No. 18/13

State History, Intergenerational Transmission and Institutional Trust in Africa

by

Gian Luca Tedeschi

Abstract

This research advances the hypothesis that intergenerationally transmitted beliefs act as a mechanism through which pre-colonial institutions influence modern day attitudes towards state figures in African countries. To address the question empirically, I combine contemporary individual-level survey data with historical data on pre-colonial political centralisation by ethnic group. By employing an identification strategy based on the current location of individuals and the disease environment of the ethnic groups' historical homelands, I establish that individuals belonging to ethnic groups that were characterised by a pre-colonial state show significantly higher levels of trust in current day institutions.

State History, Intergenerational Transmission and Institutional Trust in Africa

by

Gian Luca Tedeschi

Outline

The Author

Ph.D. candidate in Economics at the University of Nottingham.

E-mail: Gianluca.tedeschi@nottingham.ac.uk

State History, Intergenerational Transmission and Institutional Trust in Africa [*]

Gian Luca Tedeschi [†]

This research advances the hypothesis that intergenerationally transmitted beliefs act as a mechanism through which pre-colonial institutions influence modern day attitudes towards state figures in African countries. To address the question empirically, I combine contemporary individual-level survey data with historical data on pre-colonial political centralisation by ethnic group. By employing an identification strategy based on the current location of individuals and the disease environment of the ethnic groups' historical homelands, I establish that individuals belonging to ethnic groups that were characterised by a pre-colonial state show significantly higher levels of trust in current day institutions.

[†]School of Economics, University of Nottingham, University Park, Nottingham, NG7 2RD. E-mail: gianluca.tedeschi@nottingham.ac.uk.

I. INTRODUCTION

Since the late 19th century, virtually the entire world's land and population has been separated into states by borders. While most of human history has been characterised by stateless societies lacking centralised authorities, nowadays people are predominantly subject to a modern form of nation-state. To maintain authority, modern countries require recognition of the state's legitimacy by its citizens. Confidence in the institutions of the state is important for its effective functioning (Lipset, 1959; Almond and Verba, 1963; Easton, 1965; Putnam, 1993, 2000; Fukuyama, 1995). It is a commonly accepted view that confidence in state institutions and recognition of authority is determined by current political, economic and institutional factors. Over the last years, however, there has been ample research in economics documenting the influence of history on current economic and non-economic outcomes, with a large number of studies showing how historical factors can shape persistent cultural traits (Alesina et al., 2013; Giuliano and Nunn, 2017; Nunn and Wantchekon, 2011; Becker et al., 2016; Guiso et al., 2016). Is recognition of the state's legitimacy also a persistent trait shaped by historical events?

In this paper I advance the hypothesis that state history, i.e. for how long a population has experienced a state, shapes present attitudes towards state institutions through the mechanism of intergenerational transmission. I establish that modern day individuals whose ancestors experienced earlier exposure to a state show overall higher confidence in institutional figures. To guide my analysis, I develop a model, showing how reliance on tradition can cause beliefs to reflect experiences of past generations. The model predicts that, by relying on tradition, individuals with an ancestry of state history will be more likely to have positive attitudes towards the state. The idea builds on insights from the cultural anthropology literature, postulating that when information acquisition is either costly or imperfect, individuals employ heuristic decision-making strategies (Boyd and Richerson, 1985, 1995, 2005).

To address the question empirically, Africa provides an attractive setting, since several hundreds of relatively self-contained ethnic groups pursued different life-ways as recently as 5 generations ago. Descendants of those ethnic groups, with wildly varied ancestral backgrounds, now live together in the setting of modern African countries. I use data from the 2005 and 2008 Afrobarometer surveys to examine whether individuals belonging to an ethnic group that was politically centralised in the period preceding European colonisation show higher levels of trust in current-day state institutions. To measure pre-colonial political centralisation I employ data from the work of the anthropologist Murdock (1959, 1967), which mapped the spatial distribution of ethnic groups in Africa and quantified their economic, political and cultural characteristics. Specifically, by using the "Jurisdictional Hierarchy

Beyond the Local Community Level" index contained in Murdock's (1967) Ethnographic Atlas, I construct a measure categorising ethnic groups as either politically centralised or stateless.

In order to isolate a causal effect of the ethnic groups' state history on present day individuals' institutional trust, I address three significant challenges to identification. Of these challenges, the first relates to the fact that historical state formation may have also influenced current economic and institutional development which, in turn, may have influenced trust in those institutions. To distinguish intergenerational transmission from the effect of the current environment, I exploit the fact that cultural beliefs and values are internal to the individual. Murdock's (1959) ethnolinguistic map of Africa allows me to determine which area ethnic groups historically inhabited. Using geo-located data from the Afrobarometer surveys, I can determine whether respondents are still living in the homeland historically inhabited by their ethnic group. By examining individuals living outside of their ethnic group's historical homeland, I reduce the concern of the indirect effect of historical state formation working through current economic and institutional development.

The second empirical challenge I address relates to the existence of unobserved contemporary and historical ethnic characteristics driving the association I examine. To deal with this issue, first, I include an array of controls relating to individual characteristics and contemporary economic conditions. Second, I include country fixed effects, which account for time-invariant country-specific features. Third, I control for pre-colonial ethnic characteristics. Specifically, I take into account historic ethnic features contained in the Ethnographic Atlas, such as agricultural practices, settlement patterns and practice of slavery, and other ethnicity-level factors such as agricultural suitability of the historical ethnic homeland. In the context of this analysis, the most important omitted factor is possibly the colonial rule. Specifically, not taking into account how colonisation interacted with the pre-existing political structures I measure could severely bias my results. Among the sources of controversy regarding the Ethnographic Atlