studies and Training Center, Brown University, 64 Waterman St., Providence, RI 02912, USA, and Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, DK-1353 Copenhagen K, Denmark.

# 1 Introduction

Political institutions have been widely viewed as major determinants of economic growth and comparative economic development. The origins of existing variations in the nature of political institutions across the globe have been attributed to a variety of factors, underlying the contribution of economic prosperity and the threat of revolution to the onset of democracy, as well as the role of inequality, ethnic fractionalization, and class stratification in the formation of autocratic institutions.<sup>1</sup> Furthermore, attention has been drawn towards historical, legal, and geographical factors that have affected the prevailing political institutions across former colonies, highlighting economic incentives that induced colonial powers to impose extractive institutions in some regions of the world and inclusive ones in others.<sup>2</sup>

This research explores the origins of variation in the nature and the persistence of pre-colonial political institutions, highlighting one of the deepest roots of autocracy, molded during the dawn of the dispersion of anatomically modern humans across the globe.<sup>3</sup> The study advances the hypothesis and establishes empirically that variation in population diversity across human societies, as determined in the course of the exodus of *Homo sapiens* out of Africa tens of thousands of years ago, shaped the distribution of political institutions in early stages of development and has persistently affected the attributes of contemporary institutions across societies.

The hypothesized contribution of population diversity to the emergence of autocratic institutions rests on two fundamental building blocks. First, in view of the adverse effect of population diversity on social cohesiveness and aggregate productivity, the emergence of formal or informal institutions, and their associated code of conduct, have plausibly mitigated the detrimental effects of social non-cohesiveness on productivity. Second, population diversity, and its manifestation in heterogeneity in cognitive as well as physical traits, has conceivably fostered the degree of inequality in society, amplifying class stratification, the scope for domination and the implementation of autocratic rules.<sup>4</sup> Thus, the dual effect of population diversity on the demand for institutions as well as on the scope for domination has plausibly contributed to the emergence of autocratic institutions.

The association between population diversity and the degree of autocratic institutions is examined empirically across ethnic groups during the pre-colonial era, as well as across nations in the contemporary period. This empirical setting has several virtues. First, the examination of the

---

<sup>1</sup>See Lipset (1960); Aghion et al. (2004); Alesina and Giuliano (2015); Bentzen et al. (2017).

<sup>2</sup>See Engerman and Sokoloff (1997); La Porta et al. (1999); Acemoglu et al. (2001).

<sup>3</sup>More generally, the origins of state formation have been attributed to the establishment of sedentary communities and the subsequent rise in social complexity in the post-Neolithic Revolution era (Mann, 1986; Belfer-Cohen and Bar-Yosef, 2002). In particular, it has been associated with the rise in population density in the post-Neolithic period (Diamond, 1997), the rise in food surplus, due to climatic shocks, technological advancements, the gains from trade (Gosden, 1989; Allen, 1997; Arnold, 1993; Fenske, 2014; Litina, 2014), and the existence of storable crops (Mayshar et al., 2017).

<sup>4</sup>The hypothesis that population diversity is associated with the scope for domination is supported by evidence from one of the closest species to human beings. The common chimpanzee, whose level of genetic diversity of 0.82 is larger than the one present in the human species (Pemberton et al., 2013), is characterized by extreme social stratification associated with a dominating alpha male. In contrast, the bonobo, that diverged from the common chimpanzee due to their geographical isolation dictated by the Congo River, are characterized by greater genetic homogeneity and are more egalitarian.

role of population diversity in the pre-colonial as well as the modern era permits the analysis to shed light on the association between population diversity and the evolution of institutions over the course of human history. Second, the focus on nations as well as ethnic groups permit the exploration of the association between population diversity and the emergence of autocratic institutions in societies of different scales. Third, in view of the important effect of colonialism on the nature of institutions, the pre-colonial analysis, by construction, is immune from the potentially confounding effect of colonialism on the association between population diversity and autocracy. Fourth, the intertemporal setup permits the examination of the persistent effect of pre-colonial institutions on the contemporary ones, isolating the direct association between population diversity and contemporary institutions from its lingering association via the persistence of past institutions. Fifth, the focus on ethnic groups permits the analysis to disentangle the role of phenotypic diversity (within an ethnic group), from the potential role of ethnic diversity (across groups), in the emergence of autocratic institutions.

The empirical analysis is conducted in several layers, based on a novel geo-referenced dataset consisting of ethnic groups, for which population diversity is either observed, or can be predicted, and for which geographic, ethnographic and institutional characteristics have been recorded. The first layer, as outlined in Figure 1, explores the association between population diversity and the degree of autocratic institutions across pre-colonial ethnic groups as reported by the *Ethnographic Atlas*. Consistent with the first element of the proposed mechanism, according to which ethnic groups characterized by higher population diversity are more likely to form institutions that would mitigate the adverse effect of social non-cohesiveness on productivity, the empirical analysis establishes that ethnic groups that are characterized by a higher level of observed population diversity tend to possess more elaborate institutions, as captured by the degree of jurisdictional hierarchy in those societies. Further, in line with the second element of the proposed mechanism, which highlights the potential effect of population diversity on social stratification and the scope for domination, the empirical analysis establishes that ethnic groups that are characterized by a higher level of observed population diversity tend to have a higher level of class stratification and a greater intensity of slavery.

The ethnic-level empirical analysis further explores the potential effect of population diversity on the emergence of autocratic institutions. In line with the proposed mechanism, the empirical analysis establishes that the association between population diversity and the prevalence of autocratic institutions could have plausibly operated through its dual association with the formation of institutions as well as with class stratification and the scope for domination. In particular, the extent of jurisdictional hierarchy as well as the degree of social stratification and the intensity of slavery, as reported by the *Ethnographic Atlas*, are associated with the presence of autocratic institutions as captured by various measures such as: (i) the degree of absence of checks on leader's power, (ii) the difficulty of removal of leaders, (iii) the leader's exercise of authority, (iv) the degree of lack of community decisions, (v) the perception of leader's power, and (vi) indigenous autocracy. Moreover, there exists a positive reduced-form association between population diversity and the

various measures of autocracy. Thus, consistent with the proposed hypothesis, the first layer of the empirical analysis suggests that population diversity contributed to the degree of pre-colonial autocratic institutions across ethnic groups, while lending credence to the hypothesized mechanism that governs this reduced-form relationship, according to which population diversity contributed to the demand for institutions as well as for the scope for domination, giving rise to institutions of the autocratic type.

The research exploits several empirical strategies to mitigate concerns about the potential role of reverse causality, omitted cultural, geographical and human characteristics, as well as sorting in the observed association between population diversity and autocracy. In particular, the positive associations between the extent of the observed population diversity and the degree of autocracy may reflect reverse causality from autocracy to population diversity. It is not inconceivable that in the course of human history autocratic regimes had fostered domination and conquests of a wide range of populations and ethnic groups, and have therefore affected the observed levels of population diversity. Hence, in order to remove the concern about reverse causality, the analysis exploits predicted population diversity rather than observed diversity. In particular, since observed population diversity within a geographically indigenous contemporary ethnic group decreases with distance along ancient migratory paths from East Africa, as established by the serial founder effect (e.g., Harpending and Rogers, 2000; Ramachandran et al., 2005; Prugnolle et al., 2005; Ashraf and Galor, 2013), migratory distance from Africa is exploited to predict population diversity for the 1,267 ethnic groups in the *Ethnographic Atlas*, and this *predicted* level of population diversity is shown to be positively associated with the extent of autocracy.

Furthermore, the associations between ethnic level population diversity and the degree of autocracy may be governed or biased by omitted cultural, geographical and human characteristics. Thus, in order to mitigate these concerns, the empirical analysis exploits two related strategies. In light of the serial founder effect, the analysis exploits the migratory distance from Africa to each ethnic group as an instrumental variable for the observed level of population diversity, and as a predictor for its level of diversity. Nevertheless, there are several plausible scenarios that would weaken this identifying strategy. First, selective migration out of Africa, or natural selection along the migratory paths, could have affected human traits and therefore institutional development independently of the effect of migratory distance from Africa on the degree of diversity in human traits. Second, migratory distance from Africa could be correlated with distances from focal historical locations (e.g., distances from technological frontiers) and could therefore capture the effect of these distances on the process of development and the formation of institutions, rather than the indirect effect of these migratory distances via population diversity.

These potential concerns are mitigated by the following observations. First, while migratory distance from Africa has a significant negative association with the degree of genetic diversity, conditional on the distance from the equator, it has no association with the mean level of human traits, such as height, weight, skin reflectiveness, and IQ (Ashraf and Galor, 2013). Second, conditional on migratory distance from East Africa, migratory distances from historical technological

frontiers in the years 1, 1000, and 1500 are not significantly associated with autocracy, reinforcing the justification for the reliance on the out of Africa hypothesis and the serial founder effect.

Moreover, a highly implausible threat to the identification strategy would emerge if the actual migration path out of Africa would have been correlated with geographical characteristics that are directly conducive to economic development and hence to the development of institutions (e.g., soil quality, climatic conditions, and propensity to trade). This, however, would have implausibly necessitated that the conduciveness of these geographical characteristics to autocracy would be aligned along the main root of the migratory path out of Africa, as well as along each of the main forks that emerge from this primary path. In particular, in several important forks in the course of this migration process (e.g., from the fertile crescent and to associated eastward migration towards east Asia and western migration towards Europe), the geographical characteristics that are conducive to autocracy would have to diminish symmetrically along these diverging migratory roots. Nevertheless, in order to further mitigate this highly implausible concern, the analysis establishes that the results are unaffected qualitatively, if it accounts for the potentially confounding effects of a wide range of geographical factors in the homeland of each ethnic group, such as absolute latitude, average elevation, terrain ruggedness, coastal length, as well as climatic conditions captured by the average and standard deviations of temperature. Moreover, the analysis accounts for spatial auto-correlation as well as regional fixed effects, capturing time-invariant unobserved heterogeneity in each region and hence identifying the association within a geographical region rather than across regions. Furthermore it establishes that selection on unobservables is not a concern.

The observed associations between population diversity and the extent of autocratic institutions may further reflect the sorting of diverse populations into geographical niches characterized by autocratic institutions. While sorting would not affect the existence of a positive association between population diversity and the extent of autocracy, it could weaken the proposed mechanism. However, in view of the serial founder effect and the tight negative association between migratory distance from Africa and population diversity, sorting would necessitate that the ex-ante spatial distribution of autocratic institutions would have to be negatively correlated with migratory distance from Africa. As argued above, this would have implausibly necessitated that the conduciveness of geographical characteristics to autocracy would be aligned with the primary migratory path out of Africa, as well as with each of its diverging forks, and would diminish symmetrically along these diverging migratory roots. Nevertheless, to further mitigate this highly implausible scenario, the empirical analysis accounts for the potentially confounding effects of a wide range of geographical characteristics, as well as regional fixed effects.

The second layer of the empirical analysis, as outlined in Figure 1, explores the importance of the association between population diversity and pre-colonial autocratic institutions across ethnic groups for the understanding of the contemporary variation in autocratic institutions across nations. In particular, it examines the persistence of ethnic institutions, that were formed in the pre-colonial era, and their association with contemporary national institutions. Aggregating pre-colonial ethnic institutions into pre-colonial national institutions, the analysis suggests that indeed pre-colonial

ethnic institutions have contributed to contemporary institutions, beyond the persistent association with population diversity. In particular, autocratic institutions, and the absence of executive constraints in the contemporary period, are positively and significantly associated with the extent of autocratic institutions in the pre-colonial era, accounting for the potentially confounding effects of geographical characteristics and population diversity. Moreover, the findings suggest that the persistence of institutions can be partly attributed to the direct association between population diversity and both pre-colonial and contemporary institutions.

The third layer of the empirical analysis, as outlined in Figure 1, examines the reduced-form relationship between population diversity and the nature of contemporary national institutions. Consistent with the proposed hypothesis, it establishes that population diversity at the national level, as captured by ancestry-adjusted predicted population diversity, has a significant direct positive association with the degree of autocracy and with the absence of executive constraints across countries, accounting for a large number of confounding geographical characteristics, regional fixed effects, colonial history, legal origins, pre-colonial development and ethnolinguistic fractionalization (and its geographical origins). Moreover, the association remains nearly intact if one accounts for arguably endogenous controls such as income per capita and education. Thus, the third layer of the empirical analysis suggests that the spatial distribution of population diversity across the globe has also contributed to the contemporary variation in the degree of autocracy across countries. This reduced-form positive association of population diversity and the prevalence of contemporary autocratic institutions across nations may reflect either the persistence of institutions from the pre-colonial to the modern era, as established in the second layer of the analysis, or a direct association between population diversity and contemporary autocratic institutions, capturing the association between diversity and the demand for institutions as well as the scope for domination.

Finally, in light of the negative association between migratory distance from East Africa and the duration of settlements, one could have argued that societies at greater migratory distance from Africa had shorter time to evolve and to form autocratic institutions. Thus, the negative association between the migratory distance from Africa, genetic diversity, and the extent of autocracy may reflect the shorter duration of settlements at greater migratory distance from Africa. Nevertheless, while the duration of settlement is indeed negatively associated with autocracy, it has no qualitative effect on the association between diversity and autocracy. Moreover, in view of the potential association between population diversity and the onset of the Neolithic Revolution, one could have argued that the emergence of sedentary communities in the course of the Neolithic Revolution, rather than the dual effect of diversity, contributed to the onset of autocracy. However, accounting for the time elapsed since the Neolithic Revolution has no qualitative effect on the association between population diversity and autocracy.

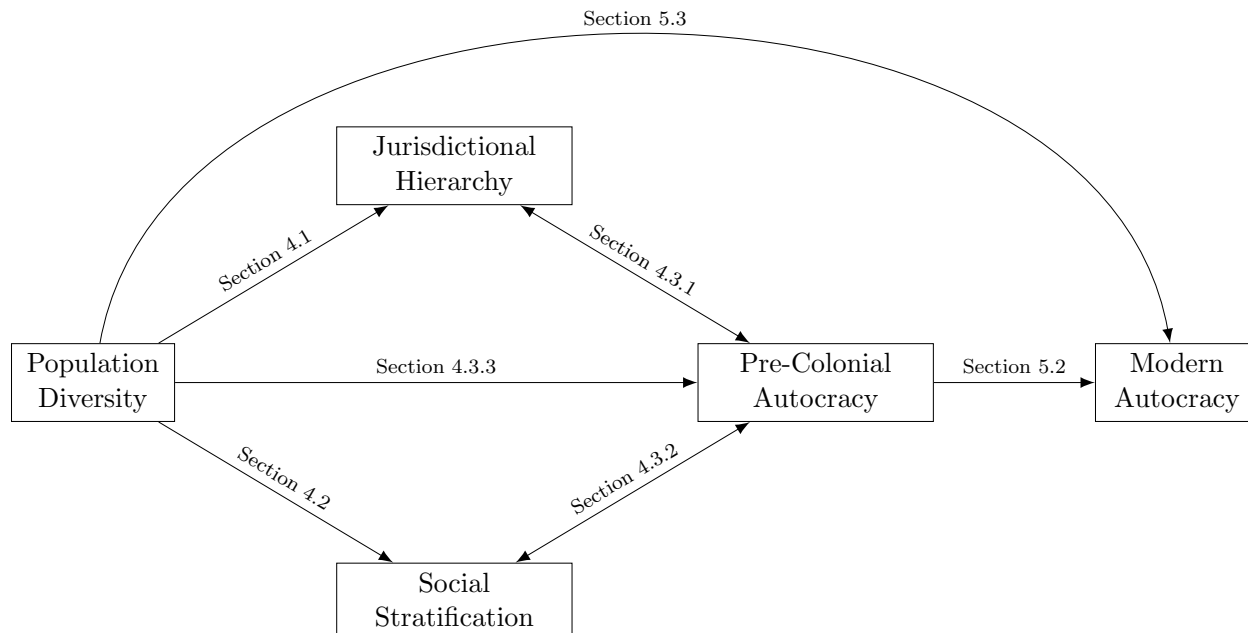


Figure 1: Overview of the structure of the empirical analysis.

## 2 Data

This section presents the novel data set that is used in the empirical analysis of the association between population diversity and the emergence and the persistence of autocratic institutions. In particular, it introduces the data on observed population diversity at the ethnic-group level, predicted population diversity for the entire set of ethnic groups in the *Ethnographic Atlas*, and ancestry adjusted predicted population diversity as well for nation states. Moreover it presents a range of measures of institutions, autocracy, social stratification, and scope for domination at the ethnic group level, as well as measures of autocracy at the national level.

### 2.1 Measures of Population Diversity

This research highlights the pivotal contribution of population diversity for the emergence of autocratic institutions. While population diversity can be captured by ethnolinguistic fractionalization, ethnolinguistic polarization, or genetic diversity, several reasons suggests that genetic diversity ought to be used in order to properly capture the hypothesized effects of population diversity.

First, while population diversity at the national level can be captured by each of the three measures, diversity within ethnic groups can be captured only by existing measures of genetic diversity. Second, for the country-level analysis, measures of ethnolinguistic fractionalization captures primarily the proportional representation of each ethnic group in the population, while measures of ethnic polarization incorporate proxies for pairwise dissimilarities amongst ethnic groups within

the population. In contrast, the genetic diversity of a national population is an index that incorporates information on all three dimensions of heterogeneity at the country level: the proportional representation of each ethnic group, the pairwise dissimilarities across these groups (as captured by genetic distance), and most importantly, the degree of interpersonal diversity within each group (as captured by genetic diversity within the group). Third, since the hypothesized effect of population diversity on inequality and social stratification is operating via heterogeneity in cognitive as well as physical traits, the measure of population diversity ought to reflect phenotypic diversity. A-priori the degree of either fractionalization or polarization does not necessarily reflect the degree of phenotypic diversity, whereas genetic diversity is correlated with phenotypic diversity.

Thus the study employs various measures of genetic diversity (i.e., observed diversity within ethnic groups, predicted diversity within ethnic groups, predicted diversity within countries, and ancestry-adjusted predicted diversity within countries) to capture population diversity.

### **2.1.1 Observed Population Diversity within Ethnic Groups**

Population geneticists use an index known as expected heterozygosity to measure the extent of diversity in genetic material across individuals in a given population (e.g., an ethnic group). Genetic diversity captures the probability that two individuals, selected at random from a given population, differ from one another with respect to a spectrum of genetic traits. In particular, the overall expected heterozygosity for a given population is the average gene-specific heterozygosity (based on the proportional representations of different alleles of this trait in the population) over multiple DNA loci.

Existing measures of expected heterozygosity for indigenous ethnic groups are created by population geneticists utilizing data on allelic frequencies within a particular class of DNA loci labeled microsatellites, located in non-protein-coding regions of the human genome and are largely regarded as selectively neutral. This attribute has a major advantage, assuring that this measure of genetic diversity is unaffected by factors that are correlated with political institutions that could have governed the process of natural selection. Nevertheless, a conceptually meaningful measure of genetic diversity (i.e., a measure that can capture the effect of political and economic outcomes) ought to reflect diversity in phenotypically expressed traits.

Reassuringly, diversity in microsatellites is positively correlated with heterogeneity in phenotypically expressed genomic material. In particular, similarly to expected heterozygosity in neutral genetic markers, evidence suggests that a serial founder effect associated with migratory distance from East Africa has a negative association with various forms of morphological and cognitive diversity (Henn et al., 2012), including diversity in skeletal features pertaining to cranial characteristics (Manica et al., 2007; von Cramon-Taubadel and Lycett, 2008; Betti et al., 2013), dental characteristics (Hanihara, 2008), and pelvic attributes (Betti et al., 2013), as well as phonemic diversity (Atkinson, 2011). Furthermore, as reported in Tables A.1 and A.2, in line with the proposed hypothesis, genetic diversity of the ancestral population is associated with a lower level of inter-



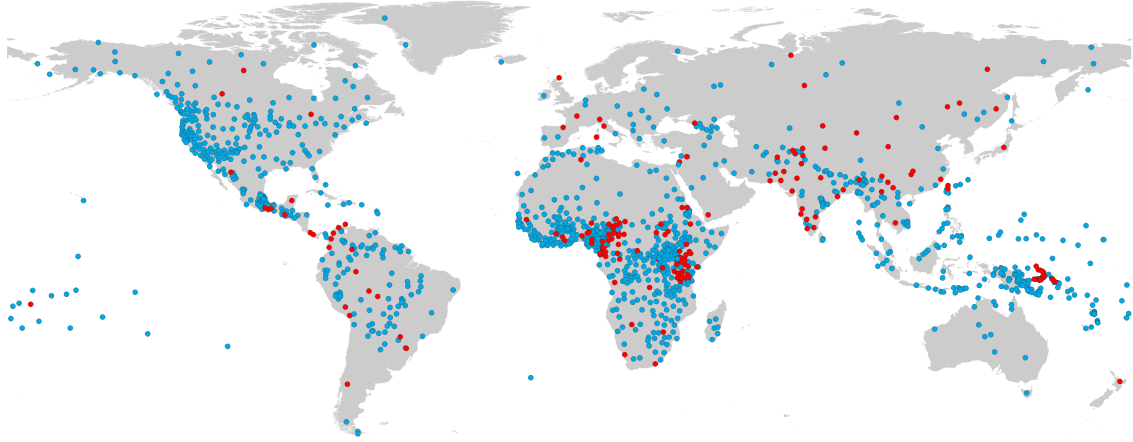


Figure 2: The interior centroids of the historical homelands of ethnic groups with both observed and predicted diversity (red) and only predicted diversity (blue).

personal trust (and therefore higher levels of social non-cohesiveness), among second-generation migrants to the US and among Africans migrants residing in Africa.

This research employs newly assembled data (Pemberton et al., 2013) on observed diversity in 232 predominantly indigenous ethnic groups across the globe that have been largely isolated from genetic flows from other ethnic groups.<sup>5</sup> The distribution of these ethnic groups across the globe is depicted in Figure 2 and the summary statistics of this measure of genetic diversity as documented in Table A.3 establishes that observed diversity ranges from 0.77 in Africa to 0.58 in South America.<sup>6</sup> Moreover, the study creates a novel geo-referenced dataset consisting of ethnic groups, for which observed population diversity is matched to geographic, ethnographic and institutional characteristics.<sup>7</sup>

### 2.1.2 Predicted Population Diversity within Ethnic Groups

The research exploits the tight negative relationship between migratory distance from East Africa and observed diversity across the ethnic group in the sample of (Pemberton et al., 2013) in order to generate a measure of predicted diversity for all ethnic group in the *Ethnographic Atlas*, overcom-

<sup>5</sup>This dataset combines eight human genetic diversity datasets based on the 645 loci that they share, including the HGDP-CEPH Human Genome Diversity Cell Line Panel used by Ashraf and Galor (2013).

<sup>6</sup>The analysis includes all observations on ethnic groups in Pemberton et al. (2013), excluding two ethnicities (the Surui and the Ache of South America) that are largely viewed as extreme outliers in terms of genetic diversity (e.g. Wang et al., 2007). The exclusion of these ethnicities is not particular to our study. In particular, Ramachandran et al. (2005) omits the Surui, being “an extreme outlier in a variety of previous analyses”, and did not include the Ache either. Furthermore, these ethnicities have the lowest levels of genetic diversity in the clean sample and the largest residuals of an OLS regression of genetic diversity on migratory distance from Addis Ababa. Including these observations, nevertheless, does not affect the qualitative analysis.

<sup>7</sup>A geo-referenced mapping of the ethnographic characteristics of the ethnic groups to shed light on contemporary cultural traits was first introduced by Nunn (2008).

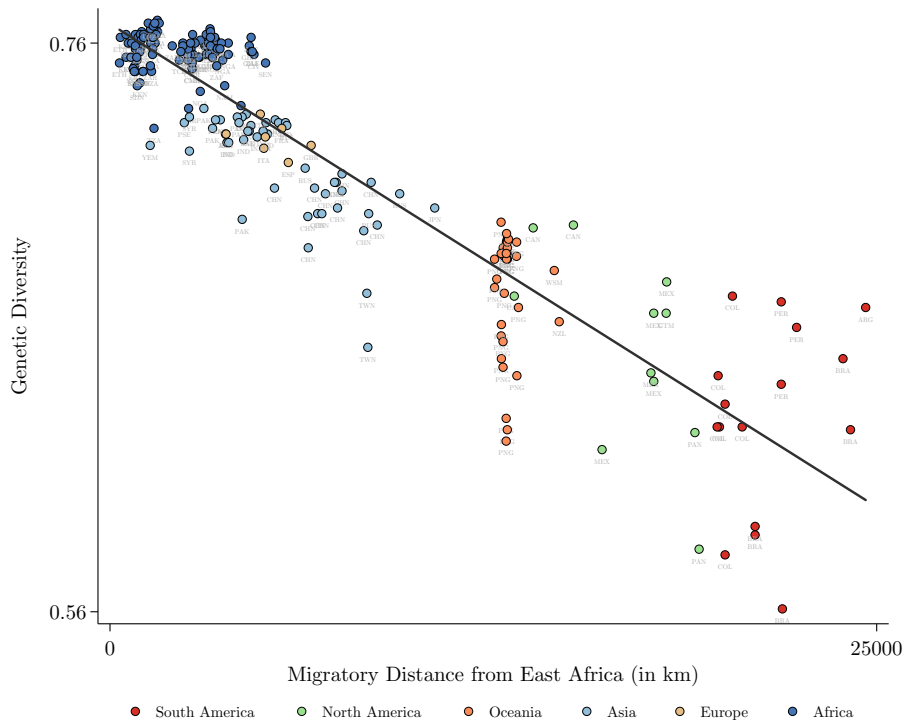


Figure 3: The negative association between migratory distance from East Africa and observed genetic diversity across the 230 ethnic groups in the sample.

ing sample limitations as well as potential concerns about selection on unobservables and reverse causality that may affect the relationship between observed diversity and political institutions.<sup>8</sup>

The composition of genetic traits within populations has evolved in the course of a stepwise migration process of anatomically modern human out of Africa 90,000–60,000 BP. This “out of Africa” migration was associated with a decline in the degree of genetic diversity in populations that settled at greater migratory distances from Africa. In particular, following the serial founder effect, since the spatial diffusion of humans occurred in a sequence of steps, in which a subgroup of individuals left their parental colony to establish a new settlement farther away, carrying only a subset of the genetic diversity of their parental colony, the extent of genetic diversity observed within an indigenous ethnic group decreases with its migratory distance from East Africa (e.g., Harpending and Rogers, 2000; Ramachandran et al., 2005; Prugnolle et al., 2005; Ashraf and Galor, 2013).<sup>9</sup>

<sup>8</sup>Since predicted population diversity for each of the ethnic groups in the sample is a generated regressor, the analysis based on predicted population diversity employs a two-step bootstrapping algorithm to compute the standard errors (Table A.4, A.5, and A.6 in the Appendix).

<sup>9</sup>In the pre-colonial era, the geographical locations of societies reflected the locations to which their ancestral populations had arrived at the culmination of their prehistoric “out of Africa” migration from the cradle of humankind, and as such, the diversity of a precolonial society was presumably determined by the ancient serial founder effect originating in East Africa.

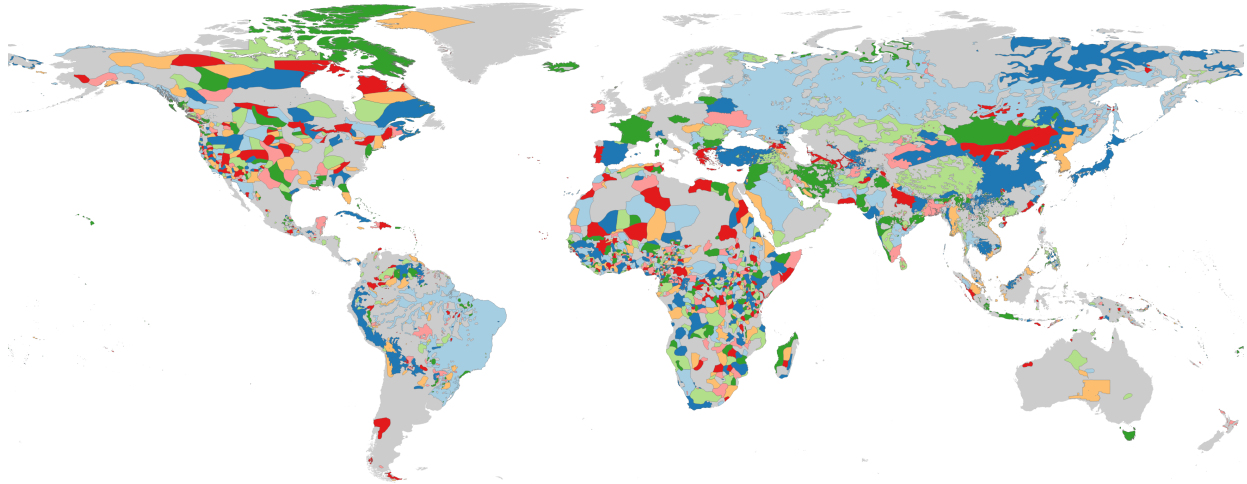


Figure 4: The historical homelands of ethnic groups in the dataset.

Reflecting this chain of ancient population bottlenecks originating in East Africa, the scatter plot in Figure 3 depicts the highly significant negative association between migratory distance and the cradle of mankind in East Africa on genetic diversity in the ethnic-group sample of Pemberton et al. (2013).<sup>10</sup> In particular, regressing genetic diversity on migratory distance from Addis Ababa in this sample suggest that expected heterozygosity falls by 6 percentage points for every 10,000 km increase in migratory distance from East Africa. Furthermore, migratory distance explains 85 percent of the cross-group variation in observed diversity.<sup>11</sup>

The distribution of ethnic groups across the globe in this predicted diversity sample is depicted in Figure 2, and the summary statistics of this measure of genetic diversity documented in Table A.3, establishes that observed diversity ranges from 0.77 in Africa to 0.59 in South America. Furthermore, the geographical characteristics of these ethnic groups are based on their geographical homelands as depicted in Figure 4 and further described in Appendix C.

### 2.1.3 Predicted Population Diversity within Countries

While the observed and predicted population diversity for ethnic groups is sufficient in order to conduct the ethnic-level analysis, the examination of the association between diversity and autocratic institutions across countries requires the use of a genetic diversity measure for national populations. However, national contemporary populations are composed of multiple ethnicities which may not be indigenous to their current geographical locations. Thus, one needs to construct

<sup>10</sup>Figure B.1 in the Appendix shows the corresponding figure including two ethnicities (the Surui and the Ache of South America) that are largely viewed as extreme outliers in terms of genetic diversity (e.g. Wang et al., 2007).

<sup>11</sup>As further elaborated in the Appendix, in estimating the migratory distance from Addis Ababa (East Africa) for each of the ethnic groups in the data, the shortest traversable paths from Addis Ababa to the interior centroid of each ethnic group was computed. Given the limited ability of humans to travel across large bodies of water, the traversable area included bodies of water at a distance of 100 km from land mass (excluding migration from Africa into Europe via Italy or Spain).

an index of genetic diversity for contemporary national populations that accounts for the proportional representation of each ethnic group within the country, the expected heterozygosity within each subnational group, as well as the diversity that arises from the genetic distances between the precolonial ancestral populations. Hence, the country-level analysis employs the measure of genetic diversity, as constructed by Ashraf and Galor (2013), accounting for these three important elements of population diversity within a national population.<sup>12</sup> The summary statistics of this measure of genetic diversity is documented in Table A.3.

## 2.2 Measures of Political Institutions, Autocracy, and Social Stratification

In view of the hypothesis that diversity contributed to demand for institutions as well as to the emergence of pre-colonial autocratic institutions which persisted over time and affected the nature of contemporary institutions, measures of pre-colonial institutions and pre-colonial and modern autocracy will be used.

For the analysis of pre-colonial institutions, the study exploits the arguably largest and most comprehensive collection of ethnographic tabulations found in the *Ethnographic Atlas*, consisting of ethnographic data for 1,267 worldwide ethnic groups (Murdock, 1967). Pre-colonial ethnic level institutions are captured by “Jurisdictional Hierarchy Beyond Local Community”, as reported by the *Ethnographic Atlas*. This widely used measure of pre-colonial institution (or statehood) consists of five gradations: (i) no political authority beyond the local community, (ii) one level (e.g., petty chiefdoms), (iii) two levels (e.g., larger chiefdoms), (iv) three levels (e.g., states), and (v) four levels (e.g., large states). The geographical distribution of the pre-colonial institutions measure is depicted in Figure B.2 in the Appendix.

The presence of autocratic institutions in the pre-colonial era is captured by various measures such as: (i) degree of absence of checks on leader’s power, (ii) difficulty of removal of leaders, (iii) leader’s exercise of authority, (iv) degree of lack of community decisions, and (v) perception of leader’s power, as reported by the *Standard Cross Cultural Survey* (Murdock and White, 1969). Furthermore, the autocratic institutions in the pre-colonial era is captured by a measure of indigenous autocracy based on the variable “Succession to the Office of Local Headman” from the *Ethnographic Atlas*. The geographical distribution of the indigenous autocracy measure is depicted in Figure B.3 in the Appendix. Moreover, the degree of contemporary autocratic institutions is captured by the indexes of “Autocracy” and “Constraints on the Executive” as reported by the gold standard in comparative research in political institutions: The Polity IV Project dataset (Marshall et al., 2014). The geographical distribution of the autocracy measure is depicted in Figure B.4 in the Appendix.

In light of the hypothesis that diversity contributed to the degree of social stratification and the scope for domination, contributing to the emergence of pre-colonial autocratic institutions,

---

<sup>12</sup>Since predicted population diversity for each country is a generated regressor, the empirical analysis based on predicted population diversity employs a two-step bootstrapping algorithm to compute the standard errors (Table A.4, A.5, and A.6 in the Appendix).

two ethnographic characteristics are used in order to capture these elements. First, “Class Stratification” as reported by the *Ethnographic Atlas* is used. This measure of social stratification is aggregated into three gradations: (i) absence of stratification, (ii) the presence of wealth distinctions or elite, and (iii) the presence of complex social classes or hereditary aristocracy. The geographical distribution of the social stratification measure is depicted in Figure B.5 in the Appendix. Second the intensity of slavery as reported by the *Ethnographic Atlas* is used. The Intensity of Slavery variable is aggregated into three gradations: (i) absence or near absence of slavey, (ii) incipient or nonhereditary, and (iii) hereditary and socially significant. The geographical distribution of the indigenous autocracy measure is depicted in Figure B.6 in the Appendix.

### 3 Empirical Framework

#### 3.1 Empirical Strategy

This research advances the hypothesis that diversity contributed to the concentration of power over the course of human history within social groups, such as ethnic groups or nation states. Furthermore, it suggests that this effect was governed by the impact of population diversity on the emergence of institutions as well as on the degree of social stratification.

The association between population diversity and the degree of autocratic institutions is examined empirically across ethnic groups during the pre-colonial era, as well as across nations in the contemporary period. This empirical setting has several virtues. First, the examination of the association between population diversity in the pre-colonial as well as the modern era permits the analysis to shed light on the association between population diversity on the evolution of institutions over the course of human history. Second, the focus on nations as well as ethnic group, permit the exploration of the association between diversity and the emergence of autocratic institutions in societies of different scales. Third, in view of the important effect of colonialism on the nature of institutions, the pre-colonial analysis, by construction, is immune from the potentially confounding effect of colonialism on the association between population diversity and autocracy. Fourth, the intertemporal setup permits the examination of the persistent effect of pre-colonial institutions on the contemporary ones, isolating the direct association between diversity and contemporary institutions from its lingering association via the persistence of past institutions. Fifth, the focus on ethnic groups permits that analysis to disentangle the association between the emergence of autocratic institutions and phenotypic diversity (within an ethnic group), from its potential association with ethnic diversity (across groups).

The empirical analysis of the contribution of population diversity for the emergence and the persistence of autocratic institutions exploits several strategies to mitigate potential concerns regarding the potential roles of reverse causality, omitted variables, and sorting.

First, the positive associations between the extent of the observed population diversity and the degree of autocracy may reflect reverse causality from autocracy to population diversity. In particular, it is not inconceivable that in the course of human history autocratic regimes had fostered

domination and conquests of a wide range of populations and ethnic groups, and have therefore affected the observed levels of population diversity.

Thus, in order to remove this potential concern about reverse causality, the study exploits variations in predicted population diversity. In particular, rather than relying on observed population diversity within ethnic groups, that may conceivably be endogenous to the extent of autocracy, the analysis is conducted based on predicted diversity for each of the 1,267 ethnic groups in the *Ethnographic Atlas* based on their pre-historical migratory distances from east Africa, which are exogenous to the observed level of population diversity.

The onset of the migration of anatomically modern human from Africa, 60,000-90,000 years ago, was inherently associated with a reduction in the extent of genetic diversity in populations that settled at greater migratory distances from Africa. In particular, as follows from a serial founder effect, since the spatial diffusion of humans to the rest of the world occurred in a series of discrete steps, where in each step a subgroup of individuals left their parental colony to establish a new settlement farther away, carrying with them only a subset of the genetic diversity of their parental colony, the extent of genetic diversity observed within a geographically indigenous contemporary ethnic group decreases with distance along ancient migratory paths from East Africa (e.g., Harpending and Rogers, 2000; Ramachandran et al., 2005; Prugnolle et al., 2005; Ashraf and Galor, 2013). Indeed, migratory distance from the cradle of humankind in East Africa to indigenous settlements across the globe had a highly significant linear negative association with population diversity, capturing 86% of the variation in genetic diversity among the 53 ethnic groups in the Human Genome diversity Project and 84% of the variation in genetic diversity among the 232 ethnic groups in the expanded sample of Pemberton et al. (2013).

Second, the associations between ethnic-level population diversity and the degree of autocracy may be governed or biased by omitted cultural, geographical, and human characteristics. Thus, in order to mitigate these concerns, the empirical analysis exploits two related strategies. In light of the serial founder effect, the analysis exploits the migratory distance from Africa to each of ethnic group as: (i) an instrumental variable for the observed level of population diversity, and (ii) a predictor its level of diversity.

This identification strategy is based on the identifying assumption that the migratory distance from Africa affected economic or institutional outcomes only via its effect on genetic diversity. There are several plausible scenarios that would represent threats to this identifying assumption: First, selective migration out of Africa, or natural selection along the migratory paths, could have affected human traits and therefore institutional development independently of the effect of migratory distance from Africa on the degree of diversity in human traits. Second, migratory distance from Africa could be correlated with distances from focal historical locations (e.g., distances from technological frontiers) and could therefore capture the effect of these distances on the process of development and the formation of institutions, rather than the indirect effect of these migratory distances via population diversity.

These potential concerns are mitigated by the following observations. First, while migratory distance from Africa has a significant negative association with the degree of genetic diversity, conditional on the distance from the equator, it has no association with the mean level of human traits, such as height, weight, skin reflectiveness, and IQ (Ashraf and Galor, 2013). Second, conditional on migratory distance from East Africa, migratory distances to historical technological frontiers in the years 1, 1000, and 1500 are not significantly associated with autocracy, reinforcing the reliance on the out-of-Africa hypothesis and the serial founder effect.

Moreover, a highly implausible threat to the identification strategy would emerge if the actual migration path out of Africa would have been correlated with geographical characteristics that are directly conducive to economic development and hence to the development of institutions (e.g., soil quality, climatic conditions, and propensity to trade). This would have implausibly necessitated that the conduciveness of these geographical characteristics to autocracy would be aligned along the main root of the migratory path out of Africa, as well as along each of the main forks that emerge from this primary path. In particular, in several important forks in the course of this migration process (e.g., from the fertile crescent and to associated eastward migration towards east Asia and western migration towards Europe) the geographical characteristics that are conducive to autocracy would have to diminish symmetrically along these diverging migratory roots; a requirement that is at odds with the climatic, topographic, and geographical characteristics, in general and at some of these forks, in particular. Nevertheless, in order to mitigate further this highly implausible concern, the analysis explores the robustness of the results to the potentially confounding effects of a wide range of geographical factors in the homeland of each ethnic group. In particular, it accounts for geographical factors such as absolute latitude, average elevation, terrain ruggedness, coastal length, as well as climatic conditions captured by the average and standard deviations of temperature and precipitation. Moreover, the analysis accounts for: (i) regional fixed effects, capturing time invariant unobserved heterogeneity in each region, and hence identifying the association within a geographical region rather than across regions, (ii) spatial auto-correlation, and (iii) selection on unobservables.

The observed associations between population diversity and the extent of autocratic institutions may further reflect the sorting of diverse populations into geographical niches characterized by autocratic institutions. While sorting would not affect the existence of a positive association between population diversity and the extent of autocracy, it could weaken the proposed mechanism. However, in view of the serial founder effect and the tight negative association between migratory distance from Africa and population diversity, sorting would necessitate that the ex-ante spatial distribution of autocratic institution would have to be negatively correlated with migratory distance from Africa. As was argued above, this would have implausibly necessitated that the conduciveness of geographical characteristics to autocracy would be aligned with the primary migratory path out of Africa, as well as with each of its diverging forks, and will diminish symmetrically along these diverging migratory roots. Nevertheless, to further mitigate this highly implausible scenario, the

empirical analysis accounts for the potentially confounding effects of a wide range of geographical characteristics, as well as regional fixed effects.

Finally, in light of the impact of the serial founder effect on the duration of settlements, one could have argued that societies at greater migratory distance from Africa had shorter time to evolve and to form autocratic institutions. Thus, the negative association between the migratory distance from Africa, genetic diversity, and the extent of autocracy may reflect the shorter duration of settlements at greater migratory distance from Africa. Hence, the empirical analysis accounts for the duration of settlement and its potential effect on the emergence of autocracy.

## 3.2 Baseline Regression Specifications

This section presents the baseline econometric models that will be used in the empirical analysis of the relationship between population diversity and autocracy in the pre-colonial era.

### 3.2.1 Population Diversity and Ethnographic Characteristics

In estimating the association between observed population diversity and the pre-colonial levels of (i) jurisdictional hierarchy, (ii) social stratification, and (iii) the intensity of slavery, the following empirical specification is adopted and estimated initially via ordinary least squares (OLS):<sup>13</sup>

$$Y_i = \beta_0 + \beta_1 G_i + X_i' \beta_2 + \varepsilon_i, \quad (1)$$

where  $Y_i$  is a measure of either jurisdictional hierarchy, social stratification, or the intensity of slavery, for ethnicity  $i$ ;  $G_i$  is the observed population diversity for ethnicity  $i$ ,  $X_i$  is a vector of potentially confounding geographical characteristics for ethnicity  $i$ ; and  $\varepsilon_i$  is an error term for ethnicity  $i$ .

Moreover, considering the remarkably strong predictive power of migratory distance from East Africa for observed genetic diversity, the analysis estimates the association between predicted population diversity (in an extended sample of the entire set of ethnic groups in the *Ethnographic Atlas*) and the pre-colonial levels of (i) jurisdictional hierarchy, (ii) social stratification, and (iii) the intensity of slavery, based on the following OLS specification:<sup>14</sup>

$$Y_i = \beta_0 + \beta_1 \hat{G}_i + X_i' \beta_2 + C_i' \beta_3 + \varepsilon_i, \quad (2)$$

where  $\hat{G}_i$  is the level of population diversity predicted by migratory distance from East Africa for ethnicity  $i$ ;  $X_i$  is a vector of potentially confounding geographical characteristics for ethnicity  $i$ ;  $C_i$  is a vector of regional dummy variables for ethnicity  $i$ , and  $\varepsilon_i$  is an error term for ethnicity  $i$ .

---

<sup>13</sup>As established in Table A.7 in the Appendix, the results are robust to the use of an alternative estimation method (i.e., ordered probit).

<sup>14</sup>As established in Table A.8 in the Appendix, the results are robust to the use of an alternative estimation method (i.e., ordered probit).



Furthermore, the association between population diversity and the levels of pre-colonial jurisdictional hierarchy, social stratification, and the intensity of slavery is estimated via a 2SLS procedure, instrumenting observed population diversity in ethnicity  $i$ , with the migratory distance from East Africa to the centroid of ethnicity  $i$ . In particular, the second stage of the 2SLS regression is estimated by equation (1), while the first stage of the 2SLS regression is estimated by the equation

$$G_i = \alpha_0 + a_1 Z_i + X_i' \alpha_2 + \eta_i, \quad (3)$$

where  $Z_i$  is the migratory distance from East Africa to the centroid of the homeland of ethnicity  $i$ ,  $X_i$  is a vector of potentially confounding geographical characteristics for ethnicity  $i$ , and  $\eta_i$  is an error term for ethnicity  $i$ .

### 3.2.2 Ethnographic Characteristics and Pre-Colonial Autocracy

In estimating the association between pre-colonial autocracy and the pre-colonial levels of (i) jurisdictional hierarchy, (ii) social stratification, and (iii) the intensity of slavery, the following empirical specification is adopted and estimated via OLS:<sup>15</sup>

$$A_i = \beta_0 + \beta_1 Y_i + X_i' \beta_2 + \varepsilon_i, \quad (4)$$

where  $A_i$  is a measure of pre-colonial autocracy for ethnicity  $i$ ;  $Y_i$  is a measure of either jurisdictional hierarchy, social stratification, or the intensity of slavery, for ethnicity  $i$ ;  $X_i$  is a vector of potentially confounding geographical characteristics for ethnicity  $i$ ; and  $\varepsilon_i$  is an error term for ethnicity  $i$ .

### 3.2.3 Population Diversity and Pre-Colonial Autocracy

In estimating the association between predicted population diversity and pre-colonial autocracy, the following empirical specification is adopted and estimated via ordinary least squares OLS:<sup>16</sup>

$$A_i = \beta_0 + \beta_1 \hat{G}_i + X_i' \beta_2 + \varepsilon_i, \quad (5)$$

where  $A_i$  is a measure of pre-colonial autocracy for ethnicity  $i$ ;  $\hat{G}_i$  is the level of population diversity predicted by migratory distance from East Africa for ethnicity  $i$ ;  $X_i$  is a vector of potentially confounding geographical characteristics for ethnicity  $i$ ; and  $\varepsilon_i$  is an error term for ethnicity  $i$ .

## 4 Population Diversity and Autocracy in the Pre-Colonial Era

This section explores the association between population diversity and the degree of pre-colonial autocratic institutions across ethnic groups. Moreover, it examines the hypothesized mechanism

<sup>15</sup>As established in Table A.9 in the Appendix, the results are robust to the use of an alternative estimation method (i.e., ordered probit).

<sup>16</sup>As established in Table A.10 in the Appendix, the results are robust to the use of an alternative estimation method (i.e., ordered probit).

that may govern this reduced-form relationship. In particular, it investigates: (i) the association between population diversity and pre-colonial jurisdictional hierarchy, (ii) the association between population diversity on pre-colonial social stratification and slavery, (iii) the associations between pre-colonial jurisdictional hierarchy, social stratification and slavery, on the one hand, and pre-colonial autocracy on the other hand, and (iv) the association of population diversity on pre-colonial autocracy.

In view of the conjecture that ethnic groups characterized by higher population diversity are more likely to form institutions that would mitigate the adverse effect of non-cohesiveness on productivity, the empirical analysis first examines whether ethnic groups that are characterized by a higher level of observed population diversity tend to possess more elaborate institutions, as captured by the degree of jurisdictional hierarchy in those societies.

In light of potential concerns about the endogeneity of observed population diversity as well as the size and the representativeness of the ethnic group sample with observed population diversity, the research exploits two empirical strategies to explore the association between population diversity and autocracy and to demonstrate the robustness of the estimated association. First, migratory distance from Africa is exploited as an instrumental variable for observed population diversity, in order to examine the association between diversity and the degree of jurisdictional hierarchy, accounting for the potentially confounding effects of geographical characteristics as well as regional fixed effects. Second, using migratory distance from Africa to predict population diversity for 1,267 ethnic groups in the *Ethnographic Atlas*, the analysis explores the robustness of the result for this extended sample.

Further, in light of the second element of the proposed mechanism about the association between population diversity and social stratification, the empirical analysis explores whether ethnic groups that are characterized by a higher level of observed population diversity tend to have a higher level of class stratification and a higher intensity of slavery. Moreover, exploiting migratory distance from Africa as: (i) an instrumental variable for observed population diversity, and (ii) as a predictor of population diversity for all ethnic groups in the *Ethnographic Atlas*, the empirical analysis explores the association between population diversity on class stratification and the intensity of slavery, accounting for the potentially confounding effects of geographical characteristics as well as regional fixed effects.

Finally, the ethnic-level empirical analysis explores the contribution of population diversity to the emergence of autocratic institutions. In light of the proposed mechanism, the empirical analysis explores the association between jurisdictional hierarchy and the presence of autocratic institutions as captured by: (i) degree of absence of checks on leader's power, (ii) difficulty of removal of leaders, (iii) leader's exercise of authority, (iv) degree of lack of community decisions, (v) perception of leader's power, (vi) and indigenous autocracy. Furthermore, it examines the association between social stratification and the intensity of slavery and these measures of autocracy. Moreover, the empirical analysis explores the reduced-form association between predicted population diversity and each of these measures of autocracy.

Table 1: Observed Population Diversity and Jurisdictional Hierarchy

	Log Number of Levels of Jurisdictional Hierarchy					
	OLS					IV
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	2.565*** (0.934)	3.680*** (0.915)	3.705*** (0.956)	4.491*** (1.237)	4.367*** (1.215)	4.328*** (1.550)
Absolute Latitude		0.010*** (0.004)	0.011** (0.004)	0.010** (0.004)	0.030*** (0.006)	0.030*** (0.006)
Agricultural Suitability			0.004 (0.019)	-0.001 (0.020)	0.003 (0.019)	0.003 (0.018)
Elevation				-0.020 (0.123)	0.227 (0.144)	0.227 (0.143)
Ruggedness				0.501 (0.503)	0.075 (0.494)	0.070 (0.513)
Distance to Waterway				-0.828 (1.006)	-1.216 (1.125)	-1.213 (1.086)
Average Temperature					0.044*** (0.013)	0.044*** (0.012)
<i>N</i>	133	133	133	133	133	133
Adjusted $R^2$	0.033	0.096	0.089	0.087	0.166	0.166
1st Stage $F$ -statistic (K-P)						173.525

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

## 4.1 Population Diversity and Jurisdictional Hierarchy

This subsection explores the association between population diversity and pre-colonial jurisdictional hierarchy across ethnic groups.

### 4.1.1 Observed Population Diversity and Jurisdictional Hierarchy

The first layer of this empirical analysis establishes that ethnic groups that are characterized by a higher level of observed population diversity tend to possess more elaborate institutions, as captured by the degree of jurisdictional hierarchy in those societies.

Table 1 presents the results from OLS regression analyses of the log number of levels of jurisdictional hierarchy in the pre-colonial era on observed population diversity. Consistent with the

prediction of the proposed hypothesis, column 1 establishes a highly statistically and economically significant correlation between the measure of jurisdictional hierarchy and observed diversity, based on the 133 ethnic groups for which information on both population diversity and jurisdictional hierarchy is available. In particular, a 1 percentage point increase in observed population diversity is associated with a 2.6 percent increase in the number of levels of jurisdictional hierarchy. In light of the potentially confounding effects of geographical characteristics on this association, columns 2–5 establish that this association is robust to the gradual inclusion of control variables, capturing a range of geographical factors. In particular, the association remains highly significant while accounting for the absolute latitude of the centroid of the homeland of the ethnicity (column 2), agricultural suitability (column 3), homeland elevation, ruggedness, distance to waterways (column 4), and the average temperature (column 5).

Thus, Table 1 establishes that the estimated association between observed diversity and pre-colonial jurisdictional hierarchy is highly statistically and economically significant, accounting for a wide range of potentially confounding geographical characteristics and continental fixed effects. In particular, a 1 percentage point increase in observed population diversity is associated with a 4.4 percent increase in the number of levels of jurisdictional hierarchy. This partial association between jurisdictional hierarchy and population diversity, as derived in column 5, is plotted in Panel A of Figure B.7.

Furthermore, in light of potential concerns about the endogeneity of observed population diversity, the second layer of this empirical analysis exploits an instrumental variable strategy to explore the association between population diversity and jurisdictional hierarchy. In view of the negative association between population diversity and migratory distance from the cradle of humankind in East Africa to various settlements across the globe, migratory distance from Africa is exploited as an instrumental variable for observed population diversity, establishing a highly significant positive association between diversity and the degree of jurisdictional hierarchy.

Column 6 presents the results from 2SLS regression analyses of the log number of levels of jurisdictional hierarchy in the pre-colonial era on observed population diversity. It establishes that migratory distance from East Africa is a very strong instrument for genetic diversity (the Kleibergen-Paap  $F$ -statistic is 174) and that there is a highly statistically and economically significant association between observed diversity and the log number of levels of jurisdictional hierarchy. In particular, a 1 percentage point increase in the level of observed diversity increases the number of levels of jurisdictional hierarchy by 4.3 percent, conditional on the full set of control variables.

#### **4.1.2 Predicted Population Diversity and Jurisdictional Hierarchy**

In light of potential concerns about the endogeneity of observed population diversity and the size and the representativeness of the ethnic group sample with observed population diversity, the third layer of this empirical analysis exploits an additional empirical strategy to identify the association between population diversity and jurisdictional hierarchy and to demonstrate the robustness of the estimated association. Using migratory distance from Africa to project population diversity for

Table 2: Predicted Population Diversity and Jurisdictional Hierarchy

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	4.799*** (0.290)	5.024*** (0.325)	5.114*** (0.317)	5.096*** (0.342)	5.091*** (0.342)	4.177*** (1.115)
Absolute Latitude		0.002* (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.013*** (0.002)	0.015*** (0.003)
Agricultural Suitability			0.032*** (0.004)	0.035*** (0.004)	0.036*** (0.004)	0.038*** (0.005)
Elevation				-0.072*** (0.027)	-0.011 (0.031)	0.076** (0.031)
Ruggedness				0.149 (0.118)	0.084 (0.117)	-0.126 (0.118)
Distance to Waterway				0.630*** (0.174)	0.445** (0.175)	0.222 (0.189)
Average Temperature					0.017*** (0.005)	0.014*** (0.005)
Regional FE	No	No	No	No	No	Yes
$N$	1116	1116	1116	1116	1116	1116
Adjusted $R^2$	0.178	0.180	0.218	0.228	0.238	0.292

This table presents the results of an ethnic-group level OLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

1,267 ethnic groups in the *Ethnographic Atlas*, the analysis further establishes the robustness of the highly significant association between population diversity and jurisdictional hierarchy in this extended sample.

Table 2 presents the results from OLS regressions of the log number of levels of jurisdictional hierarchy in the pre-colonial era on predicted population diversity using the extended sample with predicted diversity. Reassuringly, in comparison to the estimates in Table 1, the estimated coefficients on predicted diversity in Table 2 are of the same order of magnitude and remain highly statistically significant. Furthermore, the estimates are very stable across specifications.

Column 1 establishes that population diversity, as predicted by migratory distance from East Africa, has a highly statistically and economically significant association with the log number of levels of jurisdictional hierarchy. In particular, a 1 percentage point increase in the level of predicted population diversity is associated with a 4.8 percent increase in the number of levels of jurisdictional hierarchy. Furthermore, columns 2–6 establish that this association is robust to the

inclusion of control variables capturing a range of geographical factors. In particular, the estimated association between predicted diversity and jurisdictional hierarchy remains highly significant while controlling for the absolute latitude of the centroid of the homeland of the ethnicity (column 2), agricultural suitability (column 3), elevation of the homelands, terrain ruggedness, distance to waterways (column 4), average temperature (column 5), and continental fixed effects (column 6).

Thus, Table 2 establishes that the estimated association between predicted population diversity and pre-colonial jurisdictional hierarchy is highly statistically and economically significant, accounting for a wide range of potentially confounding geographical characteristics and continental fixed effects. In particular, a 1 percentage point increase in predicted population diversity is associated with a 4.2 percent increase in the number of levels of jurisdictional hierarchy. This partial association between jurisdictional hierarchy and predicted population diversity, derived in column 6, is plotted in Panel B of Figure B.7.

### 4.1.3 Robustness

In light of the impact of the serial founder effect on the duration of settlements, one could have argued that societies at greater migratory distance from Africa had shorter time to evolve and to form autocratic institutions. Thus, the negative association between the migratory distance from Africa, genetic diversity, and the extent of autocracy may reflect the shorter duration of settlements at greater migratory distance from Africa. Nevertheless, as reported in Tables A.11 and A.12, although the duration of settlement has a highly significant association with the level of jurisdictional hierarchy, accounting for the duration of settlement and its potential effect on the emergence of autocracy has no qualitative effect on the established association.<sup>17</sup>

Furthermore, the findings are unaffected qualitatively by alternative geographical characteristic that has been shown to be correlated with the emergence of the state in general and the presence of autocracy in particular. First, as suggested by Fenske (2014), ecological diversity is associated with jurisdictional hierarchy. Nevertheless, accounting for the potentially confounding effect of ecological diversity does not alter the qualitative association between population diversity and jurisdictional hierarchy in the predicted as well as the observed samples (Tables A.13 and A.14). Second, the suitability of land for tubers may be associated with jurisdictional hierarchy. Accounting for the potentially confounding effect of major crop types does not alter the qualitative association between population diversity and jurisdictional hierarchy in the predicted as well as the observed samples (Tables A.15 and A.16).

The findings are robust to additional confounding geographical and ethnographic characteristics. First, variability of soil suitability that, as established by Michalopoulos (2012), contributes to ethnolinguistic fractionalization and thus population diversity has no qualitative impact on the findings (Tables A.17 and A.18). Second, while the scale of each society may be associated with jurisdictional hierarchy, it has no qualitative impact on the association between population diversity and jurisdictional hierarchy (Tables A.19 and A.20). Third, the exclusion of the African continent

---

<sup>17</sup>The Neolithic Revolution is accounted for in the country-level analysis and has no effect on the findings.

has no impact on the qualitative results (Tables A.21 and A.22). Fourth, the year of description of each ethnic group, as recorded in the *Ethnographic Atlas*, has no qualitative impact on the findings (Tables A.23 and A.24). Fifth, omitted variable statistics indicate that the results are not driven by omitted variables (Tables A.4 and A.5). Sixth, the results are robust to accounting for alternative distances (Tables A.25 and A.26). Seventh, the results are robust to accounting for the fact that the explanatory variable is a generated regressor using bootstrapping (Table A.5).

The findings are further robust to alternative estimation methods. In particular, the results are robust to accounting for spatial autocorrelation (Tables A.27 and A.28). Furthermore, the use of ordered probit rather than OLS has no impact on the results (Table A.7 and A.8), and the analysis is unaffected by the use of the number (rather than the logarithm of the number) of levels of jurisdictional hierarchy as the outcome variable (Tables A.29 and A.30).

## 4.2 Population Diversity and Social Stratification & Slavery

This subsection establishes the association between population diversity and the pre-colonial degree of social stratification and the intensity of slavery across ethnic groups.

### 4.2.1 Observed Population Diversity and Social Stratification & Slavery

The first layer of this empirical analysis establishes that ethnic groups that are characterized by a higher level of observed population diversity tend to be characterized by a higher degree of social stratification and by the presence of slavery.

Table 3 presents the results from OLS regression analyses of social stratification and slavery in the pre-colonial era on observed population diversity. Consistent with the prediction of the proposed hypothesis, columns 1–3 establish a highly statistically and economically significant correlation between the measure of social stratification and observed diversity, based on the 129 ethnic groups for which information on both population diversity and social stratification are available. Column 1 presents the correlation accounting for absolute latitude. Column 2 establishes that the association remains highly statistically significant while accounting for the baseline geographical controls. The point estimate implies that a 1 percent increase in observed population diversity is associated with a 0.06 higher score on the social stratification scale. This partial association between social stratification and population diversity is plotted in Panel A of Figure B.8.

Furthermore, columns 4–6 establish that there is a highly statistically and economically significant correlation between population diversity and the intensity of slavery, based on the 130 ethnic groups for which information on both measures is available. Column 4 presents the correlation accounting for absolute latitude. Column 5 establishes that the association remains highly statistically significant while accounting for the baseline geographical controls. The point estimate implies that a 1 percent increase in observed population diversity is associated with a 0.6 higher score on the intensity of slavery.

In light of potential concerns about the endogeneity of observed population diversity, the second layer of this empirical analysis exploits an instrumental variable strategy to identify the association

Table 3: Observed Diversity and Stratification &amp; Slavery

	Social Stratification			Intensity of Slavery		
	OLS		IV	OLS		IV
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	3.402** (1.478)	5.469*** (1.542)	3.855** (1.803)	5.365*** (1.121)	6.325*** (1.367)	5.946*** (1.670)
Absolute Latitude	0.015*** (0.005)	0.051*** (0.009)	0.050*** (0.009)	0.006 (0.005)	0.034*** (0.011)	0.034*** (0.010)
Agricultural Suitability		-0.003 (0.030)	-0.003 (0.029)		-0.024 (0.027)	-0.024 (0.026)
Elevation		0.371** (0.164)	0.384** (0.155)		0.190 (0.125)	0.193 (0.122)
Ruggedness		0.362 (0.654)	0.188 (0.659)		-0.126 (0.607)	-0.186 (0.600)
Distance to Waterway		-3.785* (2.051)	-3.579* (1.959)		-2.806** (1.381)	-2.756** (1.342)
Average Temperature		0.081*** (0.017)	0.081*** (0.017)		0.067*** (0.020)	0.067*** (0.019)
<i>N</i>	129	129	129	130	130	130
Adjusted $R^2$	0.052	0.187	0.182	0.066	0.174	0.174
1st Stage $F$ -statistic (K-P)			192.956			166.096

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of measures of pre-colonial stratification and inequality on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

between population diversity and social stratification and slavery. In view of the negative association between population diversity and migratory distance from the cradle of humankind in East Africa to various settlements across the globe, migratory distance from Africa is exploited as an instrumental variable for observed population diversity, establishing a highly significant positive association between diversity and the degree of social stratification and the intensity of slavery.

The analysis reveals that the association between genetic diversity and social stratification and the intensity of slavery is statistically significant. In particular, column 3 establishes that migratory distance from East Africa is a very strong instrument for genetic diversity (the Kleibergen-Paap  $F$ -statistic is 193) and that there is a statistically and economically significant association between observed diversity and social stratification, controlling for the baseline control variables. The point



estimate implies that a 1 percent increase in observed population diversity is associated with a 0.04 higher score on the social stratification scale.

Similarly, column 6 confirms that migratory distance from East Africa is a very strong instrument for genetic diversity (the Kleibergen-Paap  $F$ -statistic is in this case 166) and establishes that there is a highly statistically and economically significant association between observed diversity and the intensity of slavery, controlling for the baseline controls. The point estimate implies that a 1 percent increase in observed population diversity is associated with a 0.6 higher score on the intensity of slavery scale.

#### 4.2.2 Predicted Population Diversity and Stratification & Slavery

In light of potential concerns about the endogeneity of observed population diversity and the size and the representativeness of the ethnic group sample with observed population diversity, the third layer of this empirical analysis exploits an additional empirical strategy to identify the association between population diversity and social stratification and the degree of slavery and to demonstrate the robustness of the estimated association. Using migratory distance from Africa to project population diversity for 1,267 ethnic groups in the *Ethnographic Atlas*, the analysis further establishes the robustness of the highly significant association between population diversity and social stratification and the degree of slavery in this extended sample.

Table 4 presents the results from OLS regressions of social stratification and slavery in the pre-colonial era using the extended sample of predicted population diversity. Reassuringly, in comparison to the estimates in Table 3, the estimated coefficients on predicted diversity in Table 4 are largely of the same order of magnitude and remain highly statistically significant. Furthermore, the estimates are very stable across specifications. In particular, columns 1–3 establish that predicted population diversity has a highly statistically and economically significant association with the measure of social stratification. Column 1 presents the association accounting for absolute latitude. Column 2 establishes that the point estimate is very similar and remains highly significant in the presence of the baseline geographical control variables, and column 3 establishes that the point estimate is of the same order of magnitude and remain highly significant as one accounts for continental fixed effects. This partial association between social stratification and population diversity is plotted in Panel B of Figure B.8.

Similarly, columns 4 to 6 establish that predicted population diversity has a highly statistically and economically significant association with the intensity of slavery. Column 5 establishes that the point estimate is very similar and remains highly significant in the presence of the baseline controls. Furthermore, column 6 establishes that the point estimate is nearly unchanged and remains highly significant as one accounts for continental fixed effects.

Thus, Table 4 establishes that the estimated association between predicted population diversity and pre-colonial social stratification and slavery is highly statistically and economically significant, accounting for a wide range of potentially confounding geographical characteristics and continental fixed effects. In particular, a 1 percentage point increase in predicted population diversity increases

Table 4: Predicted Population Diversity and Stratification & Slavery

	Social Stratification			Intensity of Slavery		
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	5.154*** (0.562)	5.789*** (0.564)	5.379*** (2.042)	7.672*** (0.506)	7.615*** (0.515)	5.824*** (2.140)
Absolute Latitude	0.007*** (0.001)	0.018*** (0.004)	0.023*** (0.004)	0.001 (0.001)	0.011*** (0.004)	0.017*** (0.004)
Agricultural Suitability		0.033*** (0.008)	0.036*** (0.008)		0.008 (0.007)	-0.001 (0.007)
Elevation		-0.201*** (0.054)	-0.084 (0.055)		-0.035 (0.039)	-0.158*** (0.051)
Ruggedness		1.690*** (0.210)	1.326*** (0.231)		0.595*** (0.224)	1.021*** (0.247)
Distance to Waterway		-0.250 (0.351)	-0.289 (0.395)		0.568* (0.310)	0.099 (0.309)
Average Temperature		0.023*** (0.008)	0.021*** (0.007)		0.024*** (0.007)	0.020*** (0.007)
Regional FE	No	No	Yes	No	No	Yes
$N$	1074	1073	1073	1083	1082	1082
Adjusted $R^2$	0.065	0.156	0.188	0.189	0.209	0.292

This table presents the results of an ethnic-group level OLS regression analysis of measures of pre-colonial stratification and inequality on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

the score on the social stratification index by 0.06 (column 2) and the intensity of slavery index by 0.08 (column 5).

### 4.3 Jurisdictional Hierarchy, Stratification, and Pre-Colonial Autocracy

This part of the pre-colonial ethnic-level empirical analysis explores the association between the emergence of intuitions as well as the degree of social stratification with the emergence of pre-colonial autocratic institutions. In line with the proposed mechanism, the empirical analysis establishes that the number of levels of jurisdictional hierarchy and the degree of social stratification and slavery are associated with the presence of pre-colonial autocratic institutions, as captured by: (i) degree of absence of checks on leader's power, (ii) difficulty of removal of leaders, (iii) leader's exercise of authority, (iv) degree of lack of community decisions, (v) perception of leader's power,

and (vi) a measure of indigenous autocracy based on the rules of succession to the office of the local headman. Moreover, the empirical analysis establishes the reduced form association between predicted population diversity and each of these measures of pre-colonial autocracy.

#### **4.3.1 Jurisdictional Hierarchy and Pre-Colonial Autocratic Institutions**

First, the empirical analysis establishes that the number of levels of jurisdictional hierarchy is associated with the presence of pre-colonial autocratic institutions at the community level.

Table 5 presents the results from OLS regression analyses of the various measures of pre-colonial autocratic institutions on jurisdictional hierarchy.<sup>18</sup> Columns 1 and 2 establish that pre-colonial jurisdictional hierarchy is highly significantly associated with the degree of absence of checks on the leader's power. Columns 3 and 4 suggest that pre-colonial jurisdictional hierarchy is highly significantly associated with the difficulty of removal of leaders. Columns 5 and 6 find that pre-colonial jurisdictional hierarchy is highly significantly associated with the leader's exercise of authority. Columns 7 and 8 establish that pre-colonial jurisdictional hierarchy is highly significantly associated with the degree of lack of community decisions. Columns 9 and 10 suggest that pre-colonial jurisdictional hierarchy is highly significantly associated with the perception of the leader's power. Finally, columns 11 and 12 establish that pre-colonial jurisdictional hierarchy is highly significantly associated with the level of indigenous autocracy.

Thus, Table 5 establishes that jurisdictional hierarchy, as reflected by the number of levels of jurisdictional hierarchy, is highly statistically significantly associated with the various measures of pre-colonial autocratic institutions.

#### **4.3.2 Stratification and Pre-Colonial Autocratic Institutions**

Second, the empirical analysis establishes that the degree of social stratification and slavery is associated with presence of pre-colonial autocratic institutions at the community level. As presented in Table 6, social stratification is significantly associated with the degree of absence of checks on the leader's power (columns 1 and 2), and highly significantly associated with the difficulty of removal of leaders (columns 3 and 4); the leader's exercise of authority (columns 5 and 6); the degree of lack of community decisions (columns 7 and 8); the perception of the leader's power (columns 9 and 10); and indigenous autocracy (columns 11 and 12). Similarly, as presented in Table 7, the intensity of slavery is highly significantly associated with: the degree of absence of checks on the leader's power (columns 1 and 2); the difficulty of removal of leaders (columns 3 and 4); the leader's exercise of authority (columns 5 and 6); the degree of lack of community decisions (columns 7 and 8); the perception of the leader's power (columns 9 and 10); and indigenous autocracy (columns 11 and 12). Thus, Table 6 and Table 7 establish that the associations between social stratification, the intensity of slavery, and various measures of pre-colonial autocratic institutions are mostly highly statistically significant.

---

<sup>18</sup>Given the limited number of observations in the SCCS-dataset, the analysis cannot account for continental fixed effects.

Table 5: Jurisdictional Hierarchy and Autocratic Institutions

	Degree of Absence of Checks on Leader's Power		Difficulty of Removal of Leaders		Leader's Exercise of Authority		Degree of Lack of Community Decisions		Perception of Leader's Power		Indigenous Autocracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log Levels of Jurisdictional Hierarchy	0.864*** (0.149)	0.886*** (0.153)	0.942*** (0.180)	0.841*** (0.162)	0.894*** (0.130)	0.907*** (0.138)	0.850*** (0.136)	0.843*** (0.142)	0.899*** (0.136)	0.942*** (0.138)	0.121*** (0.034)	0.089** (0.035)
Absolute Latitude	-0.004 (0.004)	-0.000 (0.012)	-0.005 (0.005)	-0.002 (0.015)	-0.007* (0.004)	-0.001 (0.013)	-0.006* (0.004)	-0.020 (0.012)	-0.004 (0.004)	0.003 (0.012)	-0.004*** (0.001)	0.002 (0.003)
Agricultural Suitability		-0.062*** (0.021)		-0.016 (0.031)		-0.033 (0.022)		-0.029 (0.019)		-0.024 (0.022)		0.018*** (0.005)
Elevation		0.070 (0.160)		-0.167 (0.210)		0.011 (0.124)		-0.186 (0.129)		0.024 (0.131)		-0.018 (0.030)
Ruggedness		0.670 (0.597)		2.078*** (0.738)		0.557 (0.551)		0.985* (0.561)		-0.299 (0.518)		-0.035 (0.145)
Distance to Waterway		-1.025*** (0.372)		-0.412 (0.770)		-1.409*** (0.479)		0.114 (0.472)		-1.411*** (0.474)		0.047 (0.239)
Average Temperature		0.024 (0.023)		0.019 (0.029)		0.021 (0.022)		-0.020 (0.023)		0.018 (0.024)		0.009* (0.005)
<i>N</i>	82	82	74	74	83	83	86	86	86	86	855	855
Adjusted <i>R</i> <sup>2</sup>	0.293	0.336	0.256	0.290	0.354	0.365	0.338	0.340	0.336	0.334	0.040	0.055

This table presents the results of an ethnic-group level OLS regression analysis of measures of pre-colonial autocracy on a measure of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy), conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table 6: Social Strata and Autocratic Institutions

	Degree of Absence of Checks on Leader's Power		Difficulty of Removal of Leaders		Leader's Exercise of Authority		Degree of Lack of Community Decisions		Perception of Leader's Power		Indigenous Autocracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Social Stratification	0.539*** (0.091)	0.569*** (0.092)	0.559*** (0.114)	0.475*** (0.132)	0.470*** (0.085)	0.499*** (0.090)	0.510*** (0.083)	0.507*** (0.094)	0.476*** (0.094)	0.531*** (0.099)	0.095*** (0.019)	0.083*** (0.020)
Absolute Latitude	-0.008* (0.004)	-0.004 (0.013)	-0.008 (0.005)	-0.004 (0.016)	-0.010** (0.004)	-0.007 (0.013)	-0.010** (0.004)	-0.024* (0.012)	-0.008* (0.004)	-0.001 (0.012)	-0.005*** (0.001)	0.002 (0.003)
Agricultural Suitability		-0.054* (0.027)		0.004 (0.037)		-0.025 (0.028)		-0.014 (0.029)		-0.007 (0.032)		0.018*** (0.005)
Elevation		0.215 (0.159)		-0.053 (0.209)		0.177 (0.140)		-0.093 (0.133)		0.124 (0.129)		0.002 (0.030)
Ruggedness		-0.152 (0.591)		1.584* (0.915)		-0.198 (0.627)		0.382 (0.662)		-0.903 (0.586)		-0.139 (0.146)
Distance to Waterway		-1.315*** (0.456)		-0.478 (0.912)		-1.628** (0.650)		-0.004 (0.613)		-1.508** (0.664)		0.181 (0.256)
Average Temperature		0.018 (0.024)		0.014 (0.030)		0.012 (0.023)		-0.025 (0.023)		0.013 (0.024)		0.010** (0.005)
<i>N</i>	83	83	74	74	84	84	87	87	87	87	827	827
Adjusted <i>R</i> <sup>2</sup>	0.315	0.348	0.246	0.246	0.270	0.272	0.335	0.313	0.262	0.261	0.053	0.068

This table presents the results of an ethnic-group level OLS regression analysis of measures of pre-colonial autocracy on a measure of pre-colonial social stratification, conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table 7: Slavery and Autocratic Institutions

	Degree of Absence of Checks on Leader's Power		Difficulty of Removal of Leaders		Leader's Exercise of Authority		Degree of Lack of Community Decisions		Perception of Leader's Power		Indigenous Autocracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Intensity of Slavery	0.259** (0.115)	0.276** (0.121)	0.441*** (0.123)	0.414*** (0.139)	0.260** (0.111)	0.268** (0.121)	0.325*** (0.105)	0.331*** (0.113)	0.275** (0.106)	0.303*** (0.114)	0.115*** (0.021)	0.110*** (0.022)
Absolute Latitude	-0.008* (0.004)	0.001 (0.014)	-0.007 (0.005)	-0.001 (0.016)	-0.010** (0.004)	-0.003 (0.014)	-0.009** (0.004)	-0.021 (0.013)	-0.007 (0.004)	0.001 (0.014)	-0.004*** (0.001)	0.003 (0.003)
Agricultural Suitability		-0.049 (0.030)		0.005 (0.041)		-0.022 (0.031)		-0.014 (0.030)		-0.007 (0.032)		0.017*** (0.005)
Elevation		0.016 (0.185)		-0.320 (0.230)		0.007 (0.171)		-0.270* (0.161)		-0.054 (0.155)		-0.015 (0.030)
Ruggedness		0.347 (0.692)		2.189** (0.994)		0.362 (0.747)		0.910 (0.780)		-0.210 (0.718)		-0.096 (0.146)
Distance to Waterway		-0.990** (0.425)		-0.271 (0.712)		-1.327** (0.543)		0.238 (0.421)		-1.207** (0.478)		-0.027 (0.238)
Average Temperature		0.029 (0.026)		0.018 (0.031)		0.019 (0.026)		-0.021 (0.023)		0.017 (0.026)		0.011** (0.005)
<i>N</i>	82	82	73	73	83	83	86	86	86	86	832	832
Adjusted $R^2$	0.080	0.077	0.146	0.178	0.101	0.078	0.149	0.133	0.082	0.061	0.062	0.078

This table presents the results of an ethnic-group level OLS regression analysis of measures of pre-colonial autocracy on a measure of the intensity of slavery, conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

### 4.3.3 Predicted diversity and Pre-Colonial Autocracy

Third, the empirical analysis establishes that predicted population diversity has a positive association with the presence of pre-colonial autocratic institutions at the community level. Given the limited number of observations in the SCCS-dataset, the analysis uses predicted diversity, rather than observed diversity.

As reported in Table 8, predicted population diversity has a significant association with the degree of absence of checks on the leader’s power (columns 1 and 2); the difficulty of removal of leaders (columns 3 and 4); the leader’s exercise of authority (columns 5 and 6); the degree of lack of community decisions (columns 7 and 8); the perception of the leader’s power (columns 9 and 10); and indigenous autocracy (columns 11 and 12).<sup>19</sup> Overall, Table 8 establishes that predicted population diversity has significant association with the various measures of pre-colonial autocratic institutions at the community level.

As suggested by Bentzen et al. (2017) irrigation suitability, and its potential effect on the emergence of a landed elite, is associated with the presence of autocracy across contemporary countries and regions. However, as reported in Table A.31, irrigation suitability is not associated with measures of autocracy across ethnic groups. Moreover, accounting for the potentially confounding effect of irrigation suitability does not alter qualitatively association between population diversity and autocracy.

Furthermore, the results are robust to accounting for the time since settlement (Table A.32), the approximate year of description as reported in the *Ethnographic Atlas* (Table A.33), and alternative distances (Table A.34).

Hence, the findings establish the association between predicted population diversity and the level of autocratic institutions, while indicating that this association could have plausibly operated through the association between population diversity and the formation of institutions as well as stratification.

## 5 Roots of Autocracy in the Modern Era

This section explores the determinants of the nature of national institutions.<sup>20</sup> It examines the importance of the association between population diversity and pre-colonial autocratic institutions across ethnic groups for the understanding of the contemporary variation in autocratic institutions across nations. In particular, it examines the persistence of ethnic institutions, that were formed in the pre-colonial era, and their association with contemporary national institutions. Moreover, it analyzes the direct association between population diversity and both pre-colonial and contemporary national institutions.

---

<sup>19</sup>It should be noted that the association between predicted population diversity and pre-colonial autocracy is not quadratic. In particular, if predicted diversity squared is included to the baseline specifications, the estimated coefficient of this square term is insignificant in all specifications.

<sup>20</sup>The overall negative association between autocratic institutions and economic development is reflected in Figure B.4 in the Appendix.

Table 8: Predicted Diversity and Autocratic Institutions

	Degree of Absence of Checks on Leader's Power		Difficulty of Removal of Leaders		Leader's Exercise of Authority		Degree of Lack of Community Decisions		Perception of Leader's Power		Indigenous Autocracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Predicted Genetic Diversity	4.611** (1.847)	4.217** (1.829)	5.590*** (1.998)	5.190** (2.002)	5.029*** (1.730)	5.196*** (1.780)	6.208*** (1.373)	6.009*** (1.470)	4.953*** (1.583)	5.737*** (1.745)	1.100*** (0.404)	1.301*** (0.410)
Absolute Latitude	-0.006 (0.005)	0.000 (0.014)	-0.006 (0.006)	0.004 (0.016)	-0.009** (0.004)	-0.002 (0.014)	-0.008* (0.004)	-0.019 (0.012)	-0.006 (0.004)	0.004 (0.013)	-0.004*** (0.001)	0.004 (0.003)
Agricultural Suitability		-0.044* (0.025)		0.013 (0.034)		-0.020 (0.025)		-0.010 (0.024)		-0.003 (0.023)		0.020*** (0.005)
Elevation		0.047 (0.190)		-0.276 (0.234)		0.032 (0.180)		-0.235 (0.160)		-0.025 (0.158)		-0.020 (0.030)
Ruggedness		0.913 (0.707)		2.698*** (0.883)		0.644 (0.658)		1.161* (0.651)		-0.052 (0.607)		-0.066 (0.144)
Distance to Waterway		-1.057** (0.442)		-0.426 (0.779)		-1.584*** (0.526)		-0.078 (0.460)		-1.524*** (0.492)		0.028 (0.242)
Average Temperature		0.026 (0.028)		0.027 (0.032)		0.020 (0.027)		-0.019 (0.023)		0.020 (0.026)		0.011** (0.005)
<i>N</i>	83	83	74	74	84	84	87	87	87	87	898	898
Adjusted <i>R</i> <sup>2</sup>	0.075	0.091	0.079	0.156	0.114	0.113	0.175	0.168	0.095	0.084	0.039	0.062

This table presents the results of an ethnic-group level OLS regression analysis of measures of pre-colonial autocracy on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.



## 5.1 Baseline Regression Specifications

### 5.1.1 Persistence of Autocracy

In estimating the persistence of institutions from the pre-colonial to the modern era, the following empirical specification is adopted and estimated via ordinary least squares OLS:

$$A_{m,i} = \beta_0 + \beta_1 A_{p,i} + \beta_2 G_{m,i} + X'_{m,i} \beta_3 + H'_i \beta_4 + \varepsilon_i, \quad (6)$$

where  $A_{m,i}$  is the level of modern autocracy for country  $i$ ;  $A_{p,i}$  is the level of pre-colonial autocracy for country  $i$ ;  $G_{m,i}$  is the predicted level of ancestry-adjusted genetic diversity for country  $i$ ;  $X_i$  is a vector of geographical characteristics for country  $i$ ;  $H_i$  is a vector of non-geographical control variables for country  $i$ ;  $C_i$  is a vector of regional dummy variables for country  $i$ ; and  $\varepsilon_i$  is a country-specific error term.

### 5.1.2 Population Diversity and Modern Autocracy

In estimating the association between population diversity and contemporary institutions, the following empirical specification is adopted and estimated via OLS:

$$A_{m,i} = \beta_0 + \beta_1 G_{m,i} + X'_{m,i} \beta_2 + H'_i \beta_3 + \varepsilon_i. \quad (7)$$

Hence, this specification captures the overall association between population diversity,  $\beta_1$  and the level of autocracy. This level is the sum of the direct association between population diversity and the level of autocracy (i.e.,  $\beta_2$  in equation (6)) as well as the indirect one via the persistence of pre-colonial institutions on contemporary ones (i.e.,  $\beta_1$  in equation (6)) multiplied by the association between population diversity and indigenous autocracy.

## 5.2 Persistence of Autocracy

The empirical analysis of the determinants of modern institutions establishes initially the importance of the association between population diversity and pre-colonial autocratic institutions across ethnic groups for the understanding of the contemporary variation in autocratic institutions across nations, accounting for a large number of possibly confounding geographical characteristics, regional fixed effects, colonial history (i.e., duration and colonizer nation), legal origins, pre-colonial development and the degree of ethnolinguistic fractionalization and its potential geographical origins. In particular, it suggests that ethnic institutions that were formed in the pre-colonial era persisted over time and are associated with contemporary national institutions.

Aggregating the level of pre-colonial ethnic autocracy into the level of national pre-colonial autocracy, based on the weighted average of the level of autocracy in the ancestral population of modern countries (following the methodology of Giuliano and Nunn (2016)), the analysis suggests that indeed the pre-colonial level of autocracy has contributed to the contemporary level of autoc-

racy, beyond the persistent effects of geographical and human characteristics.<sup>21</sup> In particular, the levels of autocracy and the absence of executive constraints in the contemporary period are positively and significantly associated with the degree of autocratic institutions in the pre-colonial era, accounting for potentially confounding effects of geographical characteristics as well as population diversity.

The presence of autocratic institutions in the pre-colonial era could be captured by either the variable “Succession to the Office of Local Headman” in the *Ethnographic Atlas* or the variable “degree of absence of checks on leader’s power” in the *SCCS*. Given the limited number of observations of the latter variable, the baseline analysis will be based on the former one.<sup>22</sup>

The degree of contemporary autocratic institutions is captured by the indexes of “Constraints on the Executive” and “Autocracy” as reported by the *Polity IV Project* dataset. Table 9 establishes the presence of institutional persistence in the sample of countries with information on both indigenous and modern autocracy. Column 1 establishes that, unconditionally, the level of indigenous autocracy is negatively and highly significantly associated with the executive constraints in the modern period. Column 2 establishes that the association is robust to controlling for the baseline control variables for the country-level analysis. Furthermore, column 3 establishes that institutional persistence remains significant at the 10% significance level when also controlling for predicted diversity. That column also establishes that predicted diversity has a significant association with executive constraints, controlling for the baseline control variables as well as indigenous autocracy. This latter result foreshadows the findings of the next section, which establish the association between genetic diversity and executive constraints in modern countries.

Similarly, column 4 establishes a highly significant negative unconditional association between the indigenous autocracy and the index of autocracy for the modern period. Moreover, column 5 suggests that the association established in column 4 is robust to controlling for the baseline control variables. Furthermore, column 6 establishes that institutional persistence remains significant at the 10% significance level when also controlling for predicted diversity. That column also establishes that predicted diversity has a significant association with executive constraints, controlling for the baseline control variables as well as indigenous autocracy. This latter result also foreshadows the findings of the next section, which establish the association between genetic diversity and the level of autocracy in modern countries.

### 5.3 Contemporary Population Diversity and Modern Autocracy

The empirical analysis of the determinants of modern institutions further explores the contribution of modern population diversity to autocratic national institutions in the modern era. It examines whether population diversity at the national level, as captured by predicted population diversity,

---

<sup>21</sup>Table A.35 in the Appendix establishes the persistence of autocratic institutions on the purely geographical level, i.e. using an alternative aggregation method of pre-colonial institutions based on the precolonial level autocracy of ethnic groups that resided within the borders of the modern countries.

<sup>22</sup>As established in Table A.35 in the Appendix, these findings are robust to using the degree of absence of checks on leader’s power, rather than indigenous autocracy.

Table 9: Persistence of Autocracy

	Log Executive Constraints			Log Autocracy		
	(1)	(2)	(3)	(4)	(5)	(6)
Indigenous Autocracy	-0.255*** (0.071)	-0.195** (0.087)	-0.170* (0.090)	0.438*** (0.148)	0.395** (0.178)	0.343* (0.185)
Predicted Genetic Diversity			-3.408** (1.408)			6.926** (2.809)
Absolute Latitude		0.008* (0.004)	0.008** (0.004)		-0.011 (0.009)	-0.012 (0.009)
Agricultural Suitability		0.026** (0.011)	0.027*** (0.010)		-0.041* (0.022)	-0.042** (0.021)
Elevation		0.082 (0.096)	0.062 (0.099)		-0.055 (0.181)	-0.015 (0.186)
Ruggedness		0.000 (0.000)	0.000 (0.000)		-0.000 (0.001)	-0.000 (0.001)
Distance to Waterway		0.000 (0.003)	0.000 (0.003)		-0.002 (0.005)	-0.002 (0.005)
Temperature		-0.001 (0.009)	-0.004 (0.010)		0.007 (0.019)	0.013 (0.019)
Colony		0.205 (0.157)	0.239 (0.148)		-0.188 (0.285)	-0.257 (0.265)
Legal Origin FE	No	Yes	Yes	No	Yes	Yes
Regional FE	No	Yes	Yes	No	Yes	Yes
$N$	153	153	153	153	153	153
Adjusted $R^2$	0.072	0.435	0.458	0.047	0.455	0.477

This table presents the results of an ethnic-group level OLS regression analysis of measures of contemporary autocracy on a measure of pre-colonial autocracy, conditional on a range of geographical control variables as well as predicted diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

has a significant association with the degree of autocracy and the absence of executive constraints across countries, accounting for a large number of potentially confounding geographical characteristics, regional fixed effects, colonial history (i.e., duration and colonizer nation), legal origins, pre-colonial development and the degree of ethnolinguistic fractionalization and its potential geographical origins.

The country-level analysis employs the measure of genetic diversity, as constructed by Ashraf and Galor (2013), accounting for three important elements of population diversity with a national population: the proportional representation of each ethnic group within the country, the expected heterozygosity within each subnational group, as well as the diversity that arises from the genetic distances between the pre-colonial ancestral populations.

### 5.3.1 Population Diversity and Constraint on the Executive

This subsection establishes that consistent with the proposed hypothesis population diversity at the national level has a highly significant negative association with the degree of executive constraints, accounting for a large number of confounding factors.

As reported in Table 10, column 1 establishes based on data from 155 countries that, unconditionally, the level of predicted diversity within a country in the modern era has a highly significant negative association with the constraint on the chief executive.<sup>23</sup> The estimated association indicates that a 1 percentage point increase in predicted diversity is associated with a 3.7 percent decrease in the average level of the “Constraint on the Chief Executive” over the period 1994–2013. Columns 2–4 establish that the association remains highly significant once additional confounding geographical characteristics are accounted for. In addition, column 5 indicates that the negative association of predicted population diversity remains highly significant while accounting for continental fixed effects, capturing unobserved heterogeneity across continents. Reassuringly, as reported in columns 6 and 7, the association of genetic diversity on contemporary executive constraints is unaffected by colonial history and legal origins fixed effects. This partial association between constraints on the executive and population diversity, as derived in column 7, is plotted in Figure B.9.

In light of the potential effect of autocracy on population diversity via the cross-continental migration in the post-1500 era, the empirical analysis exploits an instrumental variable strategy to further identify the association between population diversity and constraints on the executive. In view of the negative association between population diversity and migratory distance from the cradle of humankind in East Africa to various settlements across the globe, migratory distance from Africa is exploited as an instrumental variable for population diversity.

The association between population diversity and the constraint on the executive is established in column 8, which presents the estimation results from 2SLS regression analyses instrumenting

---

<sup>23</sup>It should be noted that the association between predicted population diversity and autocracy in the modern era is not quadratic. In particular, if predicted diversity squared is included to the baseline specifications, the estimated coefficient of this square term is insignificant in all specifications.

population diversity by the migratory distance from East Africa. The column establishes that the level of population diversity has a highly significant association with the constraint on the chief executive in 1994–2013, controlling for the baseline controls.

Furthermore, Table A.36 in the Appendix establishes that the findings in Table 10 are robust to focusing on constraint on the executive in 2013. Moreover, accounting for the potentially confounding effect of irrigation suitability does not alter qualitatively the association between population diversity and autocracy (Table A.37). In addition, the findings are robust to the use of additional geographical controls, such as the percentage of land near a waterway (Table A.38), inequality of land suitability (Table A.39), and percentages of population living in various climate zones (Table A.40). Furthermore, they are robust to the inclusion of additional measures of colonial history, such as colonizer nation (Table A.41). Moreover, the findings are robust to the inclusion of arguably endogenous controls, such as income per capita (Table A.42), years of schooling (Table A.43), population density in 1500 (Table A.44), and social infrastructure (Table A.45). Finally, the use of ethnolinguistic fractionalization as an alternative measure of population diversity suggests that, while fractionalization has no association with executive constraints, the association with genetic diversity remains nearly intact (Table A.46).<sup>24</sup>

Moreover, in view of the potential association between population diversity and the onset of the Neolithic Revolution, one could have argued that the emergence of sedentary communities in the course of the Neolithic Revolution, rather than the dual effect of diversity, contributed to the onset of autocracy. However, as established in Tables A.47, accounting for the time elapsed since the Neolithic Revolution has no qualitative effect on the association between population diversity and constraint on the executive.<sup>25</sup>

Finally, the negative association between population diversity and the constraint on the executive is robust to the inclusion of yearly data since 1830.<sup>26</sup> In particular, as depicted in Figure B.10 in the Appendix, this yearly association between population diversity and constraints on the executive is negative for 97% of the years in the Polity IV data since 1830, and is increasingly more significant statistically as the number of observations increases.<sup>27</sup>

---

<sup>24</sup>While the positive association between fractionalization and autocracy is significant in the absence of geographical controls, consistent with Aghion et al. (2004), once geographical controls are introduced only genetic diversity remains significant.

<sup>25</sup>Since data on the time elapsed since the Neolithic Revolution is not available at the ethnic group level, one cannot establish directly the robustness of the analysis the Neolithic Revolution. However, since the duration of settlements is plausibly correlated with the onset of the Neolithic Revolution, this potential effect is accounted for.

<sup>26</sup>The data contains less than 30 observations per year for the pre-1830 period, rendering estimations potentially unreliable given the number of control variables and fixed effects that is accounted for in the regressions.

<sup>27</sup>The analysis focuses on data for the modern period since the historical data is available only for a small and selected group of countries whose institutions were sufficiently growth promoting so as to be included in the sample. In particular, those countries are generally developed countries with lower levels of autocracy today. Furthermore, the ethnic-level analysis captures already some of this early period since the description of ethnic groups in the sample is primarily based on their characteristics around the turn to the 20th century.

Table 10: Predicted Population Diversity and Constraint on the Executive (1994–2013)

	Log Constraint on Chief Executive							
	OLS						IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-3.666*** (0.814)	-4.136*** (0.831)	-3.165*** (0.867)	-3.165*** (0.867)	-3.375*** (0.824)	-3.353*** (0.831)	-3.663*** (1.294)	-5.775*** (2.236)
Absolute Latitude		0.007*** (0.001)	0.007*** (0.002)	0.007*** (0.002)	0.012*** (0.002)	0.014*** (0.003)	0.009*** (0.003)	0.010*** (0.003)
Agricultural Suitability			0.039*** (0.009)	0.039*** (0.009)	0.050*** (0.010)	0.051*** (0.010)	0.030*** (0.010)	0.030*** (0.010)
Elevation			-0.084 (0.072)	-0.084 (0.072)	-0.021 (0.081)	-0.003 (0.080)	0.056 (0.091)	0.051 (0.087)
Ruggedness			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance to Waterway			0.003 (0.002)	0.003 (0.002)	0.002 (0.002)	0.002 (0.002)	0.001 (0.003)	0.001 (0.003)
Colony						0.103 (0.123)	0.202 (0.140)	0.212 (0.129)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	155	155	155	155	155	155	155	155
Adjusted $R^2$	0.054	0.140	0.243	0.243	0.337	0.337	0.447	0.436
1st Stage $F$ -statistic (K-P)								55.825

This table presents the results of a country-level OLS and 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

### 5.3.2 Population Diversity and Autocracy

This subsection establishes that consistent with the proposed hypothesis population diversity at the national level has a highly significant negative association with the index of autocracy, accounting for a large number of confounding geographical characteristics, regional fixed effects, colonial history, legal origins, pre-colonial development and the degree of ethnolinguistic fractionalization and its potential geographical origins. Moreover, the association remains nearly intact if one accounts for arguably endogenous controls such as income per capita and education.

As reported in Table 11, column 1 establishes based on data from 155 countries that, unconditionally, the level of predicted diversity within a country in the modern era has a highly significant negative association with the level of autocracy. The estimated association indicates that a 1 percentage point increase in predicted diversity is associated with a 7.7 percent increase in the average level of the autocracy measure for the period 1994–2013. Columns 2–4 establish that the association remains highly significant once additional confounding geographical characteristics are accounted for. In addition, column 5 indicates that the negative association of predicted population diversity remains highly significant while accounting for continental fixed effects, capturing unobserved heterogeneity across continents. Reassuringly, as reported in columns 6 and 7, the association between genetic diversity and the contemporary level of autocracy is unaffected by colonial history and legal origins fixed effects. This partial association between autocracy and population diversity, as derived in column 7, is plotted in Figure B.11.

The association between population diversity and autocracy is further established in column 8, which presents the estimation results from a 2SLS regression analysis instrumenting predicted diversity by the migratory distance from East Africa. The column establishes that the level of predicted diversity has a highly significant association with autocracy in 1994–2013, controlling for potentially confounding geographical characteristics, continental fixed effects, and legal origins fixed effects.

Thus, the second layer of the empirical analysis of the determinants of contemporary institutions suggests that the spatial distribution of population diversity across the globe has also contributed to contemporary variation in the degree of autocracy across countries. This reduced-form association between population diversity and the prevalence of contemporary autocratic institutions across nations may reflect either persistence of institutions from the pre-colonial to the modern era, as established in the first layer of the analysis, or a direct association between population diversity and contemporary autocratic institutions, capturing the association between diversity and the demand for institutions as well as the scope for domination.

Furthermore, Table A.48 in the Appendix establishes that the findings in Table 11 are robust to focusing on the level of autocracy in 2013. Moreover, accounting for the potentially confounding effect of irrigation suitability does not alter qualitatively the association between population diversity and autocracy (Table A.49). In addition, the findings are robust to the use of additional geographical controls, such as the percentage of land near a waterway (Table A.50), inequality of land suitability (Table A.51), and percentages of population living in various climate zones (Table

Table 11: Predicted Population Diversity and Autocracy (1994–2013)

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	7.742*** (1.837)	8.564*** (1.876)	6.724*** (1.951)	6.724*** (1.951)	7.039*** (1.828)	7.028*** (1.842)	7.240*** (2.675)	9.147** (4.652)
Absolute Latitude		-0.013*** (0.003)	-0.013*** (0.003)	-0.013*** (0.003)	-0.022*** (0.004)	-0.023*** (0.007)	-0.015** (0.007)	-0.016** (0.007)
Agricultural Suitability			-0.079*** (0.019)	-0.079*** (0.019)	-0.103*** (0.020)	-0.103*** (0.021)	-0.047** (0.021)	-0.047** (0.020)
Elevation			0.190 (0.156)	0.190 (0.156)	0.048 (0.165)	0.039 (0.165)	-0.024 (0.172)	-0.020 (0.162)
Ruggedness			-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Distance to Waterway			-0.008** (0.004)	-0.008** (0.004)	-0.007 (0.004)	-0.007 (0.005)	-0.004 (0.004)	-0.004 (0.004)
Colony						-0.050 (0.248)	-0.139 (0.254)	-0.148 (0.235)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	155	155	155	155	155	155	155	155
Adjusted <i>R</i> <sup>2</sup>	0.056	0.115	0.216	0.216	0.313	0.309	0.464	0.462
1st Stage <i>F</i> -statistic (K-P)								55.825

This table presents the results of a country-level OLS and 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.



A.52). Furthermore, they are robust to the inclusion of additional measures of colonial history, such as colonizer nation (Table A.53). Moreover, the findings are robust to the inclusion of arguably endogenous controls, such as income per capita (Table A.54), years of schooling (Table A.55), population density in 1500 (Table A.56), and social infrastructure (Table A.57). Finally, the use of ethnolinguistic fractionalization as an alternative measure of population diversity suggests that, while fractionalization has no association with autocracy, the association with genetic diversity remains nearly intact (Table A.58).<sup>28</sup> Furthermore, it should be noted that Table A.60 in the Appendix establishes that the findings in Table 10 and 11 are robust to focusing on democracy, rather than autocracy, as the outcome variable. Finally, it should be noted that the results are robust to the use of alternative outcome measures (Tables A.61–62).

Moreover, in view of the potential association between population diversity and the onset of the Neolithic Revolution, one could have argued that the surplus that was generated in the course of the Neolithic Revolution, rather than the dual effect of diversity, contributed to the onset of autocracy. However, as established in Tables A.59, accounting for the time elapsed since the Neolithic Revolution has no qualitative effect on the association between population diversity and autocracy.

## 6 Conclusion

This research explores the origins of the variation in the prevalence and nature of political institutions across the globe. It advances the hypothesis and establishes empirically that diversity across human societies, as determined in the course of the exodus of *Homo sapiens* from Africa tens of thousands of years ago, contributed to the formation of autocratic institutions across societies. The study suggests that while population diversity has amplified the beneficial effects of institutions in mitigating the adverse effects of non-cohesiveness on productivity, the contribution of diversity to the range of cognitive and physical traits has fostered the scope for domination, leading to the formation and persistence of institutions of the autocratic type.

The analysis suggests that diversity contributed to the emergence of autocratic pre-colonial institutions. Moreover, the findings indicate that the impact of diversity on these institutions has plausibly operated through its dual effect on the formation of institutions as well as class stratification. Furthermore, reflecting the persistence of institutional, cultural, and human characteristics, the study suggests that the spatial distribution of population diversity across the globe has also contributed to contemporary variation in the degree of autocracy across countries.

---

<sup>28</sup>While the positive association between fractionalization and autocracy is significant in the absence of geographical controls, consistent with Aghion et al. (2004), once geographical controls are introduced only genetic diversity remains significant.

## References

- ACEMOGLU, D., S. JOHNSON, AND J. A. ROBINSON (2001): “The Colonial Origins of Comparative Development: An Empirical Investigation,” *American Economic Review*, 91, 1369–1401.
- AGHION, P., A. ALESINA, AND F. TREBBI (2004): “Endogenous Political Institutions,” *The Quarterly Journal of Economics*, 565–611.
- ALESINA, A. AND P. GIULIANO (2015): “Culture and institutions,” *Journal of Economic Literature*, 53, 898–944.
- ALLEN, R. C. (1997): “Agriculture and the Origins of the State in Ancient Egypt,” *Explorations in Economic History*, 34, 135–154.
- ALTONJI, J. G., T. E. ELDER, AND C. R. TABER (2005): “Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools,” *Journal of political economy*, 113, 151–184.
- ARNOLD, J. E. (1993): “Labor and the rise of complex hunter-gatherers,” *Journal of Anthropological Archaeology*, 12, 75–119.
- ASHRAF, Q. AND O. GALOR (2013): “The “Out of Africa” Hypothesis, Human Genetic Diversity, and Comparative Economic Development,” *American Economic Review*, 103, 1–46.
- ATKINSON, Q. D. (2011): “Phonemic diversity supports a serial founder effect model of language expansion from Africa,” *Science*, 332, 346–349.
- BELFER-COHEN, A. AND O. BAR-YOSEF (2002): “Early Sedentism in the Near East,” in *Life in Neolithic Farming Communities*, ed. by I. Kuijt, Springer US, Fundamental Issues in Archaeology, 19–38.
- BELLOWS, J. AND E. MIGUEL (2009): “War and local collective action in Sierra Leone,” *Journal of public Economics*, 93, 1144–1157.
- BENTZEN, J. S., N. KAARSEN, AND A. M. WINGENDER (2017): “Irrigation and autocracy,” *Journal of the European Economic Association*.
- BETTI, L., N. VON CRAMON-TAUBADEL, A. MANICA, AND S. J. LYCETT (2013): “Global geometric morphometric analyses of the human pelvis reveal substantial neutral population history effects, even across sexes,” *PloS one*, 8, e55909.
- DIAMOND, J. (1997): *Guns, Germs, and Steel: The Fates of Human Societies*, W.W. Norton.
- ENGERMAN, S. L. AND K. L. SOKOLOFF (1997): “Factor endowments, institutions, and differential paths of growth among new world economies,” *How Latin America Fell Behind*, 260–304.
- FENSKE, J. (2013): “Does land abundance explain African institutions?” *The Economic Journal*, 123, 1363–1390.
- FENSKE, J. (2014): “Ecology, Trade, and States in Pre-Colonial Africa,” *Journal of the European Economic Association*, 12, 612–640.
- GALOR, O. AND O. ÖZAK (2016): “The Agricultural Origins of Time Preference,” *American Economic Review*, 106, 3064–3103.
- GIULIANO, P. AND N. NUNN (2016): “Ancestral Characteristics of Modern Populations,” *Mimeo*.

- GOSDEN, C. (1989): “Debt, production, and prehistory,” *Journal of Anthropological Archaeology*, 8, 355–387.
- HANIHARA, T. (2008): “Morphological variation of major human populations based on nonmetric dental traits,” *American journal of physical anthropology*, 136, 169–182.
- HARPENDING, H. AND A. ROGERS (2000): “Genetic perspectives on human origins and differentiation,” *Annual review of genomics and human genetics*, 1, 361–385.
- HARRIS, I., P. JONES, T. OSBORN, AND D. LISTER (2014): “Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 Dataset,” *International Journal of Climatology*, 34, 623–642.
- HEIZER, R. F. AND W. C. STURTEVANT (1978): *Handbook of North American Indians: California*, Smithsonian Institution.
- HENN, B. M., L. L. CAVALLI-SFORZA, AND M. W. FELDMAN (2012): “The great human expansion,” *Proceedings of the National Academy of Sciences*, 109, 17758–17764.
- LA PORTA, R., F. LOPEZ-DE-SILANES, A. SHLEIFER, AND R. W. VISHNY (1999): “The Quality of Government,” *Journal of Law, Economics, and Organization*, 15, 222–279.
- LIPSET, S. M. (1960): *Political Man: The Social Bases of Politics*, Doubleday & Garden, Inc., Garden City, New York.
- LITINA, A. (2014): “The Geographical Origins of Early State Formation,” *Center for Research in Economic Analysis, University of Luxembourg*.
- MANICA, A., W. AMOS, F. BALLOUX, AND T. HANIHARA (2007): “The effect of ancient population bottlenecks on human phenotypic variation,” *Nature*, 448, 346–348.
- MANN, M. (1986): *The Sources of Social Power: Volume 1, A History of Power from the Beginning to AD 1760*, Cambridge paperback Library, Cambridge University Press.
- MARSHALL, M. G., T. R. GURR, AND K. JAGGERS (2014): “Polity IV Project: Political Regime Characteristics and Transitions, 1800–2013: Dataset User’s Manual,” Tech. rep., Center for Systemic Peace.
- MICHALOPOULOS, S. (2012): “The origins of ethnolinguistic diversity,” *The American economic review*, 102, 1508.
- MURDOCK, G. P. (1967): “Ethnographic atlas: a summary,” *Ethnology*, 109–236.
- MURDOCK, G. P. AND D. R. WHITE (1969): “Standard cross-cultural sample,” *Ethnology*, 329–369.
- MURDOCK, G. P. ET AL. (1959): “Africa: its peoples and their culture history,” .
- NUNN, N. (2008): “The long-term effects of Africa’s slave trades,” *The Quarterly Journal of Economics*, 123, 139–176.
- NUNN, N. AND L. WANTCHEKON (2011): “The Slave Trade and the Origins of Mistrust in Africa,” *The American Economic Review*, 101, 3221–3252.
- OSTER, E. (2017): “Unobservable selection and coefficient stability: Theory and evidence,” *Journal of Business & Economic Statistics*, 1–18.

- PEMBERTON, T. J., M. DEGIORGIO, AND N. A. ROSENBERG (2013): “Population structure in a comprehensive genomic data set on human microsatellite variation,” *G3: Genes— Genomes— Genetics*, g3–113.
- PRUGNOLLE, F., A. MANICA, AND F. BALLOUX (2005): “Geography Predicts Neutral Genetic Diversity of Human Populations,” *Current Biology*, 15, R159–R160.
- RAMACHANDRAN, S., O. DESHPANDE, C. C. ROSEMAN, N. A. ROSENBERG, M. W. FELDMAN, AND L. L. CAVALLI-SFORZA (2005): “Support from the Relationship of Genetic and Geographic Distance in Human Populations for a Serial Founder Effect Originating in Africa,” *Proceedings of the National Academy of Sciences*, 102, 15942–15947.
- SIEBERT, S., V. HENRICH, K. FRENKEN, AND J. BURKE (2013): “Global Map of Irrigation Areas version 5,” Tech. rep., Rheinische Friedrich-Wilhelms-University, Bonn, Germany / Food and Agriculture Organization of the United Nations, Rome, Italy.
- 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.
- WANG, S., C. M. LEWIS JR, M. JAKOBSSON, S. RAMACHANDRAN, N. RAY, G. BEDOYA, W. ROJAS, M. V. PARRA, J. A. MOLINA, C. GALLO, ET AL. (2007): “Genetic variation and population structure in Native Americans,” *PLoS Genetics*, 3, e185.
- WEIDMANN, N. B., J. K. RØD, AND L.-E. CEDERMAN (2010): “Representing ethnic groups in space: A new dataset,” *Journal of Peace Research*, 47, 491–499.

## A Additional Tables

Table A.1: Genetic Diversity and Interpersonal Trust in the US

	Trust in People						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted Genetic Diversity	-7.008*** (0.011)	-8.318*** (0.712)	-7.750*** (0.577)	-7.810*** (1.007)	-7.476*** (2.126)	-8.045*** (2.325)	-8.605*** (2.599)
Family origin continent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	No	No	Yes	Yes	Yes	Yes	Yes
Sex FE	No	No	Yes	Yes	Yes	Yes	Yes
Income FE	No	No	No	Yes	Yes	Yes	Yes
Religion FE	No	No	No	No	Yes	Yes	Yes
Education FE	No	No	No	No	No	Yes	Yes
Region in the USA FE	No	No	No	No	No	No	Yes
Observations	1149	1149	1149	906	906	906	906
Adjusted $R^2$	0.062	0.090	0.168	0.186	0.194	0.216	0.231

This table presents the results of an individual-level OLS regression analysis of interpersonal trust among second-generation migrants in the US (as reported in the General Social Survey (GSS)) on predicted population diversity (as captured by predicted genetic diversity of their parental country of origin), controlling for a range of individual-level socioeconomic characteristics (i.e., age, gender, income, religion, education), as well as time period fixed effects, regional fixed effects associated with the parental homeland, and regional fixed effects associated with the location of the second-generation migrant in the US. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.2: Genetic Diversity and Interpersonal Trust in Africa

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Intra-Group Trust								
Genetic Diversity	-23.010** (10.472)	-21.851** (10.148)	-28.775** (11.959)	-26.399*** (9.402)	-26.380*** (8.992)	-26.105*** (8.025)	-25.436*** (7.588)	-18.258*** (6.221)	-13.843*** (4.780)
Age		0.006*** (0.001)	0.007*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.000)	0.004*** (0.000)	0.004*** (0.001)	0.004*** (0.001)
Male		-0.036* (0.019)	-0.038** (0.019)	-0.036* (0.020)	-0.036* (0.020)	-0.037* (0.021)	-0.018 (0.028)	-0.029 (0.036)	-0.027 (0.037)
Slave Exports (Atlantic and Indian)			-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
School Present				-0.136*** (0.039)	-0.136*** (0.040)	-0.141*** (0.042)	-0.125*** (0.037)	-0.124*** (0.022)	-0.121*** (0.022)
Electricity Present				-0.238*** (0.063)	-0.239*** (0.048)	-0.237*** (0.046)	-0.232*** (0.040)	-0.238*** (0.041)	-0.242*** (0.042)
Piped Water Present				-0.040 (0.038)	-0.040 (0.048)	-0.038 (0.047)	-0.034 (0.048)	-0.028 (0.047)	-0.028 (0.048)
Sewage Present				0.028 (0.045)	0.028 (0.044)	0.028 (0.049)	0.034 (0.053)	0.059 (0.071)	0.059 (0.071)
Health Clinic Present				0.022* (0.012)	0.022* (0.013)	0.023* (0.012)	0.028*** (0.010)	0.039*** (0.013)	0.034*** (0.008)
Living in an Urban Area					0.003 (0.067)	-0.000 (0.069)	0.012 (0.060)	-0.005 (0.045)	-0.007 (0.044)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Living Conditions FE	No	No	No	No	No	Yes	Yes	Yes	Yes
Education FE	No	No	No	No	No	No	Yes	Yes	Yes
Religion FE	No	No	No	No	No	No	No	Yes	Yes
Home Country FE	No	No	No	No	No	No	No	No	Yes
Observations	3448	3448	3448	3448	3448	3448	3448	3448	3448
Adjusted $R^2$	0.220	0.227	0.236	0.248	0.248	0.249	0.253	0.261	0.264

This table presents the results of an individual-level OLS regression analysis of interpersonal trust towards individuals of the same ethnicity (as reported recorded in Nunn and Wantchekon (2011)) on observed population diversity in the ancestral ethnicity of these individuals, controlling for a range of individual characteristics (i.e., age, gender, living conditions, education, religion), the presence of a school, electricity, piped water, sewage, a health clinic, in the local area, whether the local area is urban, and the intensity of Atlantic and Indian slave exports. In addition, the analysis accounts for host country fixed effects as well as fixed effects associated with the country in which the homeland of the individual's ethnicity is located. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.3: Summary Statistics

	Average	S.D.	P25	P75	N
SUMMARY STATISTICS FOR THE PRE-COLONIAL ETHNIC-GROUP DATA					
Log Levels of Jurisdictional Hierarchy	0.51	0.51	0.00	0.69	1147
Social Stratification	0.79	0.86	0.00	2.00	1102
Intensity of Slavery	0.73	0.78	0.00	1.00	1113
Degree of Absence of Checks on Leader's Power	1.17	0.84	1.00	2.00	86
Difficulty of Removal of Leaders	1.29	0.98	1.00	2.00	77
Leader's Exercise of Authority	0.85	0.83	0.00	2.00	87
Degree of Lack of Community Decisions	0.87	0.80	0.00	2.00	90
Perception of Leader's Power	0.86	0.82	0.00	2.00	90
Indigenous Democracy	0.33	0.34	0.03	0.59	1188
Indigenous Democracy (Ancestry Adjusted)	0.32	0.31	0.07	0.74	1188
Area (Millions of km <sup>2</sup> )	0.07	0.35	0.00	0.05	1263
Observed Genetic Diversity	0.73	0.05	0.71	0.76	145
Predicted Genetic Diversity	0.71	0.05	0.67	0.75	1263
Elevation	635.73	731.65	216.55	1026.84	1263
Average Temperature	19.88	8.40	15.53	26.15	1253
Temperature Range	11.67	2.83	9.70	13.65	1253
Any diversity (FAO)	0.91	0.29	1.00	1.00	1097
Ecological diversity (FAO classes)	0.42	0.25	0.20	0.62	1097
Year in Ethnographic Atlas	1895.64	154.70	1890.00	1940.00	1275
Terrain Ruggedness	122799.53	133770.91	24765.56	175078.25	1263
SUMMARY STATISTICS FOR THE MODERN-COUNTRY DATA					
Log Constraint on Chief Executive	1.69	0.40	1.39	2.08	158
Log Autocracy	0.77	0.83	0.00	1.61	158
Log Democracy	1.55	0.87	0.80	2.30	158
2008 dictatorship	0.39	0.49	0.00	1.00	183
Predicted Genetic Diversity	0.73	0.03	0.72	0.74	158
Absolute Latitude	25.28	16.93	12.00	40.00	183
Agricultural Suitability	6.51	3.54	3.81	9.04	183
Elevation	0.45	0.63	0.15	0.68	183
Ruggedness	130.89	123.75	38.49	196.89	183
Distance to Waterway	2.08	8.85	0.02	0.70	183
Colony	0.69	0.46	0.00	1.00	182
Time Since Neolithic Transition (in 10.000 Years)	0.54	0.21	0.36	0.70	154
Percentage of Area Equipped for Irrigation	2.60	4.25	0.11	3.43	183
Colonial duration	1.32	1.65	0.00	2.59	183
Population density in 1500 CE	6.19	9.51	1.08	7.45	172
Ethnic fractionalization	0.44	0.26	0.19	0.66	180
Years of schooling	4.81	2.81	2.44	7.12	129

This table reports a range of summary statistics of variables included in the ethnic-level and the country-level analysis, respectively. The statistics include the average, the standard deviation (S.D.), 25th percentile (P25), the 75th percentile (P75), and the number of observations (*N*).

Table A.4: Observed Diversity and Jurisdictional Hierarchy — Selection on Unobservables and Bootstrapped Standard Errors

	Log Number of Levels of Jurisdictional Hierarchy				
	(1)	(2)	(3)	(4)	(5)
Genetic Diversity	2.565*** (0.934)	3.680*** (0.915)	3.705*** (0.956)	4.491*** (1.237)	4.367*** (1.215)
Absolute Latitude		0.010*** (0.004)	0.011** (0.004)	0.010** (0.004)	0.030*** (0.006)
Agricultural Suitability			0.004 (0.019)	-0.001 (0.020)	0.003 (0.019)
Elevation				-0.020 (0.123)	0.227 (0.144)
Ruggedness				0.501 (0.503)	0.075 (0.494)
Distance to Waterway				-0.828 (1.006)	-1.216 (1.125)
Average Temperature					0.044*** (0.013)
<i>N</i>	133	133	133	133	133
Adjusted $R^2$	0.033	0.096	0.089	0.087	0.166
AET					-6.362
Beta					5.598

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. The AET statistic captures the ratio of how large (in absolute value) selection on unobservables would need to be in order to attribute the entire effect of population diversity to selection bias (Altonji et al., 2005; Bellows and Miguel, 2009). The Beta statistic is the estimated effect of population diversity, if the proportion of selection of observables and unobservables is equal, and the maximal  $R^2$  equal to 1.3 times the observed  $R^2$  (Oster, 2017).\*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.



Table A.5: Predicted Diversity and Jurisdictional Hierarchy — Selection on Unobservables and Bootstrapped Standard Errors

	Log Number of Levels of Jurisdictional Hierarchy				
	(1)	(2)	(3)	(4)	(5)
Predicted Genetic Diversity	4.799*** (0.290)	5.024*** (0.325)	5.114*** (0.317)	5.096*** (0.342)	5.091*** (0.342)
Absolute Latitude		0.002* (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.013*** (0.002)
Agricultural Suitability			0.032*** (0.004)	0.035*** (0.004)	0.036*** (0.004)
Elevation				-0.072*** (0.027)	-0.011 (0.031)
Ruggedness				0.149 (0.118)	0.084 (0.117)
Distance to Waterway				0.630*** (0.174)	0.445** (0.175)
Average Temperature					0.017*** (0.005)
Bootstrapped Standard Error	(0.301)***	(0.339)***	(0.327)***	(0.356)***	(0.353)***
$N$	1116	1116	1116	1116	1116
Adjusted $R^2$	0.178	0.180	0.218	0.228	0.238
AET					-76.124
Beta					5.250

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. The bottom part of the table reports bootstrapped standard errors of predicted diversity, accounting for the fact that the predicted level of diversity is a generated regressor. Heteroscedasticity-robust standard errors are reported in parentheses. The AET statistic captures the ratio of how large (in absolute value) selection on unobservables would need to be in order to attribute the entire effect of population diversity to selection bias (Altonji et al., 2005; Bellows and Miguel, 2009). The Beta statistic is the estimated effect of population diversity, if the proportion of selection of observables and unobservables is equal, and the maximal  $R^2$  equal to 1.3 times the observed  $R^2$  (Oster, 2017).\*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.6: Predicted Diversity and Autocratic Institutions — Selection on Unobservables and Bootstrapped Standard Errors

	Degree of Absence of Checks on Leader's Power		Difficulty of Removal of Leaders		Leader's Exercise of Authority		Degree of Lack of Community Decisions		Perception of Leader's Power		Indigenous Autocracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Predicted Genetic Diversity	4.514**	4.217**	5.648***	5.190**	5.019***	5.196***	6.199***	6.009***	4.974***	5.737***	1.140***	1.301***
	(1.783)	(1.829)	(2.050)	(2.002)	(1.725)	(1.780)	(1.385)	(1.470)	(1.601)	(1.745)	(0.397)	(0.410)
Absolute Latitude	-0.010*	0.000	-0.005	0.004	-0.010*	-0.002	-0.008*	-0.019	-0.005	0.004	-0.002	0.004
	(0.005)	(0.014)	(0.007)	(0.016)	(0.005)	(0.014)	(0.005)	(0.012)	(0.005)	(0.013)	(0.001)	(0.003)
Agricultural Suitability	-0.039	-0.044*	0.011	0.013	-0.010	-0.020	-0.004	-0.010	0.010	-0.003	0.020***	0.020***
	(0.024)	(0.025)	(0.034)	(0.034)	(0.027)	(0.025)	(0.023)	(0.024)	(0.024)	(0.023)	(0.005)	(0.005)
Elevation		0.047		-0.276		0.032		-0.235		-0.025		-0.020
		(0.190)		(0.234)		(0.180)		(0.160)		(0.158)		(0.030)
Ruggedness		0.913		2.698***		0.644		1.161*		-0.052		-0.066
		(0.707)		(0.883)		(0.658)		(0.651)		(0.607)		(0.144)
Distance to Waterway		-1.057**		-0.426		-1.584***		-0.078		-1.524***		0.028
		(0.442)		(0.779)		(0.526)		(0.460)		(0.492)		(0.242)
Average Temperature		0.026		0.027		0.020		-0.019		0.020		0.011**
		(0.028)		(0.032)		(0.027)		(0.023)		(0.026)		(0.005)
Bootstrapped standard error	(1.795)**	(1.866)**	(2.061)***	(2.032)**	(1.816)***	(1.887)***	(1.455)***	(1.539)***	(1.661)***	(1.851)***	(0.425)***	(0.438)***
<i>N</i>	83	83	74	74	84	84	87	87	87	87	898	898
Adjusted <i>R</i> <sup>2</sup>	0.085	0.091	0.067	0.156	0.104	0.113	0.165	0.168	0.085	0.084	0.053	0.062
AET		14.181		11.337		-29.252		31.658		-7.523		-8.109
Beta		3.917		4.944		5.393		5.682		6.607		1.565

This table presents the results of an ethnic-group level OLS regression analysis of measures of pre-colonial autocracy on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables. The table includes bootstrapped standard errors that account for the uncertainty in the first stage of the prediction of genetic diversity based on the migratory distance from East Africa. Furthermore, the table includes statistics for selection on unobservables. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. The bottom part of the table reports bootstrapped standard errors of predicted diversity, accounting for the fact that the predicted level of diversity is a generated regressor. Heteroscedasticity-robust standard errors are reported in parentheses. The AET statistic captures the ratio of how large (in absolute value) selection on unobservables would need to be in order to attribute the entire effect of population diversity to selection bias (Altonji et al., 2005; Bellows and Miguel, 2009). The Beta statistic is the estimated effect of population diversity, if the proportion of selection of observables and unobservables is equal, and the maximal *R*<sup>2</sup> equal to 1.3 times the observed *R*<sup>2</sup> (Oster, 2017).\*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.7: Observed Diversity and Jurisdictional Hierarchy — Ordered Probit

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	4.862** (2.375)	8.436*** (3.113)	8.616*** (3.342)	10.140** (4.084)	10.627** (4.166)	21.837*** (7.833)
Absolute Latitude		0.024*** (0.009)	0.026** (0.011)	0.025** (0.010)	0.074*** (0.016)	0.049** (0.020)
Agricultural Suitability			0.017 (0.041)	0.007 (0.043)	0.027 (0.043)	0.018 (0.049)
Elevation				0.010 (0.277)	0.589* (0.343)	0.419 (0.347)
Ruggedness				0.771 (1.203)	-0.183 (1.260)	-0.624 (1.240)
Distance to Waterway				-1.987 (2.146)	-2.571 (2.412)	0.426 (2.828)
Average Temperature					0.100*** (0.028)	0.072** (0.033)
Regional FE	No	No	No	No	No	Yes
<i>N</i>	133	133	133	133	133	133

This table presents the results of an ethnic-group level ordered probit regression analysis of a measure of pre-colonial jurisdictional hierarchy (as captured by the number of levels of jurisdictional hierarchy) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.8: Predicted Diversity and Jurisdictional Hierarchy — Ordered Probit

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	11.587*** (0.885)	12.617*** (1.116)	13.198*** (1.134)	13.613*** (1.326)	13.543*** (1.328)	10.347*** (3.033)
Absolute Latitude		0.006** (0.003)	0.014*** (0.003)	0.015*** (0.003)	0.032*** (0.006)	0.041*** (0.007)
Agricultural Suitability			0.079*** (0.011)	0.087*** (0.011)	0.091*** (0.011)	0.105*** (0.012)
Elevation				-0.222*** (0.080)	-0.073 (0.091)	0.187* (0.096)
Ruggedness				0.513 (0.316)	0.360 (0.318)	-0.405 (0.362)
Distance to Waterway				1.378*** (0.392)	0.967** (0.398)	0.545 (0.441)
Average Temperature					0.040*** (0.012)	0.036*** (0.013)
Regional FE	No	No	No	No	No	Yes
<i>N</i>	1116	1116	1116	1116	1116	1116

This table presents the results of an ethnic-group level ordered probit regression analysis of a measure of pre-colonial jurisdictional hierarchy (as captured by the number of levels of jurisdictional hierarchy) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.9: Jurisdictional Hierarchy and Autocratic Institutions — Ordered Probit

	Degree of Absence of Checks on Leader's Power		Difficulty of Removal of Leaders		Leader's Exercise of Authority		Degree of Lack of Community Decisions		Perception of Leader's Power	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log Levels of Jurisdictional Hierarchy	1.367*** (0.253)	1.526*** (0.279)	1.234*** (0.272)	1.169*** (0.263)	1.517*** (0.292)	1.628*** (0.325)	1.458*** (0.282)	1.505*** (0.298)	1.502*** (0.292)	1.632*** (0.312)
Absolute Latitude	-0.007 (0.007)	-0.001 (0.021)	-0.009 (0.007)	-0.005 (0.021)	-0.013* (0.008)	0.001 (0.023)	-0.012* (0.007)	-0.038* (0.022)	-0.008 (0.007)	0.005 (0.021)
Agricultural Suitability		-0.116*** (0.043)		-0.019 (0.042)		-0.076 (0.052)		-0.064 (0.044)		-0.048 (0.045)
Elevation		0.118 (0.265)		-0.255 (0.296)		0.118 (0.236)		-0.383 (0.274)		0.096 (0.241)
Ruggedness		1.227 (0.958)		3.149*** (1.085)		0.754 (0.987)		1.904* (1.091)		-0.600 (0.945)
Distance to Waterway		-1.782*** (0.595)		-0.382 (1.008)		-3.210* (1.657)		0.090 (0.762)		-2.968* (1.672)
Average Temperature		0.044 (0.040)		0.025 (0.041)		0.048 (0.043)		-0.038 (0.042)		0.033 (0.042)
<i>N</i>	82	82	74	74	83	83	86	86	86	86

This table presents the results of an ethnic-group level ordered probit regression analysis of non-binary measures of pre-colonial autocracy on a measure of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy), conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.10: Predicted Diversity and Autocratic Institutions — Ordered Probit

	Degree of Absence of Checks on Leader's Power		Difficulty of Removal of Leaders		Leader's Exercise of Authority		Degree of Lack of Community Decisions		Perception of Leader's Power	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Predicted Genetic Diversity	6.452** (2.564)	6.280** (2.553)	6.449*** (2.476)	6.449** (2.725)	7.803*** (2.850)	8.569*** (2.963)	10.097*** (2.464)	10.268*** (2.658)	7.686*** (2.613)	8.966*** (2.846)
Absolute Latitude	-0.008 (0.007)	-0.001 (0.020)	-0.009 (0.007)	0.002 (0.020)	-0.013** (0.006)	-0.004 (0.021)	-0.012* (0.006)	-0.035* (0.020)	-0.009 (0.007)	0.004 (0.019)
Agricultural Suitability		-0.068* (0.039)		0.019 (0.043)		-0.044 (0.046)		-0.028 (0.045)		-0.008 (0.038)
Elevation		0.061 (0.261)		-0.371 (0.301)		0.054 (0.261)		-0.414 (0.277)		-0.011 (0.235)
Ruggedness		1.363 (0.966)		3.629*** (1.140)		0.997 (0.981)		1.960* (1.083)		-0.119 (0.887)
Distance to Waterway		-1.508** (0.615)		-0.353 (0.935)		-3.186* (1.798)		-0.190 (0.672)		-2.920* (1.722)
Average Temperature		0.037 (0.040)		0.033 (0.042)		0.031 (0.041)		-0.036 (0.038)		0.027 (0.039)
<i>N</i>	83	83	74	74	84	84	87	87	87	87

This table presents the results of an ethnic-group level ordered probit regression analysis of non-binary measures of pre-colonial autocracy on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.11: Diversity and Jurisdictional Hierarchy — Accounting for Time Since Settlement

	Log Number of Levels of Jurisdictional Hierarchy					
	OLS					IV
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	2.427** (0.938)	3.459*** (0.909)	3.455*** (0.952)	4.374*** (1.177)	4.290*** (1.152)	4.252*** (1.504)
Time Since Settlement (in 10.000 Years)	0.430*** (0.133)	0.368*** (0.119)	0.368*** (0.122)	0.449*** (0.150)	0.346** (0.147)	0.346** (0.142)
Absolute Latitude		0.009** (0.004)	0.009** (0.004)	0.008* (0.004)	0.027*** (0.006)	0.027*** (0.006)
Agricultural Suitability			-0.001 (0.019)	-0.003 (0.019)	0.001 (0.018)	0.001 (0.018)
Elevation				-0.023 (0.118)	0.198 (0.140)	0.199 (0.138)
Ruggedness				0.749 (0.473)	0.311 (0.485)	0.306 (0.498)
Distance to Waterway				-0.411 (1.082)	-0.853 (1.185)	-0.850 (1.143)
Average Temperature					0.040*** (0.013)	0.040*** (0.012)
<i>N</i>	133	133	133	133	133	133
Adjusted $R^2$	0.073	0.123	0.116	0.126	0.185	0.185
1st Stage $F$ -statistic (K-P)						171.799

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables including a measure of the time since settlement. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.12: Predicted Diversity and Jurisdictional Hierarchy — Accounting for Time Since Settlement

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	4.672*** (0.292)	4.887*** (0.327)	5.000*** (0.320)	4.948*** (0.346)	4.946*** (0.345)	3.900*** (1.141)
Time Since Settlement (in 10.000 Years)	0.220*** (0.038)	0.218*** (0.039)	0.177*** (0.040)	0.181*** (0.041)	0.176*** (0.041)	0.139*** (0.039)
Absolute Latitude		0.002 (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.012*** (0.002)	0.015*** (0.003)
Agricultural Suitability			0.030*** (0.004)	0.033*** (0.004)	0.034*** (0.004)	0.037*** (0.005)
Elevation				-0.068** (0.026)	-0.008 (0.030)	0.074** (0.031)
Ruggedness				0.179 (0.118)	0.113 (0.117)	-0.086 (0.121)
Distance to Waterway				0.667*** (0.174)	0.485*** (0.175)	0.265 (0.190)
Average Temperature					0.017*** (0.005)	0.014*** (0.005)
Regional FE	No	No	No	No	No	Yes
<i>N</i>	1115	1115	1115	1115	1115	1115
Adjusted <i>R</i> <sup>2</sup>	0.195	0.196	0.229	0.239	0.249	0.298

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables including a measure of the time since settlement. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.



Table A.13: Diversity and Jurisdictional Hierarchy — Accounting for Ecological Diversity

	Log Number of Levels of Jurisdictional Hierarchy					
	OLS					IV
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	2.380** (0.963)	3.450*** (0.954)	3.414*** (1.044)	3.964*** (1.278)	3.603*** (1.239)	3.785** (1.540)
Ecological diversity (FAO classes)	0.246 (0.279)	0.204 (0.276)	0.218 (0.316)	0.216 (0.307)	0.292 (0.280)	0.289 (0.271)
Any Diversity (FAO)	0.382** (0.175)	0.508** (0.221)	0.503** (0.236)	0.440 (0.324)	0.722** (0.335)	0.710** (0.325)
Absolute Latitude		0.010** (0.004)	0.010** (0.005)	0.009** (0.005)	0.034*** (0.006)	0.034*** (0.006)
Agricultural Suitability			-0.004 (0.022)	-0.004 (0.024)	0.002 (0.022)	0.002 (0.021)
Elevation				-0.012 (0.112)	0.277** (0.129)	0.274** (0.128)
Ruggedness				0.405 (0.475)	-0.111 (0.454)	-0.088 (0.471)
Distance to Waterway				-0.208 (1.100)	-0.305 (1.197)	-0.333 (1.140)
Average Temperature					0.052*** (0.013)	0.052*** (0.013)
<i>N</i>	131	131	131	131	131	131
Adjusted <i>R</i> <sup>2</sup>	0.042	0.100	0.093	0.081	0.191	0.191
1st Stage <i>F</i> -statistic (K-P)						184.938

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables including measures of ecological polarization and diversity. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.14: Predicted Diversity and Jurisdictional Hierarchy — Accounting for Ecological Diversity

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	4.961*** (0.285)	5.340*** (0.319)	5.531*** (0.311)	5.325*** (0.326)	5.345*** (0.325)	4.524*** (1.087)
Ecological Polarization	-0.304*** (0.111)	-0.311*** (0.111)	-0.300*** (0.111)	-0.298*** (0.112)	-0.342*** (0.111)	-0.294*** (0.105)
Ecological Diversity	0.652*** (0.135)	0.655*** (0.135)	0.561*** (0.135)	0.610*** (0.138)	0.698*** (0.137)	0.627*** (0.128)
Absolute Latitude		0.002*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.018*** (0.002)	0.018*** (0.003)
Agricultural Suitability			0.032*** (0.004)	0.035*** (0.004)	0.038*** (0.004)	0.034*** (0.005)
Elevation				-0.047 (0.029)	0.059* (0.033)	0.094*** (0.033)
Ruggedness				0.066 (0.111)	-0.068 (0.110)	-0.202* (0.116)
Distance to Waterway				0.800*** (0.179)	0.531*** (0.176)	0.290 (0.177)
Average Temperature					0.027*** (0.004)	0.022*** (0.005)
Regional FE	No	No	No	No	No	Yes
<i>N</i>	1081	1081	1081	1081	1081	1081
Adjusted $R^2$	0.228	0.233	0.270	0.280	0.304	0.343

This table presents the results of an ethnic-group level OLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables including measures of ecological polarization and diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.15: Observed Diversity and Jurisdictional Hierarchy — Accounting for Major Crop Type

	Log Number of Levels of Jurisdictional Hierarchy					
	OLS				IV	
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	2.064** (0.987)	3.318*** (1.047)	3.331*** (1.078)	3.777*** (1.313)	4.053*** (1.278)	3.972** (1.605)
Absolute Latitude		0.010*** (0.003)	0.010** (0.004)	0.009** (0.004)	0.027*** (0.007)	0.027*** (0.007)
Agricultural Suitability			0.003 (0.017)	0.002 (0.019)	0.004 (0.018)	0.004 (0.018)
Elevation				-0.031 (0.130)	0.182 (0.155)	0.182 (0.149)
Ruggedness				0.360 (0.517)	0.074 (0.528)	0.064 (0.536)
Distance to Waterway				-0.313 (1.040)	-0.814 (1.114)	-0.807 (1.061)
Average Temperature					0.038*** (0.013)	0.038*** (0.013)
Major Crop Type FE	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	133	133	133	133	133	133
Adjusted $R^2$	0.084	0.134	0.128	0.112	0.157	0.157
1st Stage $F$ -statistic (K-P)						159.030

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables including major crop type dummy variables. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.16: Predicted Diversity and Jurisdictional Hierarchy — Accounting for Major Crop Type

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	3.460*** (0.351)	4.305*** (0.386)	4.426*** (0.382)	4.359*** (0.392)	4.444*** (0.400)	4.157*** (1.105)
Absolute Latitude		0.008*** (0.001)	0.009*** (0.001)	0.009*** (0.001)	0.013*** (0.003)	0.014*** (0.003)
Agricultural Suitability			0.021*** (0.004)	0.023*** (0.004)	0.024*** (0.004)	0.030*** (0.005)
Elevation				-0.062** (0.025)	-0.029 (0.030)	0.038 (0.031)
Ruggedness				0.144 (0.113)	0.109 (0.114)	-0.052 (0.118)
Distance to Waterway				0.555*** (0.169)	0.462*** (0.175)	0.312* (0.180)
Average Temperature					0.009* (0.005)	0.008 (0.005)
Major Crop Type FE	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes
<i>N</i>	1116	1116	1116	1116	1116	1116
Adjusted $R^2$	0.237	0.269	0.283	0.290	0.292	0.320

This table presents the results of an ethnic-group level OLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables including major crop type dummy variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.17: Observed Diversity and Jurisdictional Hierarchy — Accounting for the Standard Deviation of Soil Suitability

	Log Number of Levels of Jurisdictional Hierarchy					
	OLS					IV
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	3.636*** (0.845)	4.577*** (0.810)	4.579*** (0.823)	5.355*** (1.020)	5.068*** (1.058)	5.457*** (1.322)
S.D. of Agricultural Suitability	0.160*** (0.052)	0.148*** (0.053)	0.148*** (0.053)	0.181*** (0.058)	0.143** (0.060)	0.147*** (0.056)
Absolute Latitude		0.009*** (0.003)	0.010** (0.004)	0.008** (0.004)	0.026*** (0.006)	0.026*** (0.006)
Agricultural Suitability			0.000 (0.019)	-0.005 (0.020)	-0.001 (0.019)	-0.000 (0.018)
Elevation				-0.119 (0.094)	0.111 (0.119)	0.101 (0.118)
Ruggedness				0.364 (0.449)	0.031 (0.452)	0.080 (0.462)
Distance to Waterway				-0.785 (0.889)	-1.123 (1.006)	-1.148 (0.964)
Average Temperature					0.038*** (0.013)	0.037*** (0.012)
<i>N</i>	133	133	133	133	133	133
Adjusted $R^2$	0.097	0.149	0.142	0.146	0.199	0.198
1st Stage $F$ -statistic (K-P)						193.650

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables including the standard deviation of soil suitability. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.18: Observed Diversity and Jurisdictional Hierarchy — Accounting for the Standard Deviation of Soil Suitability

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	4.913*** (0.281)	5.030*** (0.316)	5.113*** (0.311)	5.223*** (0.321)	5.218*** (0.319)	4.086*** (1.097)
S.D. of Agricultural Suitability	0.072*** (0.018)	0.068*** (0.019)	0.052*** (0.020)	0.095*** (0.024)	0.095*** (0.025)	0.097*** (0.022)
Absolute Latitude		0.001 (0.001)	0.004*** (0.001)	0.005*** (0.001)	0.012*** (0.002)	0.015*** (0.003)
Agricultural Suitability			0.030*** (0.004)	0.030*** (0.004)	0.031*** (0.004)	0.034*** (0.005)
Elevation				-0.104*** (0.026)	-0.044 (0.030)	0.044 (0.030)
Ruggedness				-0.045 (0.136)	-0.111 (0.137)	-0.334** (0.136)
Distance to Waterway				0.631*** (0.171)	0.446*** (0.172)	0.235 (0.183)
Average Temperature					0.017*** (0.004)	0.014*** (0.004)
Regional FE	No	No	No	No	No	Yes
$N$	1116	1116	1116	1116	1116	1116
Adjusted $R^2$	0.192	0.191	0.225	0.244	0.254	0.309

This table presents the results of an ethnic-group level OLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables including the standard deviation of soil suitability. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.19: Observed Diversity and Jurisdictional Hierarchy — Accounting for Scale

	Log Number of Levels of Jurisdictional Hierarchy					
	OLS					IV
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	2.693*** (0.746)	3.232*** (0.805)	3.079*** (0.821)	4.155*** (0.896)	4.202*** (0.900)	4.122*** (0.923)
Area (Millions of km <sup>2</sup> )	0.054 (0.034)	0.032 (0.032)	0.031 (0.031)	0.041 (0.030)	0.056 (0.034)	0.056* (0.032)
Absolute Latitude		0.005 (0.003)	0.003 (0.004)	0.002 (0.004)	0.009 (0.006)	0.009 (0.006)
Agricultural Suitability			-0.015 (0.017)	-0.023 (0.015)	-0.020 (0.015)	-0.020 (0.014)
Elevation				0.039 (0.077)	0.111 (0.093)	0.112 (0.087)
Ruggedness				0.617* (0.366)	0.511 (0.364)	0.499 (0.338)
Distance to Waterway				-1.464 (1.188)	-1.550 (1.215)	-1.539 (1.137)
Average Temperature					0.015 (0.012)	0.015 (0.011)
Size of Local Community FE	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	133	133	133	133	133	133
Adjusted <i>R</i> <sup>2</sup>	0.385	0.393	0.393	0.426	0.428	0.428
1st Stage <i>F</i> -statistic (K-P)						205.260

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables including ethnicity-homeland area and mean size of local communities dummy variables. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.20: Predicted Diversity and Jurisdictional Hierarchy — Accounting for Scale

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	3.856*** (0.283)	3.771*** (0.293)	3.925*** (0.292)	3.836*** (0.301)	3.866*** (0.304)	1.720* (1.022)
Area (Millions of km <sup>2</sup> )	0.111* (0.062)	0.114* (0.063)	0.122* (0.067)	0.115* (0.064)	0.121* (0.065)	0.113* (0.063)
Absolute Latitude		-0.001 (0.001)	0.001* (0.001)	0.001* (0.001)	0.004** (0.002)	0.009*** (0.002)
Agricultural Suitability			0.020*** (0.004)	0.023*** (0.004)	0.024*** (0.004)	0.031*** (0.004)
Elevation				-0.042* (0.022)	-0.021 (0.026)	0.062** (0.026)
Ruggedness				0.136 (0.106)	0.115 (0.107)	0.018 (0.111)
Distance to Waterway				0.610*** (0.180)	0.538*** (0.184)	0.364* (0.191)
Average Temperature					0.006 (0.004)	0.007* (0.004)
Size of Local Community FE	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes
<i>N</i>	1116	1116	1116	1116	1116	1116
Adjusted <i>R</i> <sup>2</sup>	0.377	0.377	0.390	0.395	0.396	0.423

This table presents the results of an ethnic-group level OLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables including ethnicity-homeland area and mean size of local communities dummy variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.



Table A.21: Observed Diversity and Jurisdictional Hierarchy — Excluding Africa

	Log Number of Levels of Jurisdictional Hierarchy					
	OLS					IV
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	8.874*** (1.177)	9.569*** (1.240)	9.703*** (1.361)	9.820*** (1.376)	8.837*** (1.509)	10.088*** (2.029)
Absolute Latitude		-0.004 (0.005)	-0.004 (0.006)	-0.006 (0.006)	0.006 (0.012)	0.002 (0.011)
Agricultural Suitability			0.009 (0.022)	0.018 (0.021)	0.009 (0.022)	0.014 (0.021)
Elevation				0.192 (0.157)	0.251 (0.170)	0.240 (0.162)
Ruggedness				-0.799 (0.610)	-0.818 (0.605)	-0.820 (0.577)
Distance to Waterway				1.351 (2.403)	0.582 (2.660)	0.673 (2.449)
Average Temperature					0.020 (0.018)	0.014 (0.018)
<i>N</i>	56	56	56	56	56	56
Adjusted $R^2$	0.359	0.359	0.349	0.367	0.368	0.362
1st Stage $F$ -statistic (K-P)						54.289

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables and excluding observations from Africa. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Given the low number of observations when excluding Africa and focusing on the sample of observed diversity, this robustness table is generated without inclusion of continental dummies. Table A.22 establishes that the results are robust to accounting for continental fixed effects when excluding Africa in the larger sample of predicted diversity.

Table A.22: Predicted Diversity and Jurisdictional Hierarchy — Excluding Africa

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	6.379*** (0.543)	6.543*** (0.551)	6.559*** (0.542)	6.777*** (0.571)	6.734*** (0.571)	6.849*** (1.311)
Absolute Latitude		-0.002** (0.001)	0.002* (0.001)	0.002** (0.001)	0.004 (0.003)	0.007** (0.004)
Agricultural Suitability			0.045*** (0.004)	0.044*** (0.005)	0.043*** (0.005)	0.045*** (0.005)
Elevation				-0.043 (0.033)	-0.034 (0.032)	0.029 (0.035)
Ruggedness				-0.169 (0.131)	-0.178 (0.130)	-0.215* (0.127)
Distance to Waterway				0.528 (0.395)	0.486 (0.399)	0.410 (0.463)
Average Temperature					0.004 (0.006)	0.003 (0.006)
Regional FE	No	No	No	No	No	Yes
<i>N</i>	629	629	629	629	629	629
Adjusted <i>R</i> <sup>2</sup>	0.202	0.206	0.304	0.312	0.311	0.357

This table presents the results of an ethnic-group level OLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables and excluding observations from Africa. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.23: Observed Diversity and Jurisdictional Hierarchy — Accounting for Year in *Ethnographic Atlas*

	Log Number of Levels of Jurisdictional Hierarchy					
	OLS				IV	
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	2.903*** (0.798)	3.846*** (0.769)	3.839*** (0.792)	4.693*** (0.995)	4.556*** (1.012)	4.751*** (1.208)
Year in Ethnographic Atlas	-0.001*** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)
Absolute Latitude		0.009** (0.004)	0.009** (0.004)	0.008* (0.004)	0.028*** (0.006)	0.028*** (0.006)
Agricultural Suitability			-0.001 (0.019)	-0.004 (0.019)	0.000 (0.018)	0.000 (0.018)
Elevation				-0.052 (0.096)	0.189 (0.118)	0.186 (0.116)
Ruggedness				0.608 (0.419)	0.187 (0.411)	0.214 (0.412)
Distance to Waterway				-0.626 (0.996)	-1.019 (1.120)	-1.032 (1.080)
Average Temperature					0.043*** (0.013)	0.043*** (0.012)
<i>N</i>	133	133	133	133	133	133
Adjusted $R^2$	0.074	0.122	0.115	0.114	0.188	0.188
1st Stage $F$ -statistic (K-P)						174.988

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables including the approximate year of description as reported in the *Ethnographic Atlas*. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.24: Predicted Diversity and Jurisdictional Hierarchy — Accounting for Year in *Ethnographic Atlas*

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	4.908*** (0.286)	5.063*** (0.321)	5.137*** (0.313)	5.111*** (0.336)	5.103*** (0.336)	3.995*** (1.119)
Year in Ethnographic Atlas	-0.000*** (0.000)	-0.000** (0.000)	-0.000** (0.000)	-0.000** (0.000)	-0.000** (0.000)	-0.000** (0.000)
Absolute Latitude		0.001 (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.012*** (0.002)	0.015*** (0.003)
Agricultural Suitability			0.031*** (0.004)	0.034*** (0.004)	0.035*** (0.004)	0.038*** (0.005)
Elevation				-0.070*** (0.026)	-0.011 (0.030)	0.076** (0.030)
Ruggedness				0.150 (0.118)	0.085 (0.117)	-0.119 (0.118)
Distance to Waterway				0.634*** (0.174)	0.455*** (0.175)	0.230 (0.189)
Average Temperature					0.017*** (0.005)	0.014*** (0.005)
Regional FE	No	No	No	No	No	Yes
<i>N</i>	1115	1115	1115	1115	1115	1115
Adjusted <i>R</i> <sup>2</sup>	0.186	0.186	0.222	0.231	0.241	0.296

This table presents the results of an ethnic-group level OLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables including the approximate year of description as reported in the *Ethnographic Atlas*. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.25: Observed Diversity and Jurisdictional Hierarchy — Accounting for Additional Distances

	Log Number of Levels of Jurisdictional Hierarchy		
	(1)	(2)	(3)
Genetic Diversity	4.366*** (1.239)	4.418*** (1.189)	4.358*** (1.231)
Distance to Technological Frontier in Year 1 (in 1000 kms)	0.000 (0.031)		
Distance to Technological Frontier in Year 1000 (in 1000 kms)		-0.024 (0.034)	
Distance to Technological Frontier in Year 1500 (in 1000 kms)			0.001 (0.032)
Absolute Latitude	0.030*** (0.007)	0.029*** (0.006)	0.031*** (0.006)
Agricultural Suitability	0.003 (0.019)	0.004 (0.019)	0.003 (0.019)
Elevation	0.227 (0.141)	0.196 (0.143)	0.228 (0.141)
Ruggedness	0.075 (0.488)	0.100 (0.487)	0.072 (0.487)
Distance to Waterway	-1.216 (1.127)	-1.236 (1.152)	-1.212 (1.153)
Average Temperature	0.044*** (0.014)	0.041*** (0.014)	0.044*** (0.013)
<i>N</i>	133	133	133
Adjusted $R^2$	0.159	0.162	0.159

This table presents the results of an ethnic-group level OLS regression analysis of a measure of pre-colonial jurisdictional hierarchy (as captured by the number of levels of jurisdictional hierarchy) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables including a number of alternative distance variables. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.26: Predicted Diversity and Jurisdictional Hierarchy — Accounting for Additional Distances

	Log Number of Levels of Jurisdictional Hierarchy		
	(1)	(2)	(3)
Predicted Genetic Diversity	4.534*** (1.225)	4.104*** (1.265)	4.382*** (1.245)
Distance to Technological Frontier in Year 1 (in 1000 kms)	0.012 (0.012)		
Distance to Technological Frontier in Year 1000 (in 1000 kms)		-0.002 (0.013)	
Distance to Technological Frontier in Year 1500 (in 1000 kms)			0.007 (0.013)
Absolute Latitude	0.016*** (0.003)	0.015*** (0.003)	0.016*** (0.003)
Agricultural Suitability	0.039*** (0.005)	0.038*** (0.005)	0.039*** (0.005)
Elevation	0.084*** (0.032)	0.074** (0.031)	0.080** (0.032)
Ruggedness	-0.146 (0.120)	-0.121 (0.121)	-0.138 (0.120)
Distance to Waterway	0.230 (0.190)	0.219 (0.189)	0.226 (0.189)
Average Temperature	0.016*** (0.005)	0.014*** (0.005)	0.015*** (0.005)
Continental FE	Yes	Yes	Yes
<i>N</i>	1116	1116	1116
Adjusted <i>R</i> <sup>2</sup>	0.292	0.292	0.292

This table presents the results of an ethnic-group level OLS regression analysis of a measure of pre-colonial jurisdictional hierarchy (as captured by the number of levels of jurisdictional hierarchy) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables including a number of alternative distance variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.27: Observed Diversity and Jurisdictional Hierarchy — Accounting for Spatial Autocorrelation

	Log Number of Levels of Jurisdictional Hierarchy					
	OLS					IV
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	6.771*** (1.690)	6.594*** (1.693)	6.744*** (1.706)	7.218*** (1.712)	6.981*** (1.640)	6.544*** (2.030)
Absolute Latitude		0.004 (0.004)	0.005 (0.005)	0.006 (0.005)	0.028*** (0.008)	0.029*** (0.009)
Agricultural Suitability			0.012 (0.019)	0.005 (0.020)	0.008 (0.019)	0.007 (0.019)
Elevation				-0.075 (0.104)	0.148 (0.120)	0.136 (0.121)
Ruggedness				0.610 (0.492)	0.260 (0.476)	0.273 (0.479)
Distance to Waterway				-1.233 (0.979)	-1.836* (0.971)	-1.933** (0.967)
Average Temperature					0.044*** (0.013)	0.044*** (0.014)
Total Impact of Genetic Diversity	5.421*** (1.988)	5.347*** (2.033)	5.515** (2.177)	5.834*** (1.760)	5.517*** (1.190)	5.631*** (2.063)
<i>N</i>	114	114	114	114	114	114

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables as well as spatial autocorrelation. Variables relating to observations associated with the same homeland polygon are averaged and a single observation is kept for each polygon. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

This table establishes that the results of Table 1 are robust to accounting for spatial autocorrelation. It shows a series of spatial autoregressive (SAR) models, with a spectral-normalized inverse-distance weighting matrix, estimated with maximum-likelihood estimation, with a spatial lag of the dependent variable and a spatially lagged error. The model treats errors as heteroskedastic. In the SAR model accounting for endogeneity in column 6, genetic diversity is instrumented by the migratory distance from East Africa.

Table A.28: Predicted Diversity and Jurisdictional Hierarchy — Accounting for Spatial Autocorrelation

	Log Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	5.899*** (0.646)	5.794*** (0.671)	7.843*** (0.816)	7.693*** (0.832)	7.153*** (0.837)	3.670** (1.621)
Absolute Latitude		0.001 (0.001)	0.001 (0.002)	0.001 (0.002)	0.013*** (0.003)	0.010*** (0.003)
Agricultural Suitability			0.030*** (0.005)	0.033*** (0.005)	0.031*** (0.005)	0.033*** (0.006)
Elevation				0.012 (0.029)	0.090*** (0.033)	0.108*** (0.035)
Ruggedness				0.070 (0.146)	0.023 (0.145)	-0.036 (0.147)
Distance to Waterway				0.325 (0.211)	0.061 (0.216)	0.066 (0.219)
Average Temperature					0.026*** (0.006)	0.021*** (0.006)
Regional FE						Yes
Total Impact of Genetic Diversity	8.045 (8.429)	8.165 (9.554)	4.479*** (0.104)	4.368*** (0.109)	4.127*** (0.116)	2.327** (0.974)
<i>N</i>	987	987	987	987	987	987

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of pre-colonial jurisdictional hierarchy (as captured by the natural logarithm of the number of levels of jurisdictional hierarchy beyond the local community) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables as well as spatial autocorrelation. Variables relating to observations associated with the same homeland polygons are averaged and a single observation is kept for each polygon. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

This table establishes that the results of Table 2 are robust to accounting for spatial autocorrelation. It shows a series of spatial autoregressive (SAR) models, with a spectral-normalized inverse-distance weighting matrix, estimated with maximum-likelihood estimation, with a spatial lag of the dependent variable and a spatially lagged error. The model treat errors as heteroskedastic. In the SAR model accounting for endogeneity in column 6, genetic diversity is instrumented by the migratory distance from East Africa.



Table A.29: Observed Diversity and Jurisdictional Hierarchy — Levels Specification

	Number of Levels of Jurisdictional Hierarchy					
	OLS					IV
	(1)	(2)	(3)	(4)	(5)	(6)
Genetic Diversity	4.096*	6.697***	6.762***	8.127***	7.851***	7.514**
	(2.083)	(2.123)	(2.259)	(2.999)	(2.934)	(3.779)
Absolute Latitude		0.024***	0.025**	0.024**	0.069***	0.069***
		(0.009)	(0.011)	(0.010)	(0.014)	(0.014)
Agricultural Suitability			0.009	-0.006	0.004	0.003
			(0.044)	(0.048)	(0.045)	(0.044)
Elevation				0.005	0.555	0.560
				(0.297)	(0.349)	(0.348)
Ruggedness				0.615	-0.335	-0.380
				(1.193)	(1.190)	(1.245)
Distance to Waterway				-2.603	-3.467	-3.441
				(1.956)	(2.194)	(2.102)
Average Temperature					0.099***	0.099***
					(0.029)	(0.028)
<i>N</i>	133	133	133	133	133	133
Adjusted $R^2$	0.014	0.084	0.077	0.072	0.152	0.152
1st Stage $F$ -statistic (K-P)						173.525

This table presents the results of an ethnic-group level OLS and 2SLS regression analysis of a measure of pre-colonial jurisdictional hierarchy (as captured by the number of levels of jurisdictional hierarchy) on observed population diversity (as captured by observed genetic diversity), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.30: Predicted Diversity and Jurisdictional Hierarchy — Levels Specification

	Number of Levels of Jurisdictional Hierarchy					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	8.869*** (0.589)	9.587*** (0.692)	9.758*** (0.681)	9.704*** (0.749)	9.693*** (0.748)	8.568*** (2.357)
Absolute Latitude		0.005** (0.002)	0.012*** (0.002)	0.012*** (0.002)	0.028*** (0.005)	0.030*** (0.005)
Agricultural Suitability			0.061*** (0.009)	0.066*** (0.009)	0.069*** (0.009)	0.067*** (0.009)
Elevation				-0.118** (0.057)	0.011 (0.063)	0.173*** (0.066)
Ruggedness				0.264 (0.244)	0.125 (0.242)	-0.375 (0.253)
Distance to Waterway				1.100*** (0.369)	0.707* (0.373)	0.218 (0.397)
Average Temperature					0.036*** (0.009)	0.027*** (0.009)
Regional FE	No	No	No	No	No	Yes
$N$	1116	1116	1116	1116	1116	1116
Adjusted $R^2$	0.146	0.150	0.183	0.189	0.200	0.271

This table presents the results of an ethnic-group level OLS regression analysis of a measure of pre-colonial jurisdictional hierarchy (as captured by the number of levels of jurisdictional hierarchy) on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.31: Predicted Diversity and Autocratic Institutions — Accounting for Irrigation Potential with Predicted Diversity

	Degree of Absence of Checks on Leader's Power		Difficulty of Removal of Leaders		Leader's Exercise of Authority		Degree of Lack of Community Decisions		Perception of Leader's Power		Indigenous Autocracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Predicted Genetic Diversity	4.397** (1.857)	3.976** (1.850)	5.249** (2.006)	4.897** (2.031)	4.729*** (1.760)	4.876*** (1.816)	6.135*** (1.400)	5.872*** (1.512)	4.760*** (1.601)	5.580*** (1.764)	1.114*** (0.404)	1.361*** (0.412)
Percentage of Area Equipped for Irrigation	0.015 (0.010)	0.016 (0.010)	0.033 (0.022)	0.025 (0.020)	0.018 (0.011)	0.018 (0.012)	0.005 (0.007)	0.008 (0.007)	0.012 (0.009)	0.009 (0.009)	-0.002 (0.003)	-0.006 (0.004)
Absolute Latitude	-0.007 (0.005)	-0.005 (0.015)	-0.008 (0.006)	-0.001 (0.016)	-0.010** (0.004)	-0.008 (0.015)	-0.008** (0.004)	-0.021 (0.013)	-0.007 (0.004)	0.001 (0.013)	-0.004*** (0.001)	0.006** (0.003)
Agricultural Suitability		-0.049** (0.024)		0.012 (0.033)		-0.025 (0.025)		-0.012 (0.024)		-0.005 (0.023)		0.021*** (0.005)
Elevation		0.021 (0.189)		-0.288 (0.233)		0.002 (0.178)		-0.248 (0.162)		-0.039 (0.158)		-0.012 (0.030)
Ruggedness		0.956 (0.710)		2.612*** (0.876)		0.693 (0.666)		1.182* (0.656)		-0.027 (0.614)		-0.083 (0.144)
Distance to Waterway		-0.930** (0.466)		-0.267 (0.807)		-1.443** (0.549)		-0.015 (0.484)		-1.453*** (0.508)		-0.029 (0.244)
Average Temperature		0.018 (0.029)		0.019 (0.033)		0.011 (0.028)		-0.023 (0.025)		0.015 (0.027)		0.014*** (0.005)
<i>N</i>	83	83	74	74	84	84	87	87	87	87	898	898
Adjusted <i>R</i> <sup>2</sup>	0.079	0.097	0.102	0.165	0.128	0.124	0.167	0.162	0.095	0.079	0.038	0.064

This table presents the results of an ethnic-group level OLS regression analysis of measures of pre-colonial autocracy on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables including a measure of irrigation potential at the ethnic-group level. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.32: Predicted Diversity and Autocratic Institutions — Accounting for Time since Settlement

	Degree of Absence of Checks on Leader's Power		Difficulty of Removal of Leaders		Leader's Exercise of Authority		Degree of Lack of Community Decisions		Perception of Leader's Power		Indigenous Autocracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Predicted Genetic Diversity	4.673** (1.923)	3.844** (1.836)	5.364** (2.068)	4.216** (1.991)	5.166*** (1.797)	5.133*** (1.915)	6.198*** (1.398)	5.764*** (1.509)	4.686*** (1.592)	5.317*** (1.802)	1.049*** (0.406)	1.268*** (0.414)
Time Since Settlement (in 10.000 Years)	-0.049 (0.272)	0.224 (0.259)	0.222 (0.252)	0.587*** (0.220)	-0.111 (0.236)	0.039 (0.267)	0.010 (0.179)	0.169 (0.211)	0.261 (0.210)	0.289 (0.242)	0.073 (0.056)	0.040 (0.056)
Absolute Latitude	-0.006 (0.005)	-0.000 (0.014)	-0.006 (0.006)	0.001 (0.016)	-0.009** (0.004)	-0.002 (0.014)	-0.008* (0.004)	-0.019 (0.012)	-0.006 (0.004)	0.003 (0.013)	-0.004*** (0.001)	0.004 (0.003)
Agricultural Suitability		-0.051* (0.025)		-0.002 (0.035)		-0.021 (0.027)		-0.014 (0.024)		-0.010 (0.022)		0.019*** (0.005)
Elevation		0.037 (0.189)		-0.288 (0.225)		0.030 (0.179)		-0.244 (0.159)		-0.040 (0.154)		-0.018 (0.030)
Ruggedness		1.092 (0.760)		3.177*** (0.897)		0.675 (0.682)		1.299* (0.687)		0.186 (0.654)		-0.072 (0.148)
Distance to Waterway		-1.097** (0.443)		-0.467 (0.786)		-1.590*** (0.528)		-0.100 (0.460)		-1.563*** (0.490)		0.035 (0.242)
Average Temperature		0.027 (0.028)		0.025 (0.032)		0.020 (0.027)		-0.019 (0.024)		0.020 (0.026)		0.011** (0.005)
<i>N</i>	83	83	74	74	84	84	87	87	87	87	897	897
Adjusted <i>R</i> <sup>2</sup>	0.063	0.087	0.072	0.185	0.105	0.101	0.165	0.162	0.096	0.087	0.040	0.061

This table presents the results of an ethnic-group level OLS regression analysis of measures of pre-colonial autocracy on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables and a measure of the time since settlement. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.33: Predicted Diversity and Autocratic Institutions — Accounting for Year in *Ethnographic Atlas*

	Degree of Absence of Checks on Leader's Power		Difficulty of Removal of Leaders		Leader's Exercise of Authority		Degree of Lack of Community Decisions		Perception of Leader's Power		Indigenous Autocracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Predicted Genetic Diversity	5.006** (1.932)	4.886** (1.880)	6.351*** (1.996)	6.002*** (1.958)	5.358*** (1.796)	5.644*** (1.809)	6.789*** (1.363)	6.664*** (1.409)	5.693*** (1.595)	6.518*** (1.745)	1.128*** (0.402)	1.303*** (0.409)
Year in Ethnographic Atlas	-0.001 (0.001)	-0.002** (0.001)	-0.002** (0.001)	-0.003* (0.001)	-0.001 (0.002)	-0.002 (0.001)	-0.002* (0.001)	-0.002** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.000*** (0.000)	-0.000** (0.000)
Absolute Latitude	-0.007 (0.005)	-0.002 (0.014)	-0.008 (0.006)	0.000 (0.016)	-0.009** (0.004)	-0.004 (0.014)	-0.009** (0.004)	-0.021* (0.012)	-0.008* (0.004)	0.001 (0.013)	-0.004*** (0.001)	0.003 (0.003)
Agricultural Suitability		-0.056** (0.026)		-0.004 (0.037)		-0.029 (0.027)		-0.022 (0.025)		-0.017 (0.024)		0.019*** (0.005)
Elevation		0.046 (0.192)		-0.286 (0.228)		0.032 (0.183)		-0.239 (0.161)		-0.029 (0.159)		-0.020 (0.030)
Ruggedness		0.922 (0.709)		2.709*** (0.867)		0.650 (0.663)		1.180* (0.650)		-0.029 (0.609)		-0.059 (0.143)
Distance to Waterway		-1.098** (0.437)		-0.490 (0.756)		-1.614*** (0.525)		-0.113 (0.454)		-1.566*** (0.483)		0.036 (0.240)
Average Temperature		0.027 (0.028)		0.027 (0.032)		0.021 (0.027)		-0.017 (0.023)		0.022 (0.026)		0.011** (0.005)
<i>N</i>	83	83	74	74	84	84	87	87	87	87	897	897
Adjusted <i>R</i> <sup>2</sup>	0.070	0.101	0.089	0.170	0.110	0.113	0.188	0.187	0.119	0.113	0.048	0.067

This table presents the results of an ethnic-group level OLS regression analysis of measures of pre-colonial autocracy on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables and the approximate year of description as reported in the *Ethnographic Atlas*. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.34: Predicted Diversity and Autocratic Institutions — Alternative Distances

	Degree of Absence of Checks on Leader's Power			Indigenous Autocracy		
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Genetic Diversity	4.379** (1.834)	3.877** (1.820)	4.272** (1.801)	1.099*** (0.413)	1.818*** (0.428)	1.308*** (0.410)
Distance to Technological Frontier in Year 1 (in 1000 kms)	-0.027 (0.062)			0.046*** (0.009)		
Distance to Technological Frontier in Year 1000 (in 1000 kms)		-0.050 (0.047)			0.042*** (0.009)	
Distance to Technological Frontier in Year 1500 (in 1000 kms)			-0.036 (0.060)			0.042*** (0.010)
Absolute Latitude	-0.002 (0.016)	-0.001 (0.014)	-0.003 (0.015)	0.010*** (0.003)	0.006** (0.003)	0.008*** (0.003)
Agricultural Suitability	-0.049** (0.024)	-0.054** (0.025)	-0.051** (0.024)	0.028*** (0.005)	0.028*** (0.005)	0.027*** (0.005)
Elevation	0.011 (0.207)	-0.002 (0.191)	-0.006 (0.211)	0.049 (0.033)	0.021 (0.031)	0.043 (0.034)
Ruggedness	0.915 (0.716)	0.913 (0.702)	0.994 (0.743)	-0.131 (0.142)	-0.102 (0.142)	-0.167 (0.144)
Distance to Waterway	-1.159** (0.461)	-1.254*** (0.435)	-1.155*** (0.435)	0.137 (0.236)	0.227 (0.242)	0.148 (0.234)
Average Temperature	0.022 (0.031)	0.021 (0.029)	0.021 (0.030)	0.021*** (0.005)	0.018*** (0.005)	0.018*** (0.005)
<i>N</i>	83	83	83	898	898	898
Adjusted $R^2$	0.082	0.094	0.085	0.081	0.081	0.076

This table presents the results of an ethnic-group level OLS regression analysis of measures of pre-colonial autocracy on predicted population diversity (as captured by predicted genetic diversity based on the migratory distance from East Africa to the interior centroid of the homeland of the ethnicity), conditional on a range of geographical control variables and alternative distances. The 2SLS analysis uses migratory distance from East Africa to the interior centroid of the homeland of each ethnicity as an instrumental variable for observed genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.35: Persistence of Institutions — Alternative Autocracy Measure

	Log Executive Constraints			Log Autocracy		
	(1)	(2)	(3)	(4)	(5)	(6)
Degree of Absence of Checks on Leader's Power	-0.191*** (0.054)	-0.172** (0.069)	-0.162** (0.070)	0.556*** (0.118)	0.519*** (0.138)	0.500*** (0.140)
Predicted Genetic Diversity			-3.495** (1.368)			7.323** (3.448)
Absolute Latitude		0.010* (0.006)	0.011** (0.005)		-0.020* (0.011)	-0.021** (0.009)
Agricultural Suitability		0.042** (0.018)	0.035* (0.018)		-0.064* (0.036)	-0.052 (0.037)
Elevation		0.010 (0.110)	0.055 (0.089)		-0.066 (0.241)	-0.161 (0.243)
Ruggedness		0.001 (0.001)	0.000 (0.001)		0.000 (0.002)	0.002 (0.002)
Distance to Waterway		0.004 (0.004)	0.003 (0.003)		-0.007 (0.006)	-0.005 (0.006)
Temperature		0.005 (0.008)	0.005 (0.008)		-0.011 (0.014)	-0.009 (0.013)
Colony		0.215 (0.212)	0.203 (0.201)		-0.392 (0.377)	-0.367 (0.349)
Legal Origin FE	No	Yes	Yes	No	Yes	Yes
<i>N</i>	48	48	48	48	48	48
Adjusted <i>R</i> <sup>2</sup>	0.124	0.256	0.349	0.210	0.299	0.374

This table presents the results of an ethnic-group level OLS regression analysis of measures of contemporary autocracy on a measure of pre-colonial autocracy, conditional on a range of geographical control variables as well as predicted diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.36: Predicted Diversity and Constraint on the Executive (2013)

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-2.758*** (0.961)	-3.202*** (0.975)	-2.161** (0.993)	-2.161** (0.993)	-2.338** (0.989)	-2.309** (0.998)	-3.076** (1.454)	-5.738** (2.466)
Absolute Latitude		0.006*** (0.002)	0.006*** (0.002)	0.006*** (0.002)	0.010*** (0.002)	0.012*** (0.003)	0.007* (0.004)	0.008** (0.004)
Agricultural Suitability			0.039*** (0.010)	0.039*** (0.010)	0.051*** (0.011)	0.052*** (0.011)	0.035*** (0.011)	0.033*** (0.011)
Elevation			-0.009 (0.076)	-0.009 (0.076)	0.055 (0.082)	0.070 (0.085)	0.124 (0.101)	0.120 (0.096)
Ruggedness			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance to Waterway			0.004 (0.002)	0.004 (0.002)	0.003 (0.002)	0.002 (0.002)	0.002 (0.003)	0.002 (0.003)
Colony						0.094 (0.125)	0.187 (0.146)	0.201 (0.135)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	150	150	150	150	150	150	150	150
Adjusted $R^2$	0.029	0.093	0.190	0.190	0.282	0.280	0.365	0.348
1st Stage $F$ -statistic (K-P)								52.875

This table presents the results of a country-level OLS and 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.



Table A.37: Predicted Diversity and Constraint on the Executive — Accounting for Irrigation Potential

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-3.657*** (0.800)	-4.228*** (0.804)	-3.288*** (0.839)	-3.288*** (0.839)	-3.527*** (0.817)	-3.501*** (0.826)	-3.647*** (1.315)	-5.933*** (2.254)
Percentage of Area Equipped for Irrigation	0.001 (0.006)	-0.005 (0.007)	-0.010 (0.007)	-0.010 (0.007)	-0.009 (0.006)	-0.009 (0.006)	0.007 (0.006)	0.007 (0.006)
Absolute Latitude		0.008*** (0.001)	0.008*** (0.002)	0.008*** (0.002)	0.012*** (0.002)	0.014*** (0.003)	0.008** (0.003)	0.010*** (0.003)
Agricultural Suitability			0.040*** (0.009)	0.040*** (0.009)	0.051*** (0.010)	0.052*** (0.010)	0.028*** (0.010)	0.027*** (0.010)
Elevation			-0.090 (0.070)	-0.090 (0.070)	-0.029 (0.080)	-0.012 (0.079)	0.060 (0.091)	0.055 (0.086)
Ruggedness			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance to Waterway			0.003 (0.002)	0.003 (0.002)	0.002 (0.002)	0.001 (0.002)	0.001 (0.003)	0.001 (0.003)
Colony						0.092 (0.126)	0.200 (0.138)	0.211* (0.127)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	155	155	155	155	155	155	155	155
Adjusted $R^2$	0.048	0.138	0.249	0.249	0.342	0.341	0.447	0.435
1st Stage $F$ -statistic (K-P)								56.818

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables including a measure of land equipped for irrigation. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.38: Predicted Diversity and Constraint on the Executive — Accounting for Percentage of Land near a Waterway

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-2.943*** (0.852)	-3.460*** (0.903)	-2.640*** (0.903)	-2.640*** (0.903)	-3.042*** (0.822)	-2.986*** (0.838)	-3.158** (1.309)	-5.328** (2.204)
Percentage of Land Near a Waterway	0.370*** (0.085)	0.300*** (0.091)	0.339** (0.142)	0.339** (0.142)	0.283** (0.130)	0.292** (0.134)	0.169 (0.120)	0.158 (0.111)
Absolute Latitude		0.006*** (0.002)	0.006*** (0.002)	0.006*** (0.002)	0.010*** (0.002)	0.013*** (0.003)	0.008** (0.003)	0.009*** (0.003)
Agricultural Suitability			0.035*** (0.009)	0.035*** (0.009)	0.046*** (0.009)	0.048*** (0.010)	0.030*** (0.010)	0.030*** (0.010)
Elevation			0.148 (0.117)	0.148 (0.117)	0.184 (0.113)	0.212* (0.115)	0.223* (0.114)	0.206* (0.108)
Ruggedness			-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.001 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Distance to Waterway			0.005** (0.002)	0.005** (0.002)	0.004* (0.002)	0.004 (0.002)	0.002 (0.002)	0.002 (0.002)
Colony						0.125 (0.123)	0.234* (0.140)	0.240* (0.129)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	149	149	149	149	149	149	149	149
Adjusted <i>R</i> <sup>2</sup>	0.166	0.215	0.277	0.277	0.375	0.377	0.474	0.462
1st Stage <i>F</i> -statistic (K-P)								53.549

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables including a measure of land near a waterway. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.39: Predicted Diversity and Constraint on the Executive — Accounting for the Inequality in Land Suitability

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-3.914*** (0.831)	-4.357*** (0.862)	-3.480*** (0.890)	-3.480*** (0.890)	-3.738*** (0.854)	-3.706*** (0.867)	-3.781*** (1.373)	-5.755** (2.317)
Land Suitability Gini	-0.214 (0.150)	-0.270** (0.137)	0.151 (0.184)	0.151 (0.184)	0.066 (0.198)	0.056 (0.199)	0.046 (0.186)	0.046 (0.172)
Absolute Latitude		0.008*** (0.002)	0.008*** (0.002)	0.008*** (0.002)	0.012*** (0.002)	0.014*** (0.003)	0.008** (0.004)	0.010*** (0.003)
Agricultural Suitability			0.045*** (0.011)	0.045*** (0.011)	0.054*** (0.011)	0.055*** (0.012)	0.034*** (0.012)	0.034*** (0.011)
Elevation			-0.040 (0.090)	-0.040 (0.090)	0.036 (0.098)	0.054 (0.096)	0.143 (0.109)	0.138 (0.104)
Ruggedness			0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Distance to Waterway			0.002 (0.003)	0.002 (0.003)	0.002 (0.002)	0.001 (0.002)	0.000 (0.002)	0.000 (0.002)
Colony						0.095 (0.127)	0.254* (0.149)	0.264* (0.138)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	147	147	147	147	147	147	147	147
Adjusted <i>R</i> <sup>2</sup>	0.076	0.184	0.249	0.249	0.346	0.345	0.463	0.455
1st Stage <i>F</i> -statistic (K-P)								58.014

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables including a measure of inequality in land suitability. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.40: Predicted Diversity and Constraint on the Executive — Accounting for Percentages of Population Living in Various Climate Zones

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-3.779*** (0.895)	-3.762*** (0.910)	-3.737*** (0.907)	-3.737*** (0.907)	-4.128*** (0.839)	-4.078*** (0.838)	-3.572*** (1.310)	-6.334*** (2.340)
Absolute Latitude		0.000 (0.004)	-0.000 (0.004)	-0.000 (0.004)	0.005 (0.005)	0.008 (0.005)	0.007 (0.005)	0.009* (0.005)
Agricultural Suitability			0.017 (0.011)	0.017 (0.011)	0.041*** (0.014)	0.041*** (0.015)	0.037** (0.016)	0.042*** (0.014)
Elevation			0.083 (0.084)	0.083 (0.084)	0.111 (0.085)	0.151* (0.087)	0.156 (0.099)	0.139 (0.092)
Ruggedness			-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Distance to Waterway			0.002 (0.002)	0.002 (0.002)	0.002 (0.002)	0.001 (0.002)	0.001 (0.003)	0.001 (0.002)
Colony						0.160 (0.119)	0.252* (0.145)	0.271** (0.131)
Climate Zone Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	149	149	149	149	149	149	149	149
Adjusted <i>R</i> <sup>2</sup>	0.316	0.312	0.307	0.307	0.385	0.391	0.461	0.445
1st Stage <i>F</i> -statistic (K-P)								47.611

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables including variables capturing the percentage of the population living in various climate zones. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.41: Predicted Diversity and Constraint on the Executive — Accounting for Colonizer Nation

	Log Constraint on Chief Executive						
	OLS						IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted Genetic Diversity	-2.937*** (1.103)	-3.193*** (1.037)	-2.891** (1.127)	-2.891** (1.127)	-2.962*** (1.115)	-4.015** (1.543)	-6.916*** (2.615)
Absolute Latitude		0.009*** (0.002)	0.010*** (0.002)	0.010*** (0.002)	0.013*** (0.003)	0.008** (0.003)	0.009*** (0.003)
Agricultural Suitability			0.042*** (0.010)	0.042*** (0.010)	0.045*** (0.011)	0.028*** (0.009)	0.027*** (0.009)
Elevation			-0.012 (0.073)	-0.012 (0.073)	0.001 (0.074)	0.066 (0.090)	0.061 (0.086)
Ruggedness			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance to Waterway			0.002 (0.002)	0.002 (0.002)	0.001 (0.002)	-0.001 (0.003)	-0.001 (0.003)
Colony					0.215 (0.158)	0.181 (0.161)	0.190 (0.150)
Colonizer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes	Yes
<i>N</i>	152	152	152	152	152	152	152
Adjusted <i>R</i> <sup>2</sup>	0.089	0.178	0.288	0.288	0.294	0.423	0.405
1st Stage <i>F</i> -statistic (K-P)							50.695

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and dummy variables indicating the colonizer nation (if any). The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.42: Predicted Diversity and Constraint on the Executive — Accounting for GDP

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-3.003*** (0.844)	-3.534*** (1.029)	-2.465** (1.049)	-2.465** (1.049)	-3.077*** (1.007)	-3.026*** (1.012)	-3.743*** (1.293)	-5.834*** (2.245)
Log Income per Capita in Year 2000	0.107*** (0.027)	0.071* (0.042)	0.078* (0.042)	0.078* (0.042)	0.033 (0.043)	0.036 (0.043)	0.024 (0.044)	0.025 (0.040)
Absolute Latitude		0.004* (0.003)	0.004 (0.003)	0.004 (0.003)	0.010*** (0.003)	0.012*** (0.003)	0.008** (0.003)	0.010*** (0.003)
Agricultural Suitability			0.042*** (0.010)	0.042*** (0.010)	0.050*** (0.010)	0.052*** (0.011)	0.032*** (0.011)	0.031*** (0.010)
Elevation			-0.015 (0.077)	-0.015 (0.077)	-0.000 (0.080)	0.024 (0.082)	0.071 (0.088)	0.068 (0.085)
Ruggedness			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance to Waterway			0.003 (0.002)	0.003 (0.002)	0.003 (0.002)	0.002 (0.002)	0.001 (0.003)	0.001 (0.003)
Colony						0.130 (0.128)	0.223 (0.143)	0.235* (0.132)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	154	154	154	154	154	154	154	154
Adjusted $R^2$	0.150	0.166	0.276	0.276	0.336	0.339	0.443	0.433
1st Stage $F$ -statistic (K-P)								56.319

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and GDP per capita in 2000. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.43: Predicted Diversity and Constraint on the Executive — Accounting for Schooling

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-2.857*** (0.678)	-2.928*** (0.779)	-2.576*** (0.802)	-2.576*** (0.802)	-2.522*** (0.766)	-2.524*** (0.768)	-3.327*** (1.084)	-4.704** (1.991)
Years of Schooling	0.064*** (0.009)	0.062*** (0.017)	0.050*** (0.016)	0.050*** (0.016)	0.060*** (0.015)	0.060*** (0.015)	0.048*** (0.016)	0.047*** (0.015)
Absolute Latitude		0.000 (0.003)	0.002 (0.002)	0.002 (0.002)	0.005* (0.003)	0.004 (0.003)	0.003 (0.003)	0.004 (0.003)
Agricultural Suitability			0.029*** (0.008)	0.029*** (0.008)	0.040*** (0.008)	0.038*** (0.009)	0.026*** (0.009)	0.027*** (0.008)
Elevation			-0.060 (0.086)	-0.060 (0.086)	-0.001 (0.079)	-0.013 (0.075)	0.005 (0.089)	0.003 (0.083)
Ruggedness			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance to Waterway			0.000 (0.002)	0.000 (0.002)	-0.000 (0.001)	0.000 (0.001)	0.001 (0.003)	0.001 (0.002)
Colony						-0.076 (0.102)	0.015 (0.122)	0.026 (0.111)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	117	117	117	117	117	117	117	117
Adjusted <i>R</i> <sup>2</sup>	0.320	0.314	0.382	0.382	0.452	0.450	0.499	0.492
1st Stage <i>F</i> -statistic (K-P)								32.563

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and the population's average years of schooling. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.44: Predicted Diversity and Constraint on the Executive — Accounting for Population Density in 1500

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-3.824*** (0.839)	-4.178*** (0.849)	-3.270*** (0.877)	-3.270*** (0.877)	-3.522*** (0.821)	-3.507*** (0.827)	-3.526*** (1.312)	-6.268*** (2.180)
Population density in 1500 CE	0.013*** (0.003)	0.009*** (0.003)	0.006** (0.003)	0.006** (0.003)	0.003 (0.003)	0.003 (0.003)	0.003 (0.004)	0.003 (0.004)
Absolute Latitude		0.006*** (0.002)	0.006*** (0.002)	0.006*** (0.002)	0.011*** (0.002)	0.013*** (0.003)	0.008** (0.003)	0.009*** (0.003)
Agricultural Suitability			0.036*** (0.009)	0.036*** (0.009)	0.049*** (0.010)	0.050*** (0.010)	0.028** (0.011)	0.028*** (0.010)
Elevation			-0.054 (0.074)	-0.054 (0.074)	0.007 (0.082)	0.020 (0.080)	0.078 (0.098)	0.071 (0.093)
Ruggedness			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance to Waterway			0.004* (0.002)	0.004* (0.002)	0.003 (0.002)	0.002 (0.002)	0.001 (0.003)	0.001 (0.003)
Colony						0.081 (0.128)	0.166 (0.141)	0.177 (0.130)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	153	153	153	153	153	153	153	153
Adjusted <i>R</i> <sup>2</sup>	0.132	0.181	0.258	0.258	0.357	0.355	0.465	0.447
1st Stage <i>F</i> -statistic (K-P)								55.520

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and population density in 1500. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.



Table A.45: Predicted Diversity and Constraint on the Executive — Accounting for Social Infrastructure

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-2.944*** (0.709)	-3.680*** (0.760)	-3.123*** (0.736)	-3.123*** (0.736)	-3.522*** (0.688)	-3.491*** (0.714)	-3.089*** (1.111)	-6.297** (2.505)
Social Infrastructure	0.561*** (0.093)	0.280* (0.152)	0.221 (0.140)	0.221 (0.140)	0.127 (0.143)	0.140 (0.170)	0.067 (0.182)	-0.080 (0.189)
Absolute Latitude		0.007*** (0.002)	0.007*** (0.002)	0.007*** (0.002)	0.009*** (0.002)	0.009*** (0.003)	0.007** (0.003)	0.008*** (0.003)
Agricultural Suitability			0.030*** (0.009)	0.030*** (0.009)	0.033*** (0.009)	0.034*** (0.010)	0.019 (0.012)	0.020* (0.010)
Elevation			-0.064 (0.078)	-0.064 (0.078)	-0.056 (0.077)	-0.053 (0.079)	-0.079 (0.086)	-0.087 (0.081)
Ruggedness			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.001 (0.000)	0.001* (0.000)
Distance to Waterway			0.002 (0.002)	0.002 (0.002)	0.001 (0.002)	0.001 (0.002)	0.002 (0.003)	0.003 (0.003)
Colony						0.023 (0.159)	0.023 (0.188)	0.032 (0.167)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	116	116	116	116	116	116	116	116
Adjusted <i>R</i> <sup>2</sup>	0.250	0.299	0.374	0.374	0.376	0.370	0.431	0.397
1st Stage <i>F</i> -statistic (K-P)								22.397

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and a measure of social infrastructure. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.46: Predicted Diversity and Constraint on the Executive — Accounting for Ethnic Fractionalization

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-2.940*** (0.825)	-3.644*** (0.863)	-3.094*** (0.887)	-3.094*** (0.887)	-3.385*** (0.844)	-3.365*** (0.849)	-3.570*** (1.295)	-5.693** (2.255)
Ethnic Fractionalization	-0.403*** (0.121)	-0.202 (0.150)	-0.033 (0.151)	-0.033 (0.151)	0.008 (0.153)	0.010 (0.150)	-0.087 (0.150)	-0.071 (0.143)
Absolute Latitude		0.006*** (0.002)	0.007*** (0.002)	0.007*** (0.002)	0.012*** (0.002)	0.014*** (0.003)	0.008** (0.003)	0.010*** (0.003)
Agricultural Suitability			0.038*** (0.009)	0.038*** (0.009)	0.050*** (0.010)	0.051*** (0.011)	0.029*** (0.011)	0.029*** (0.010)
Elevation			-0.079 (0.076)	-0.079 (0.076)	-0.023 (0.082)	-0.005 (0.082)	0.070 (0.091)	0.063 (0.086)
Ruggedness			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance to Waterway			0.003 (0.002)	0.003 (0.002)	0.002 (0.002)	0.002 (0.002)	0.001 (0.003)	0.001 (0.003)
Colony						0.103 (0.124)	0.203 (0.143)	0.211 (0.131)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	154	154	154	154	154	154	154	154
Adjusted <i>R</i> <sup>2</sup>	0.109	0.143	0.231	0.231	0.325	0.325	0.438	0.428
1st Stage <i>F</i> -statistic (K-P)								53.682

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary executive constraints on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and a measure of ethnic fractionalization. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.47: Predicted Diversity and Constraint on the Executive — Accounting for Time Since the Neolithic Transition

	Constrain							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-3.605*** (0.852)	-4.402*** (0.899)	-3.350*** (0.941)	-3.350*** (0.941)	-3.615*** (0.904)	-3.591*** (0.910)	-4.040*** (1.229)	-7.178*** (2.515)
Time Since Neolithic Transition (in 10.000 Years)	-0.036 (0.160)	-0.327* (0.171)	-0.273 (0.169)	-0.273 (0.169)	-0.188 (0.174)	-0.175 (0.173)	0.143 (0.210)	0.332 (0.210)
Absolute Latitude		0.009*** (0.002)	0.008*** (0.002)	0.008*** (0.002)	0.013*** (0.002)	0.014*** (0.003)	0.008** (0.003)	0.010*** (0.003)
Agricultural Suitability			0.035*** (0.009)	0.035*** (0.009)	0.047*** (0.010)	0.048*** (0.011)	0.030*** (0.011)	0.030*** (0.010)
Elevation			-0.105 (0.078)	-0.105 (0.078)	-0.036 (0.091)	-0.020 (0.091)	0.054 (0.099)	0.051 (0.093)
Ruggedness			0.001 (0.000)	0.001 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Distance to Waterway			0.003 (0.002)	0.003 (0.002)	0.002 (0.002)	0.001 (0.002)	0.000 (0.002)	0.001 (0.002)
Colony						0.083 (0.121)	0.186 (0.142)	0.192 (0.130)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	151	151	151	151	151	151	151	151
Adjusted $R^2$	0.044	0.151	0.241	0.241	0.327	0.325	0.441	0.421
1st Stage $F$ -statistic ( $K$ - $P$ )								25.674

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and a measure of the time elapsed since the Neolithic Transition. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.48: Predicted Diversity and Autocracy (2013)

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	6.616*** (2.006)	7.289*** (2.044)	5.525*** (2.040)	5.525*** (2.040)	5.844*** (1.957)	5.872*** (1.967)	6.519** (3.114)	9.147** (4.652)
Absolute Latitude		-0.009*** (0.003)	-0.010*** (0.004)	-0.010*** (0.004)	-0.020*** (0.004)	-0.018** (0.007)	-0.008 (0.008)	-0.016** (0.007)
Agricultural Suitability			-0.074*** (0.021)	-0.074*** (0.021)	-0.099*** (0.022)	-0.098*** (0.023)	-0.056** (0.023)	-0.047** (0.020)
Elevation			0.000 (0.175)	0.000 (0.175)	-0.149 (0.180)	-0.135 (0.180)	-0.218 (0.204)	-0.020 (0.162)
Ruggedness			-0.000 (0.001)	-0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Distance to Waterway			-0.007 (0.004)	-0.007 (0.004)	-0.005 (0.004)	-0.006 (0.005)	-0.004 (0.005)	-0.004 (0.004)
Colony						0.090 (0.270)	-0.066 (0.301)	-0.148 (0.235)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
$N$	150	150	150	150	150	150	150	155
Adjusted $R^2$	0.041	0.072	0.137	0.137	0.241	0.237	0.325	0.462
1st Stage $F$ -statistic (K-P)								55.825

This table presents the results of a country-level OLS and 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.49: Predicted Diversity and Autocracy — Accounting for Irrigation Potential

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	7.842*** (1.816)	8.877*** (1.821)	7.043*** (1.890)	7.043*** (1.890)	7.417*** (1.822)	7.411*** (1.839)	7.198*** (2.728)	9.571** (4.679)
Percentage of Area Equipped for Irrigation	0.007 (0.013)	0.018 (0.015)	0.025 (0.016)	0.025 (0.016)	0.023* (0.012)	0.023* (0.012)	-0.019 (0.012)	-0.018 (0.012)
Absolute Latitude		-0.014*** (0.003)	-0.014*** (0.003)	-0.014*** (0.003)	-0.024*** (0.004)	-0.025*** (0.006)	-0.014* (0.007)	-0.015** (0.007)
Agricultural Suitability			-0.081*** (0.019)	-0.081*** (0.019)	-0.105*** (0.020)	-0.105*** (0.021)	-0.040* (0.021)	-0.040** (0.020)
Elevation			0.204 (0.152)	0.204 (0.152)	0.066 (0.163)	0.062 (0.163)	-0.035 (0.172)	-0.030 (0.161)
Ruggedness			-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Distance to Waterway			-0.007* (0.004)	-0.007* (0.004)	-0.005 (0.004)	-0.005 (0.005)	-0.004 (0.004)	-0.004 (0.004)
Colony						-0.023 (0.251)	-0.134 (0.247)	-0.145 (0.229)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	155	155	155	155	155	155	155	155
Adjusted <i>R</i> <sup>2</sup>	0.051	0.119	0.228	0.228	0.323	0.318	0.468	0.465
1st Stage <i>F</i> -statistic (K-P)								56.818

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables including a measure of land equipped for irrigation. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.50: Predicted Diversity and Autocracy — Accounting for Percentage of Land near a Waterway

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	6.594*** (1.879)	7.438*** (1.982)	6.123*** (2.008)	6.123*** (2.008)	6.875*** (1.792)	6.839*** (1.816)	6.582** (2.702)	8.433* (4.641)
Percentage of Land Near a Waterway	-0.715*** (0.170)	-0.601*** (0.181)	-0.603* (0.310)	-0.603* (0.310)	-0.482* (0.283)	-0.488* (0.285)	-0.228 (0.254)	-0.218 (0.238)
Absolute Latitude		-0.009*** (0.003)	-0.011*** (0.004)	-0.011*** (0.004)	-0.020*** (0.005)	-0.022*** (0.007)	-0.013* (0.007)	-0.014** (0.006)
Agricultural Suitability			-0.070*** (0.020)	-0.070*** (0.020)	-0.095*** (0.020)	-0.096*** (0.021)	-0.048** (0.021)	-0.048** (0.020)
Elevation			-0.208 (0.270)	-0.208 (0.270)	-0.291 (0.247)	-0.309 (0.246)	-0.266 (0.233)	-0.252 (0.218)
Ruggedness			0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Distance to Waterway			-0.012*** (0.004)	-0.012*** (0.004)	-0.010* (0.005)	-0.009 (0.006)	-0.004 (0.004)	-0.004 (0.003)
Colony						-0.080 (0.240)	-0.191 (0.252)	-0.197 (0.233)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
$N$	149	149	149	149	149	149	149	149
Adjusted $R^2$	0.158	0.187	0.243	0.243	0.347	0.343	0.486	0.484
1st Stage $F$ -statistic (K-P)								53.549

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables including a measure of land near a waterway. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.51: Predicted Diversity and Autocracy — Accounting for the Inequality in Land Suitability

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	8.495*** (1.830)	9.265*** (1.895)	7.627*** (2.007)	7.627*** (2.007)	8.061*** (1.891)	8.055*** (1.911)	7.477** (2.916)	9.061* (4.957)
Land Suitability Gini	0.425 (0.298)	0.523* (0.280)	-0.347 (0.380)	-0.347 (0.380)	-0.127 (0.407)	-0.125 (0.409)	-0.100 (0.359)	-0.100 (0.333)
Absolute Latitude		-0.014*** (0.003)	-0.014*** (0.004)	-0.014*** (0.004)	-0.023*** (0.004)	-0.024*** (0.007)	-0.014* (0.007)	-0.015** (0.007)
Agricultural Suitability			-0.092*** (0.022)	-0.092*** (0.022)	-0.110*** (0.023)	-0.110*** (0.024)	-0.056** (0.023)	-0.056*** (0.022)
Elevation			0.128 (0.203)	0.128 (0.203)	-0.050 (0.204)	-0.053 (0.200)	-0.156 (0.216)	-0.152 (0.202)
Ruggedness			-0.000 (0.001)	-0.000 (0.001)	0.001 (0.001)	0.001 (0.001)	0.000 (0.001)	0.000 (0.001)
Distance to Waterway			-0.006 (0.004)	-0.006 (0.004)	-0.006 (0.005)	-0.006 (0.005)	-0.002 (0.004)	-0.002 (0.003)
Colony						-0.017 (0.256)	-0.199 (0.272)	-0.207 (0.251)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	147	147	147	147	147	147	147	147
Adjusted $R^2$	0.081	0.155	0.226	0.226	0.327	0.322	0.478	0.477
1st Stage $F$ -statistic (K-P)								58.014

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables including a measure of inequality in land suitability. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.52: Predicted Diversity and Autocracy — Accounting for Percentages of Population Living in Various Climate Zones

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	8.176*** (1.961)	8.231*** (1.980)	8.371*** (1.969)	8.371*** (1.969)	9.081*** (1.781)	9.031*** (1.791)	7.843*** (2.764)	10.879** (5.353)
Absolute Latitude		0.002 (0.007)	0.004 (0.008)	0.004 (0.008)	-0.008 (0.009)	-0.011 (0.011)	-0.011 (0.010)	-0.013 (0.009)
Agricultural Suitability			-0.041* (0.024)	-0.041* (0.024)	-0.092*** (0.032)	-0.092*** (0.032)	-0.073** (0.031)	-0.079*** (0.030)
Elevation			-0.137 (0.195)	-0.137 (0.195)	-0.205 (0.181)	-0.245 (0.184)	-0.175 (0.201)	-0.156 (0.186)
Ruggedness			0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.000 (0.001)	0.000 (0.001)
Distance to Waterway			-0.006 (0.004)	-0.006 (0.004)	-0.006 (0.004)	-0.005 (0.005)	-0.005 (0.004)	-0.005 (0.004)
Colony						-0.160 (0.224)	-0.242 (0.254)	-0.263 (0.229)
Climate Zone Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	149	149	149	149	149	149	149	149
Adjusted $R^2$	0.290	0.286	0.285	0.285	0.374	0.372	0.483	0.478
1st Stage $F$ -statistic (K-P)								47.611

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables including variables capturing the percentage of the population living in various climate zones. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.



Table A.53: Predicted Diversity and Autocracy — Accounting for Colonizer Nation

	Log Autocracy						
	OLS						IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted Genetic Diversity	4.574*	5.033**	4.477*	4.477*	4.591*	7.444**	11.044**
	(2.506)	(2.437)	(2.619)	(2.619)	(2.602)	(3.267)	(5.454)
Absolute Latitude		-0.016***	-0.017***	-0.017***	-0.022***	-0.012*	-0.014**
		(0.003)	(0.004)	(0.004)	(0.006)	(0.007)	(0.006)
Agricultural Suitability			-0.080***	-0.080***	-0.085***	-0.042**	-0.042**
			(0.021)	(0.021)	(0.022)	(0.019)	(0.018)
Elevation			0.114	0.114	0.093	-0.005	0.001
			(0.159)	(0.159)	(0.159)	(0.176)	(0.166)
Ruggedness			-0.000	-0.000	-0.000	-0.001	-0.001
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Distance to Waterway			-0.007*	-0.007*	-0.005	-0.001	-0.001
			(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Colonizer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes	Yes
<i>N</i>	152	152	152	152	152	152	152
Adjusted $R^2$	0.103	0.166	0.258	0.258	0.259	0.424	0.417
1st Stage $F$ -statistic (K-P)							50.695

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and dummy variables indicating the colonizer nation (if any). The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.54: Predicted Diversity and Autocracy — Accounting for GDP

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	6.678*** (1.911)	7.574*** (2.201)	5.720** (2.219)	5.720** (2.219)	6.946*** (2.055)	6.904*** (2.076)	7.380*** (2.706)	9.440** (4.705)
Log Income per Capita in Year 2000	-0.191*** (0.055)	-0.129 (0.083)	-0.124 (0.082)	-0.124 (0.082)	-0.022 (0.081)	-0.024 (0.082)	0.008 (0.084)	0.007 (0.077)
Absolute Latitude		-0.007 (0.005)	-0.008 (0.005)	-0.008 (0.005)	-0.021*** (0.006)	-0.023*** (0.007)	-0.015** (0.007)	-0.017*** (0.006)
Agricultural Suitability			-0.085*** (0.020)	-0.085*** (0.020)	-0.103*** (0.020)	-0.105*** (0.021)	-0.049** (0.022)	-0.048** (0.021)
Elevation			0.082 (0.172)	0.082 (0.172)	0.042 (0.169)	0.022 (0.172)	-0.024 (0.172)	-0.021 (0.163)
Ruggedness			-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Distance to Waterway			-0.008** (0.004)	-0.008** (0.004)	-0.007* (0.004)	-0.006 (0.005)	-0.004 (0.004)	-0.004 (0.004)
Colony						-0.111 (0.252)	-0.188 (0.264)	-0.199 (0.243)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	154	154	154	154	154	154	154	154
Adjusted $R^2$	0.128	0.137	0.241	0.241	0.309	0.305	0.456	0.454
1st Stage $F$ -statistic (K-P)								56.319

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and GDP per capita in 2000. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.55: Predicted Diversity and Autocracy — Accounting for Schooling

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	6.913*** (1.744)	6.683*** (1.885)	6.086*** (1.889)	6.086*** (1.889)	5.426*** (1.917)	5.432*** (1.910)	7.579*** (2.265)	8.288** (4.218)
Years of Schooling	-0.121*** (0.019)	-0.128*** (0.034)	-0.097*** (0.034)	-0.097*** (0.034)	-0.128*** (0.031)	-0.129*** (0.031)	-0.093*** (0.034)	-0.093*** (0.031)
Absolute Latitude		0.002 (0.005)	-0.003 (0.005)	-0.003 (0.005)	-0.010* (0.005)	-0.006 (0.007)	-0.004 (0.007)	-0.004 (0.007)
Agricultural Suitability			-0.071*** (0.019)	-0.071*** (0.019)	-0.096*** (0.021)	-0.093*** (0.021)	-0.057*** (0.021)	-0.057*** (0.019)
Elevation			0.119 (0.198)	0.119 (0.198)	-0.026 (0.166)	0.004 (0.166)	0.001 (0.183)	0.001 (0.168)
Ruggedness			-0.000 (0.001)	-0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Distance to Waterway			-0.003 (0.003)	-0.003 (0.003)	-0.003 (0.004)	-0.004 (0.004)	-0.006* (0.003)	-0.006* (0.003)
Colony						0.180 (0.219)	0.079 (0.199)	0.073 (0.179)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	117	117	117	117	117	117	117	117
Adjusted <i>R</i> <sup>2</sup>	0.269	0.263	0.340	0.340	0.440	0.438	0.553	0.553
1st Stage <i>F</i> -statistic (K-P)								32.563

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and the population's average years of schooling. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.56: Predicted Diversity and Autocracy — Accounting for Population Density in 1500

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	8.049*** (1.901)	8.670*** (1.917)	6.904*** (1.974)	6.904*** (1.974)	7.262*** (1.823)	7.267*** (1.832)	7.021** (2.745)	10.007** (4.596)
Population Density in 1500 CE	-0.022*** (0.006)	-0.016*** (0.006)	-0.010* (0.006)	-0.010* (0.006)	-0.002 (0.007)	-0.001 (0.007)	-0.005 (0.008)	-0.005 (0.007)
Absolute Latitude		-0.010*** (0.003)	-0.011*** (0.004)	-0.011*** (0.004)	-0.023*** (0.005)	-0.022*** (0.007)	-0.013* (0.007)	-0.015** (0.007)
Agricultural Suitability			-0.074*** (0.019)	-0.074*** (0.019)	-0.103*** (0.020)	-0.103*** (0.021)	-0.043* (0.022)	-0.043** (0.021)
Elevation			0.137 (0.163)	0.137 (0.163)	-0.002 (0.168)	0.002 (0.168)	-0.062 (0.184)	-0.055 (0.173)
Ruggedness			-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)
Distance to Waterway			-0.010** (0.004)	-0.010** (0.004)	-0.007 (0.005)	-0.007 (0.005)	-0.005 (0.004)	-0.005 (0.004)
Colony						0.025 (0.257)	-0.076 (0.252)	-0.087 (0.233)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	153	153	153	153	153	153	153	153
Adjusted <i>R</i> <sup>2</sup>	0.110	0.143	0.218	0.218	0.320	0.316	0.473	0.468
1st Stage <i>F</i> -statistic (K-P)								55.520

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and population density in 1500. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.57: Predicted Diversity and Autocracy — Accounting for Social Infrastructure

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	6.882*** (1.657)	7.958*** (1.776)	6.895*** (1.723)	6.895*** (1.723)	7.853*** (1.583)	7.821*** (1.674)	6.979*** (2.434)	10.557** (5.223)
Social Infrastructure	-1.009*** (0.207)	-0.598 (0.362)	-0.476 (0.347)	-0.476 (0.347)	-0.192 (0.343)	-0.206 (0.390)	0.027 (0.405)	0.191 (0.422)
Absolute Latitude		-0.010* (0.005)	-0.011** (0.005)	-0.011** (0.005)	-0.015*** (0.005)	-0.016** (0.007)	-0.011 (0.007)	-0.012* (0.006)
Agricultural Suitability			-0.059*** (0.020)	-0.059*** (0.020)	-0.069*** (0.021)	-0.069*** (0.022)	-0.023 (0.023)	-0.023 (0.021)
Elevation			0.202 (0.192)	0.202 (0.192)	0.170 (0.179)	0.167 (0.182)	0.268 (0.175)	0.276* (0.161)
Ruggedness			-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)
Distance to Waterway			-0.003 (0.004)	-0.003 (0.004)	-0.003 (0.004)	-0.003 (0.005)	-0.003 (0.006)	-0.005 (0.006)
Colony						-0.025 (0.314)	-0.035 (0.313)	-0.046 (0.279)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	116	116	116	116	116	116	116	116
Adjusted <i>R</i> <sup>2</sup>	0.207	0.226	0.291	0.291	0.300	0.293	0.453	0.444
1st Stage <i>F</i> -statistic (K-P)								22.397

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and a measure of social infrastructure.. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.58: Predicted Diversity and Autocracy — Accounting for Ethnic Fractionalization

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	6.522*** (1.875)	7.779*** (1.916)	6.686*** (1.965)	6.686*** (1.965)	7.188*** (1.841)	7.179*** (1.854)	7.319*** (2.674)	9.172** (4.677)
Ethnic Fractionalization	0.699*** (0.238)	0.339 (0.290)	0.033 (0.303)	0.033 (0.303)	-0.062 (0.298)	-0.062 (0.297)	0.183 (0.287)	0.169 (0.270)
Absolute Latitude		-0.010*** (0.004)	-0.013*** (0.004)	-0.013*** (0.004)	-0.023*** (0.004)	-0.024*** (0.007)	-0.015** (0.007)	-0.016** (0.006)
Agricultural Suitability			-0.079*** (0.020)	-0.079*** (0.020)	-0.105*** (0.021)	-0.106*** (0.022)	-0.047** (0.021)	-0.047** (0.020)
Elevation			0.182 (0.173)	0.182 (0.173)	0.055 (0.176)	0.046 (0.177)	-0.050 (0.177)	-0.044 (0.166)
Ruggedness			-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Distance to Waterway			-0.008** (0.004)	-0.008** (0.004)	-0.007 (0.005)	-0.007 (0.005)	-0.004 (0.004)	-0.004 (0.004)
Colony						-0.049 (0.248)	-0.123 (0.257)	-0.130 (0.237)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	154	154	154	154	154	154	154	154
Adjusted <i>R</i> <sup>2</sup>	0.092	0.115	0.209	0.209	0.308	0.303	0.463	0.461
1st Stage <i>F</i> -statistic (K-P)								53.682

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and a measure of ethnic fractionalization. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.59: Predicted Diversity and Autocracy — Accounting for Time Since the Neolithic Transition

	Log Autocracy							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	7.843*** (1.850)	9.386*** (1.886)	7.437*** (1.992)	7.437*** (1.992)	7.864*** (1.879)	7.869*** (1.897)	7.447*** (2.615)	-7.178*** (2.515)
Time Since Neolithic Transition (in 10.000 Years)	0.357 (0.336)	0.921** (0.363)	0.794** (0.360)	0.794** (0.360)	0.648* (0.378)	0.651* (0.381)	-0.089 (0.447)	0.332 (0.210)
Absolute Latitude		-0.017*** (0.003)	-0.016*** (0.004)	-0.016*** (0.004)	-0.026*** (0.005)	-0.025*** (0.007)	-0.014* (0.007)	0.010*** (0.003)
Agricultural Suitability			-0.067*** (0.019)	-0.067*** (0.019)	-0.092*** (0.020)	-0.091*** (0.021)	-0.043** (0.021)	0.030*** (0.010)
Elevation			0.268 (0.164)	0.268 (0.164)	0.118 (0.179)	0.121 (0.179)	0.010 (0.186)	0.051 (0.093)
Ruggedness			-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	0.000 (0.000)
Distance to Waterway			-0.007* (0.004)	-0.007* (0.004)	-0.005 (0.004)	-0.005 (0.005)	-0.002 (0.004)	0.001 (0.002)
Colony						0.017 (0.241)	-0.099 (0.255)	0.192 (0.130)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	151	151	151	151	151	151	151	151
Adjusted <i>R</i> <sup>2</sup>	0.055	0.148	0.230	0.230	0.317	0.312	0.469	0.421
1st Stage <i>F</i> -statistic (K-P)								25.674

This table presents the results of a country-level 2SLS regression analysis of a measure of contemporary autocracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables and a measure of the time elapsed since the Neolithic Transition. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.60: Predicted Diversity and Democracy (1994–2013)

	Log Democracy							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	-7.829*** (1.910)	-8.807*** (1.924)	-6.786*** (1.974)	-6.786*** (1.974)	-7.252*** (1.824)	-7.275*** (1.831)	-7.421** (2.860)	-10.942** (5.076)
Absolute Latitude		0.015*** (0.003)	0.015*** (0.003)	0.015*** (0.003)	0.027*** (0.004)	0.025*** (0.007)	0.017** (0.007)	0.019*** (0.007)
Agricultural Suitability			0.085*** (0.020)	0.085*** (0.020)	0.113*** (0.020)	0.111*** (0.021)	0.062*** (0.022)	0.061*** (0.021)
Elevation			-0.180 (0.167)	-0.180 (0.167)	-0.020 (0.176)	-0.039 (0.174)	0.041 (0.191)	0.034 (0.180)
Ruggedness			0.001 (0.001)	0.001 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
Distance to Waterway			0.008** (0.004)	0.008** (0.004)	0.007 (0.005)	0.007 (0.005)	0.005 (0.005)	0.005 (0.005)
Colony						-0.108 (0.249)	0.038 (0.266)	0.055 (0.250)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	155	155	155	155	155	155	155	155
Adjusted <i>R</i> <sup>2</sup>	0.051	0.130	0.233	0.233	0.356	0.352	0.459	0.453
1st Stage <i>F</i> -statistic (K-P)								55.825

This table presents the results of a country-level OLS and 2SLS regression analysis of a measure of contemporary democracy on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.



Table A.61: Persistence of Institutions — Alternative Aggregation Method

	Log Executive Constraints			Log Autocracy		
	(1)	(2)	(3)	(4)	(5)	(6)
Indigenous Democracy	0.361*** (0.063)	0.256*** (0.077)	0.255*** (0.079)	-0.592*** (0.137)	-0.390** (0.169)	-0.388** (0.178)
Predicted Genetic Diversity			-3.760*** (1.347)			7.649*** (2.732)
Absolute Latitude		0.006 (0.004)	0.007* (0.004)		-0.008 (0.009)	-0.010 (0.009)
Agricultural Suitability		0.026** (0.010)	0.026*** (0.010)		-0.043** (0.022)	-0.043** (0.020)
Elevation		0.056 (0.091)	0.037 (0.094)		-0.007 (0.172)	0.032 (0.177)
Ruggedness		0.000 (0.000)	0.000 (0.000)		-0.000 (0.001)	-0.000 (0.001)
Distance to Waterway		0.000 (0.002)	-0.000 (0.003)		-0.002 (0.005)	-0.001 (0.005)
Temperature		-0.003 (0.009)	-0.006 (0.009)		0.010 (0.019)	0.016 (0.018)
Colony		0.191 (0.155)	0.232 (0.143)		-0.155 (0.282)	-0.238 (0.260)
Legal Origin FE	No	Yes	Yes	No	Yes	Yes
Continental FE	No	Yes	Yes	No	Yes	Yes
$N$	153	153	153	153	153	153
Adjusted $R^2$	0.144	0.450	0.479	0.088	0.453	0.480

This table presents the results of a country-level OLS regression analysis of measures of contemporary autocracy on a measure of pre-colonial democracy, conditional on sub-Saharan Africa, Latin America, and continental fixed effects. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.62: Predicted Diversity and Dictatorship (Linear Regression)

	Log Constraint on Chief Executive							
	OLS							IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted Genetic Diversity	5.883*** (1.095)	6.477*** (1.108)	5.560*** (1.170)	5.560*** (1.170)	5.633*** (1.175)	5.696*** (1.163)	4.993*** (1.763)	5.583** (2.725)
Absolute Latitude		-0.009*** (0.002)	-0.009*** (0.002)	-0.009*** (0.002)	-0.013*** (0.002)	-0.008** (0.004)	-0.003 (0.005)	-0.003 (0.004)
Agricultural Suitability			-0.036*** (0.010)	-0.036*** (0.010)	-0.047*** (0.010)	-0.044*** (0.011)	-0.017 (0.012)	-0.016 (0.011)
Elevation			0.095 (0.093)	0.095 (0.093)	0.030 (0.099)	0.080 (0.101)	0.060 (0.109)	0.061 (0.103)
Ruggedness			-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Distance to Waterway			-0.001 (0.004)	-0.001 (0.004)	-0.000 (0.003)	-0.002 (0.004)	0.003 (0.005)	0.003 (0.005)
Colony						0.272** (0.126)	0.287** (0.138)	0.284** (0.129)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes	Yes
<i>N</i>	158	158	158	158	158	158	158	158
Adjusted <i>R</i> <sup>2</sup>	0.096	0.174	0.233	0.233	0.277	0.292	0.355	0.354
1st Stage <i>F</i> -statistic (K-P)								54.036

This table presents the results of a country-level OLS and 2SLS regression analysis of a measure of contemporary dictatorship on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

Table A.63: Predicted Diversity and Dictatorship (Probit Regression)

	Dictatorship						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted Genetic Diversity	18.465*** (4.710)	19.559*** (4.536)	17.075*** (4.541)	17.075*** (4.541)	20.338*** (5.025)	20.280*** (4.914)	21.963*** (7.335)
Absolute Latitude		-0.024*** (0.007)	-0.026*** (0.007)	-0.026*** (0.007)	-0.047*** (0.011)	-0.030** (0.014)	-0.014 (0.017)
Agricultural Suitability			-0.111*** (0.034)	-0.111*** (0.034)	-0.172*** (0.043)	-0.160*** (0.043)	-0.053 (0.046)
Elevation			0.220 (0.277)	0.220 (0.277)	-0.086 (0.321)	0.136 (0.341)	-0.037 (0.359)
Ruggedness			-0.001 (0.001)	-0.001 (0.001)	0.000 (0.002)	-0.000 (0.002)	-0.001 (0.002)
Distance to Waterway			-0.001 (0.012)	-0.001 (0.012)	-0.000 (0.012)	-0.005 (0.012)	0.019 (0.021)
Colony						1.057** (0.497)	1.012** (0.514)
Legal Origin FE	No	No	No	No	Yes	Yes	Yes
Regional FE	No	No	No	No	No	No	Yes
<i>N</i>	158	158	158	158	149	149	149
Adjusted $R^2$							
1st Stage $F$ -statistic (K-P)							

This table presents the results of a country-level probit regression analysis of a measure of contemporary dictatorship on predicted population diversity as captured by the predicted ancestry-adjusted genetic diversity (i.e., a measure that reflects: (i) the proportional representation of the descendants of each ancestral population within a country, (ii) the predicted genetic diversity of each of these ancestral populations, and (iii) the predicted pairwise genetic distances between these ancestral populations), conditional on a range of geographical control variables. The 2SLS analysis uses migratory distance from East Africa to the capital city of each country as an instrumental variable for the predicted level of genetic diversity. Heteroscedasticity-robust standard errors are reported in parentheses. \*\*\* denotes statistical significance at the 1 percent level, \*\* at the 5 percent level, and \* at the 10 percent level, for two-sided hypothesis tests.

## B Additional Figures

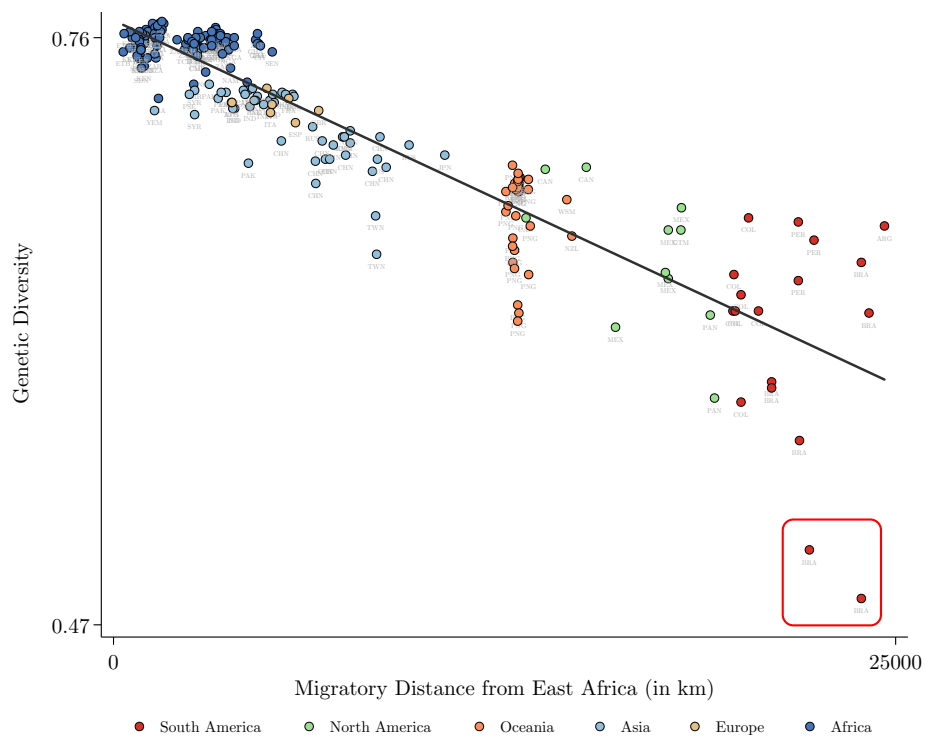


Figure B.1: This figure depicts the negative association between migratory distance from East Africa and genetic diversity across the 232 ethnic groups in the sample. It includes two ethnicities (the Surui and the Ache of South America, marked by the red square) that are largely viewed as extreme outliers in terms of genetic diversity (e.g. Wang et al., 2007).

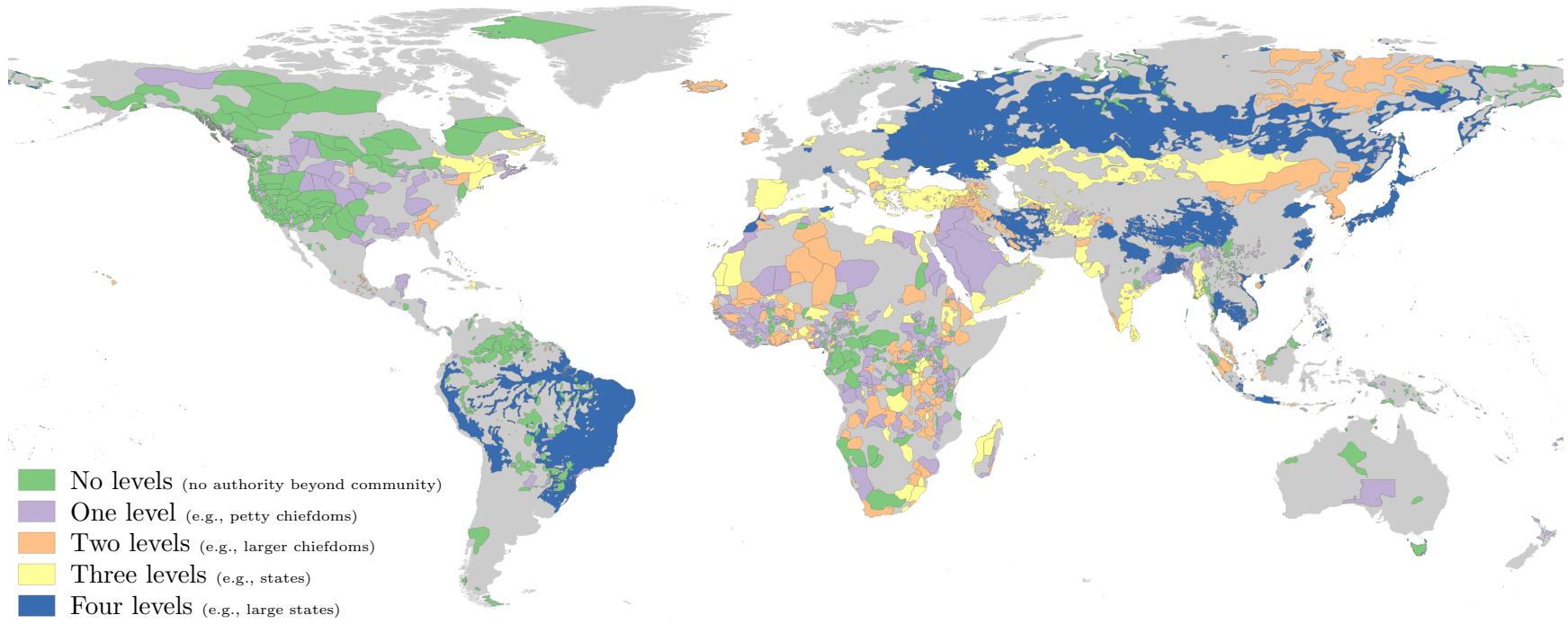


Figure B.2: The distribution of the number of levels of jurisdictional hierarchy beyond the local community across ethnic groups.

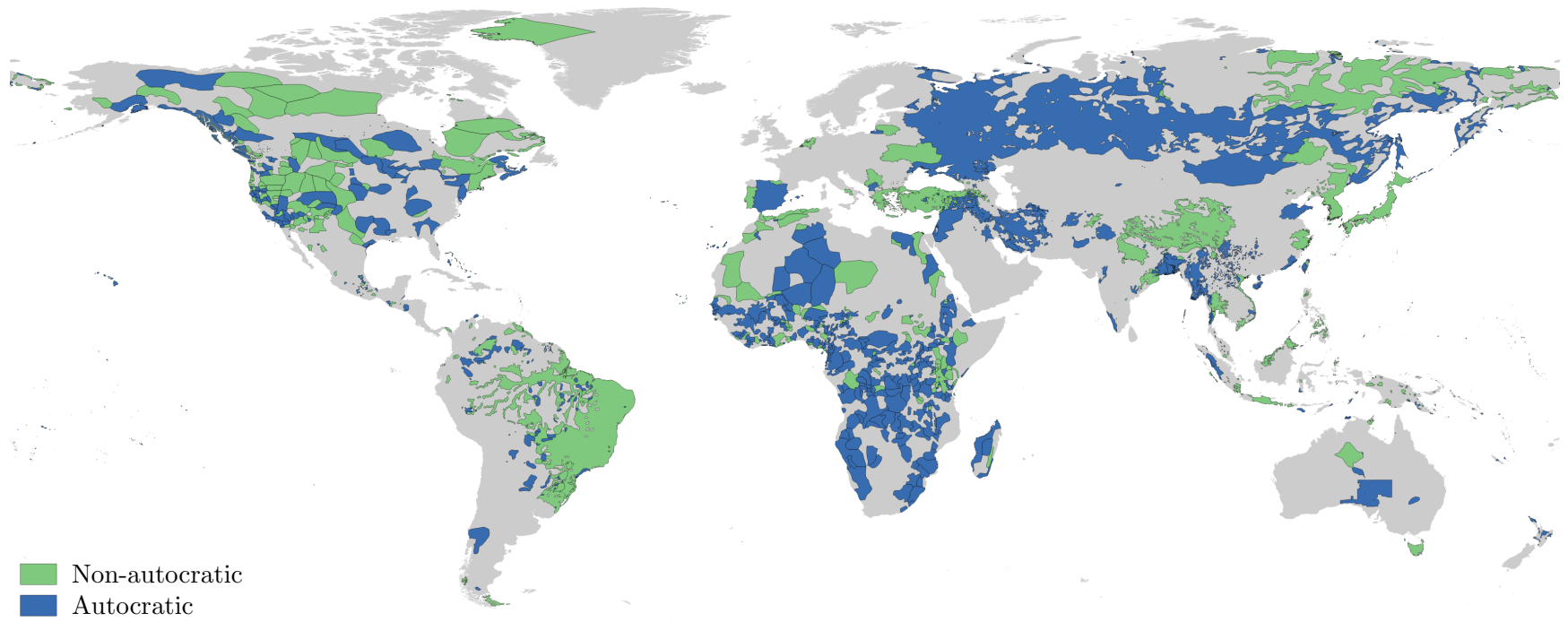
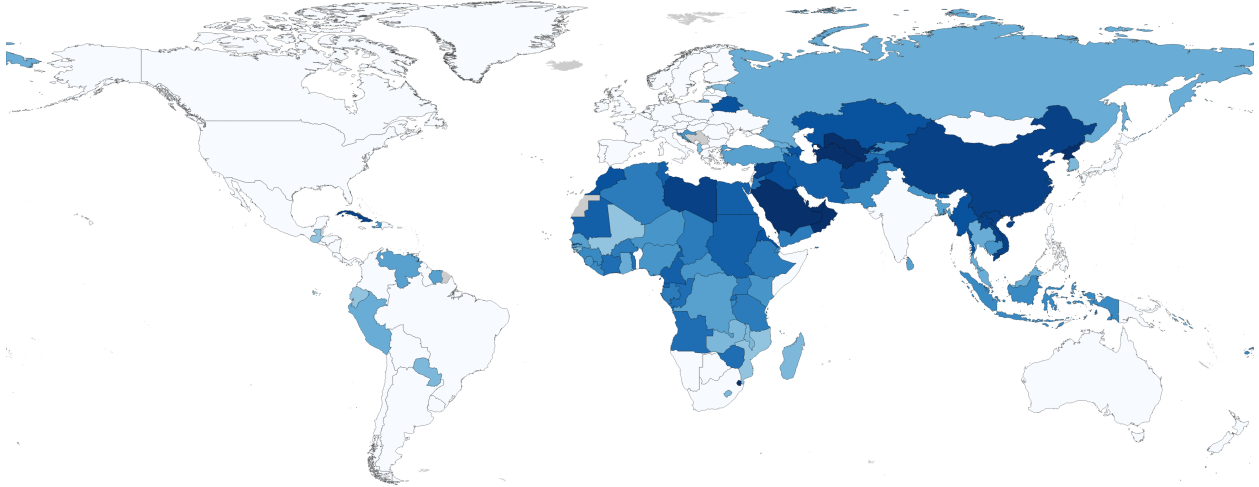


Figure B.3: The distribution of the indigenous autocracy index across ethnic groups.

Panel A: Intensity of Autocracy



Panel B: Income per Capita

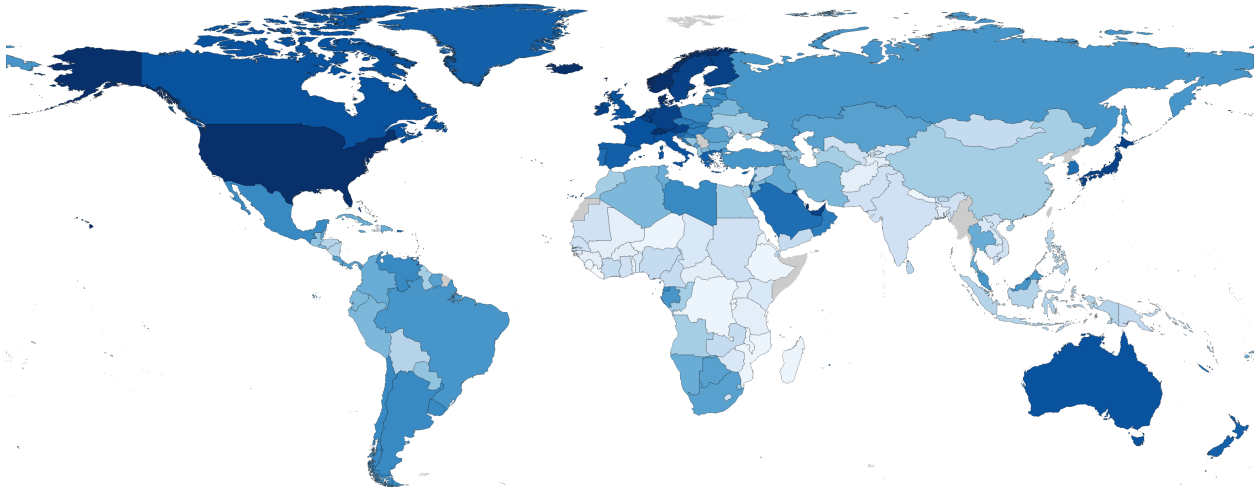


Figure B.4: The distribution of the average intensity of autocracy over the 1994–2013 period (Panel A) and the average income per capita over the 1994–2011 period (Panel B) across countries.

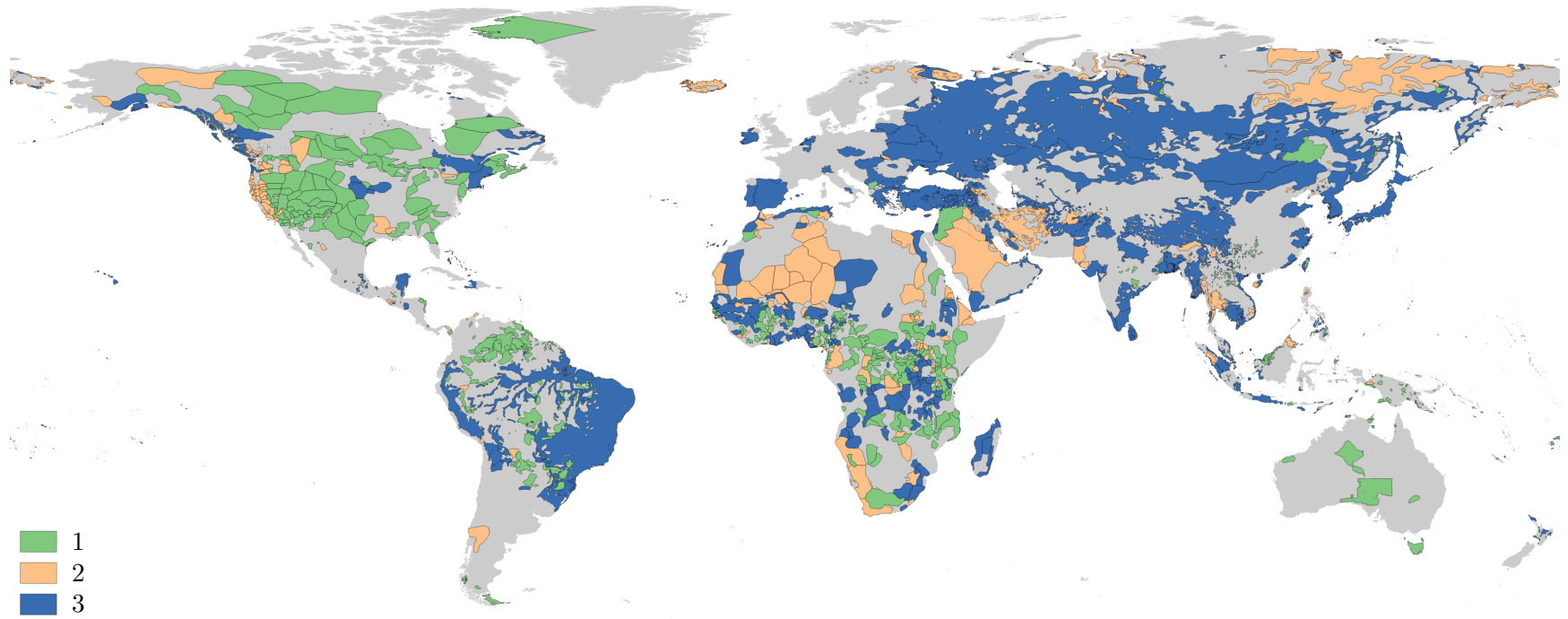


Figure B.5: The distribution of class stratification across ethnic groups.



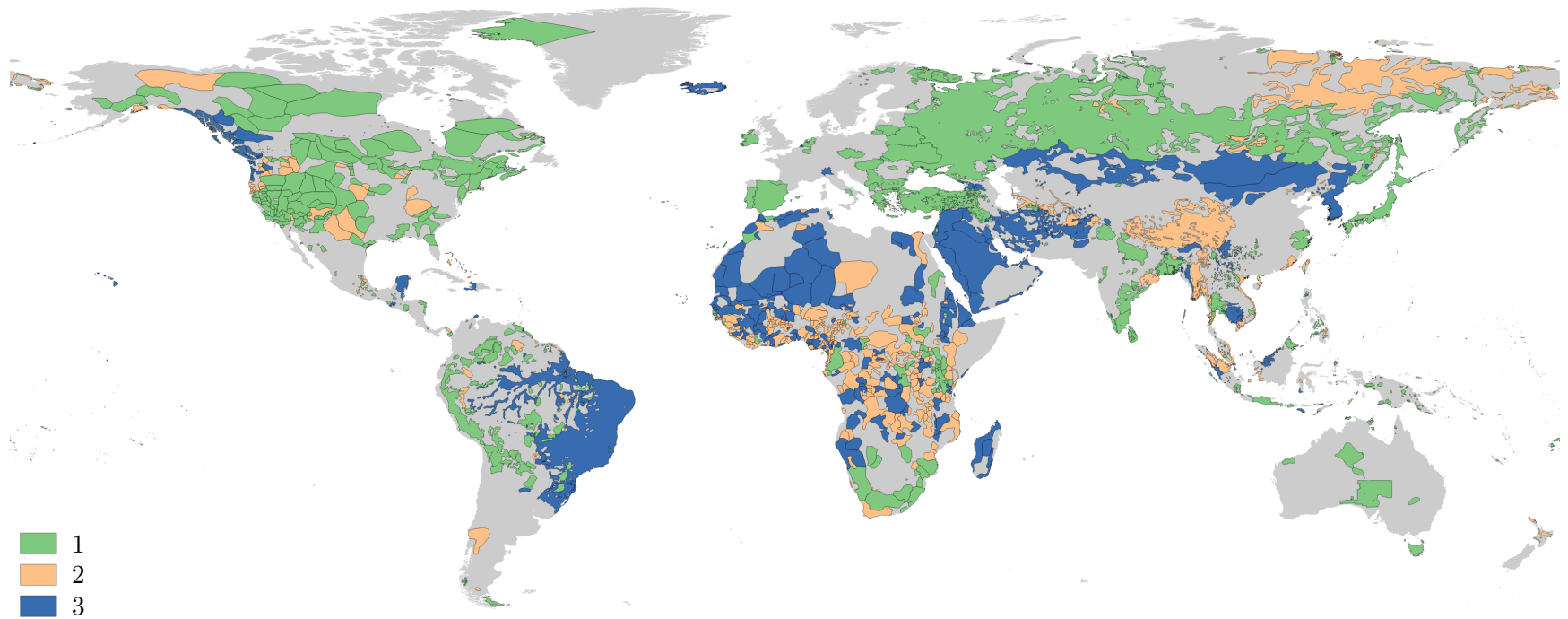
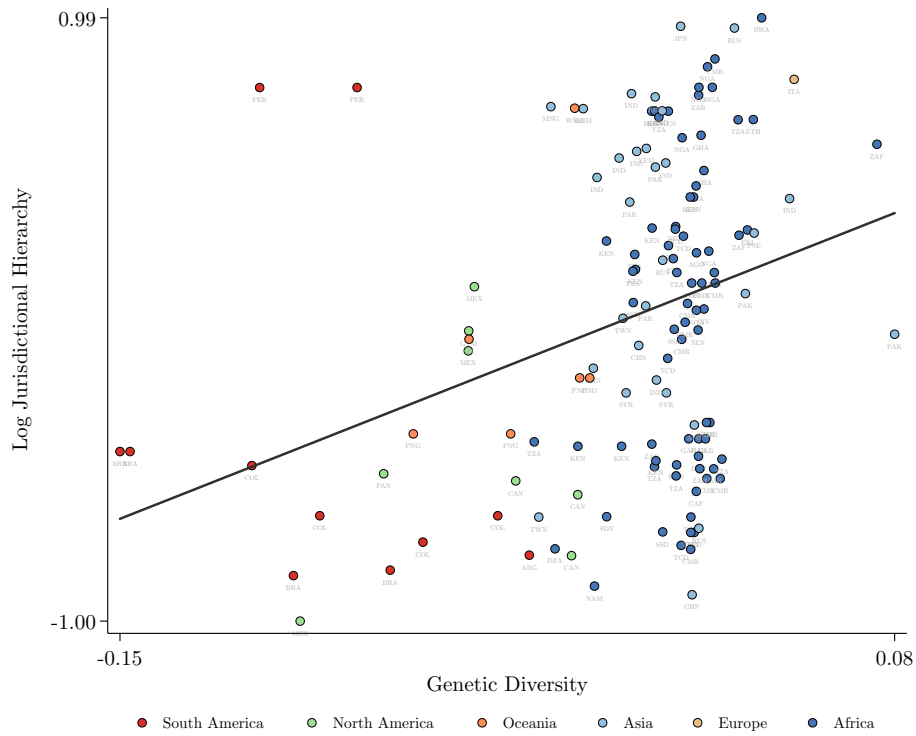
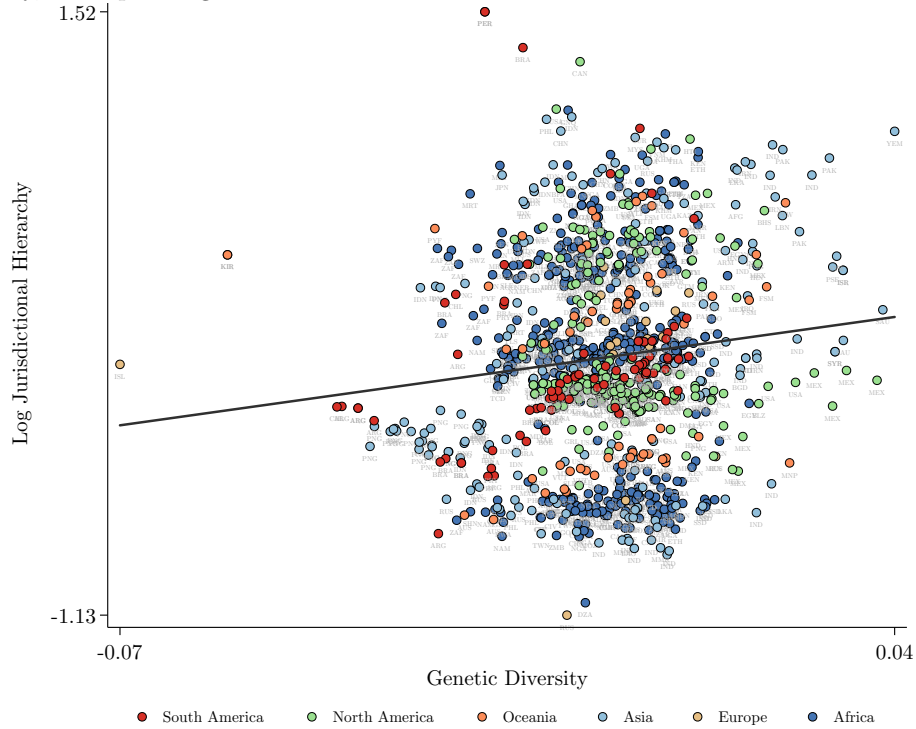


Figure B.6: The distribution of the intensity of slavery across ethnic groups.

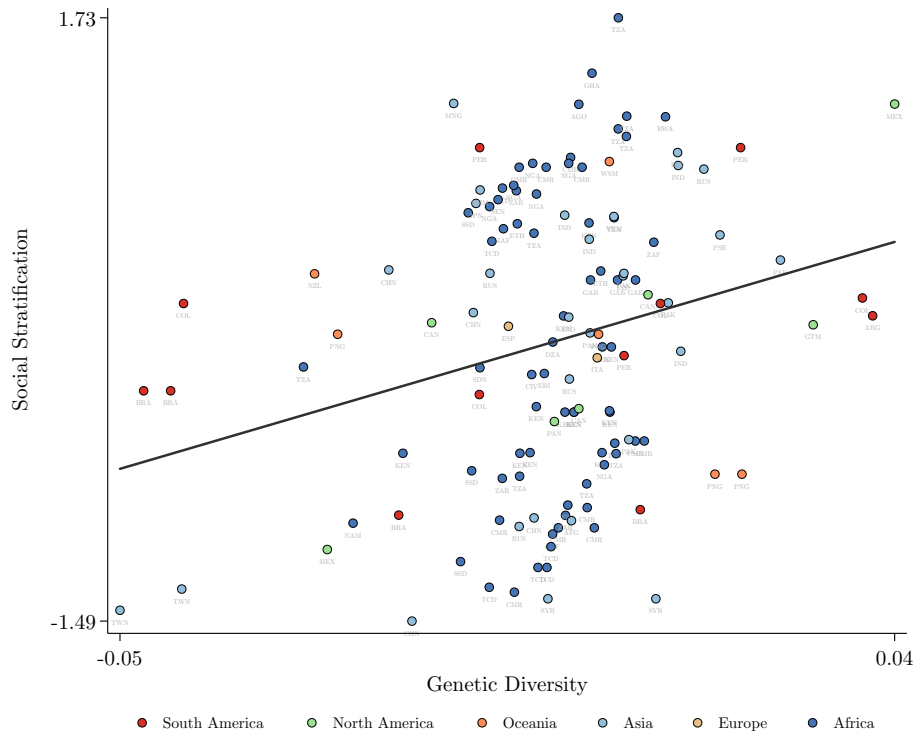


Panel A: The conditional association between genetic diversity and jurisdictional hierarchy, corresponding to column 5 of Table 1.

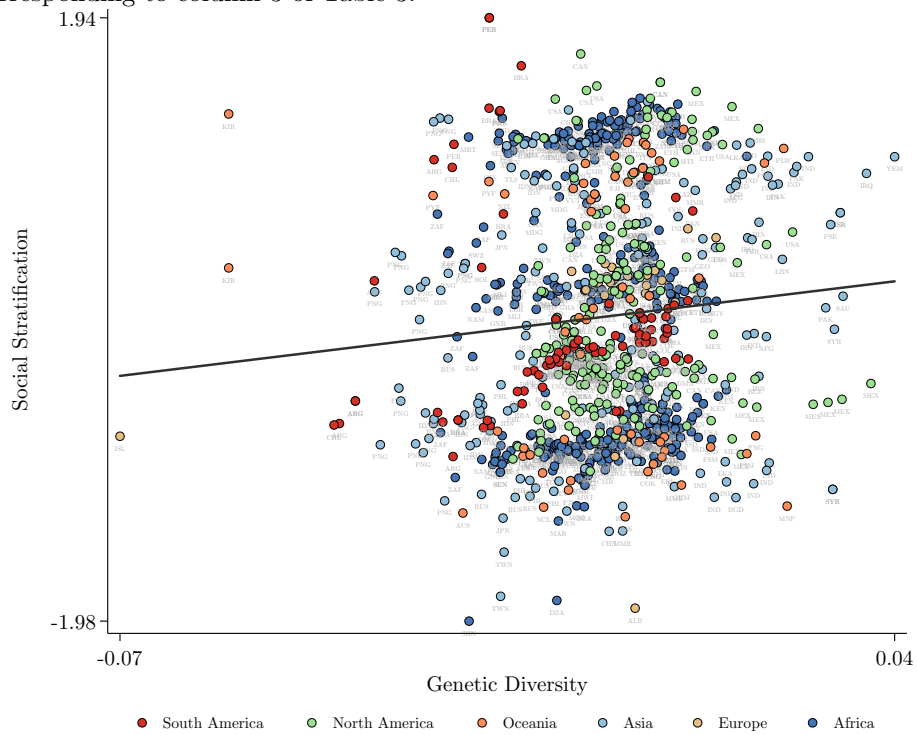


Panel B: The conditional association between predicted diversity and jurisdictional hierarchy, corresponding to column 6 of Table 2.

Figure B.7: Genetic diversity and jurisdictional hierarchy: Added variable plots.



Panel A: The conditional association between genetic diversity and social stratification, corresponding to column 3 of Table 3.



Panel B: The conditional association between predicted diversity and social stratification, corresponding to column 3 of Table 4.

Figure B.8: Genetic diversity and social stratification: Added variable plots.

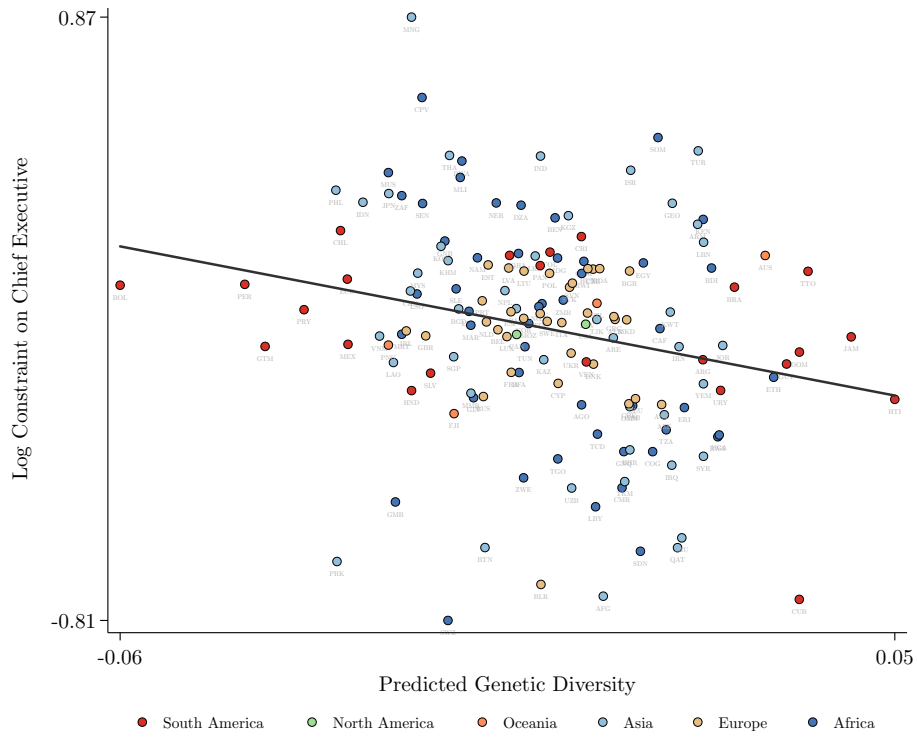


Figure B.9: The conditional association between predicted diversity and constraint on the executive, corresponding to column 7 of Table 10.

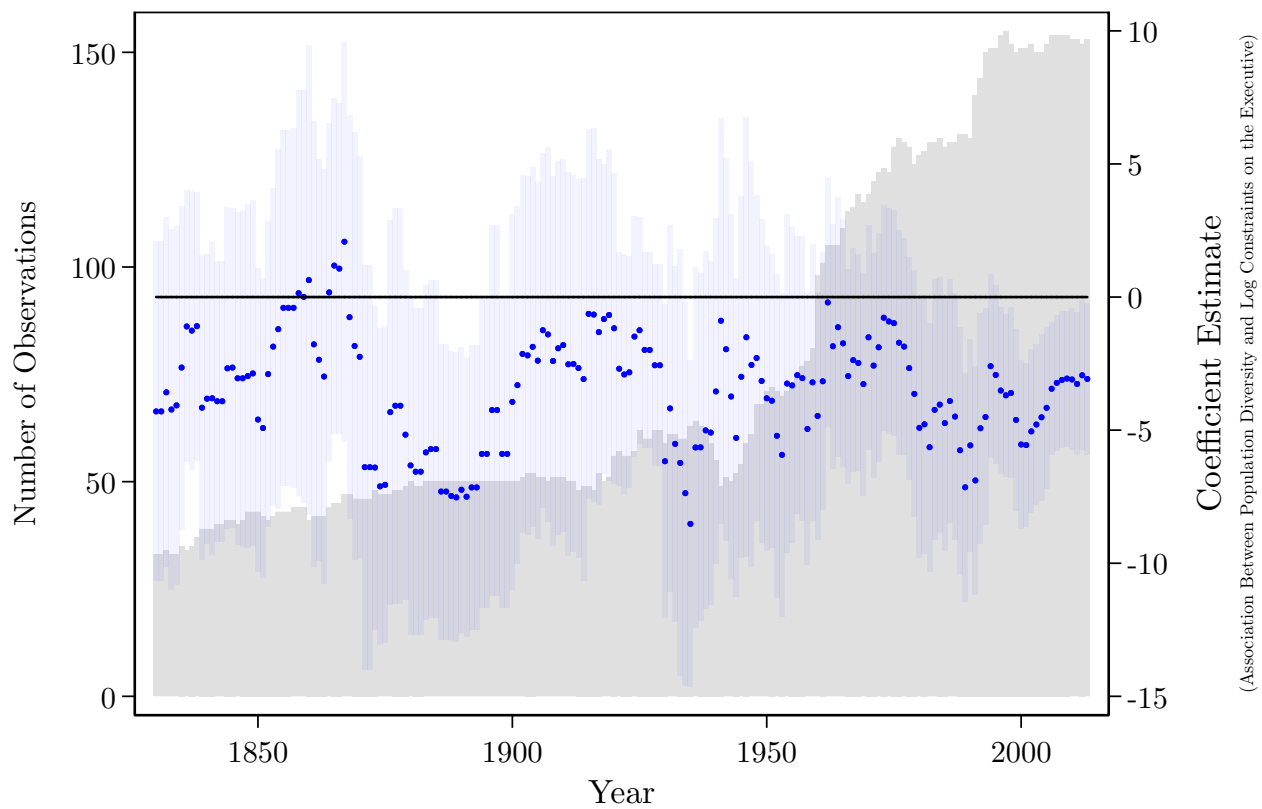


Figure B.10: The coefficient estimates (blue dots, right  $y$ -axis) and their 95% robust confidence intervals (blue shading, right  $y$ -axis) of the association between population diversity and the constraint on the executive measure, controlling for the baseline control variables, and the number of observations (gray bars, left  $y$ -axis), for each year in the period 1830–2013.

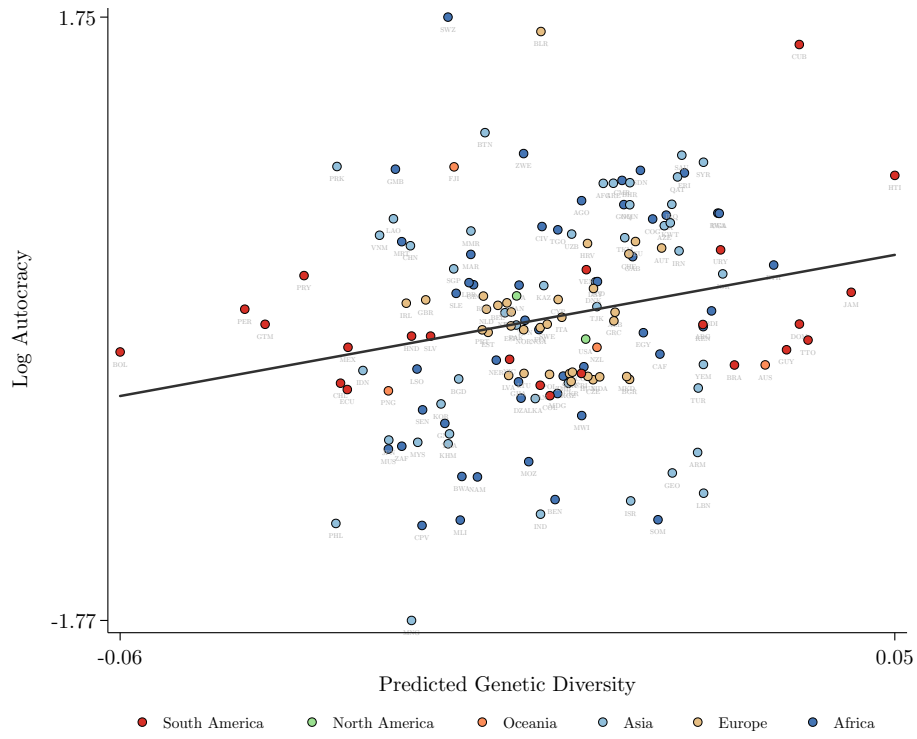


Figure B.11: The conditional association between predicted diversity and autocracy, corresponding to column 7 of Table 11.

## C Construction of Data Set

The novel geo-referenced data set of population diversity across ethnic groups is based on several sources. It links the measurements of observed genetic diversity of the 232 ethnic group (as provided Pemberton et al. (2013)), as well as the measurement of predicted diversity for the entire set of ethnic groups in the *Ethnographic Atlas* (as constructed in the current paper) to: (i) the geographical area of the historical homelands of these ethnic groups, (ii) the ethnographic characteristics of these ethnic groups, (as reported by the *Ethnographic Atlas* and the *Standard Cross-Cultural Sample*), and (iii) the geographical characteristics of the homelands of these ethnic groups.

The link between population diversity of each ethnic group and the geographical area of the historical homeland of these ethnic groups exploits several sources. Polygons for observations in the *Ethnographic Atlas* is based on Fenske (2013), who linked observations in the *Ethnographic Atlas* to the: (i) polygons found in Murdock et al. (1959), (ii) the Handbook of North American Indians (Heizer and Sturtevant, 1978), (iii) Global Mapping International’s (GMI) World Language Mapping System, (iv) the Geo-Referencing Ethnic Groups (GREG) map of Weidmann et al. (2010), and (v) data for modern administrative boundaries. We used the link between observations in the *Ethnographic Atlas* and James Fenske’s collection of polygons that was implied by the reported centroid coordinates in the data by Fenske (2013).

The matching process of observed population diversity for the 232 ethnic groups in Pemberton et al. (2013) was based on four phases. First, 65 observations from the Pemberton data was merged with name-based matches with the *Ethnographic Atlas* and via that to James Fenske’s polygons.<sup>29</sup> Second, the geocoded points of the ethnic groups reported in Pemberton et al. (2013) was overlaid with the map of James Fenske’s polygons and proximate pairs of polygons and points were classified as either separate, similar, or disparate groups, yielding 84 matches between polygons and points. Third, an additional 97 merges were achieved using a similar method with polygons from the GMI data set and their associated Ethnologue information. Fourth, for some remaining ethnic groups, a plausible polygon could be constructed based on secondary information about the ethnic group.<sup>30</sup>

---

<sup>29</sup>This matching process required the use of the various names given to each group in different sources.

<sup>30</sup>For instance, Tuscans were merged to the modern region of Toscana, Orcadians were merged to the South Orkney Islands, the Zenú were merged to the Zenú reserve, and the Sengwer were merged to the Embobut Forest area.

## D Definitions of Main Variables

This section describes the construction of the main variables.

### *Number of Levels of Jurisdictional Hierarchy*

This variable is based on variable 33 in the *Ethnographic Atlas*, denoted “Jurisdictional Hierarchy Beyond Local Community”. The Number of Levels of Jurisdictional Hierarchy variable takes on the value 1 when the original variable indicates “No levels (no political authority beyond community)”, 2 when it indicates “One levels (e.g., petty chiefdoms)”, 3 when it indicates “Two levels (e.g., larger chiefdoms)”, 4 when it indicates “Three levels (e.g., states)”, 5 when it indicates “Four levels (e.g., large states)”.

### *Genetic Diversity*

The data on Observed Diversity on the ethnic group level comes from the newly assembled data on Observed Diversity in 232 worldwide (predominantly indigenous) ethnic groups from Pemberton et al. (2013). The data on Predicted Diversity on the modern country level comes from (Ashraf and Galor, 2013).

### *Social Stratification*

This variable is based on variable 66 in the *Ethnographic Atlas*, denoted “Class Stratification”. The Social Stratification variable is grouped into the following categories. The variable takes on the value 0 when the original variable indicates “Absence among freemen”, 1 when it indicates “Wealth distinctions” or “Elite (based on control of land or other resources)”, and 2 when it indicates “Dual (hereditary aristocracy)” or “Complex (social classes)”.

### *Intensity of Slavery*

This variable is based on variable 70 in the *Ethnographic Atlas*, denoted “Type of Slavery”. The Intensity of Slavery variable is grouped into the following categories. The variable takes on the value 0 when the original variable indicates “Absence or near absence”, 1 when it indicates “Incipient or nonhereditary” or “Reported but type not identified”, and 2 when it indicates “Hereditary and socially significant”.

### *Indigenous Autocracy*

This variable is based on variable 72 in the *Ethnographic Atlas*, denoted “Succession to the Office of Local Headman”. The Indigenous Autocracy variable takes on the value 0 when the original variable indicates “Seniority or age, nonhereditary”, “Influence, wealth or social status, nonhereditary”, “Election or other formal consensus, nonhereditary”, “Informal consensus, nonhereditary”, or “Absence of any such office”. The variable takes on the value 1 when it indicates “Patrilineal heir”, “Matrilineal heir”, or “Appointment by higher authority, nonhereditary”.



### *Executive Constraints*

This variable is based on the Polity IV Project dataset (Marshall et al., 2014). The variable takes on an integer values from 1 to 7, indicating increasing extends of “institutionalized constraints on the decision-making powers of chief executives, whether individuals or collectivities” (Marshall et al., 2014).

### *Autocracy*

This variable based on the Polity IV Project dataset (Marshall et al., 2014). The variable takes on an integer value from 0 to 10, indicating increasing extends of “the presence of a distinctive set of political characteristics” characterizing autocracy. According to the definition, in their mature form, “autocracies sharply restrict or suppress competitive political participation” (Marshall et al., 2014).

### *Migratory Distance from East Africa*

In estimating the migratory distance from Addis Ababa (East Africa) for each of the ethnic groups in the data, the shortest traversable paths from Addis Ababa to the interior centroid of each ethnic group was computed. Given the limited ability of humans to travel across large bodies of water, the traversable area included bodies of water at a distance of 100km from land mass (excluding migration from Africa into Europe via Italy or Spain). Furthermore, for ethnicities that reside in a distance that exceed 100km from the traversable area connected to Addis Ababa, the distance was computed in the following way. A point set was created by clipping the extended traversable area to world boundaries and aggregating it to a resolution of 2,096,707 pixels which was then converted into points. For each ethnicity centroid, the nearest four distance points were identified and the great circle distance from the ethnicity centroid to those points were calculated. These distances was then added to the migratory distance from Addis Ababa at the distance point to obtain the total migratory distance from the ethnicity centroid from Addis Ababa to each of these four points. The point with the shortest total migratory distance from Addis Ababa was selected to represent the total migratory distance for the ethnicity.

### *Control Variables*

The control variables are based on a range of sources. For the analysis of the pre-colonial era, the developed geo-referenced dataset on within-ethnic-group genetic diversity and ethnographic information contains a wide range of variables. The data includes a range of geographic variables derived from a number of sources. These geographic variables include elevation, ruggedness, length and density of rivers in the area. Furthermore, the agricultural suitability variables are calculated as the average and standard deviations of the pre-1500 caloric suitability index constructed by Galor and Özak (2016). In addition, the average temperature, and average diurnal temperature range over the period 1901–2012 as constructed by the Climate Research Unit (see Harris et al., 2014). The irrigation

measure is based on the “area equipped for irrigation” data of the Global Map of Irrigation Areas, version 5.0 (Siebert et al., 2013).

*Regional Fixed Effects*

Dummy variables capturing location in either North America, Latin America, North Africa, sub-Saharan Africa, Asia, Europe, or Oceania.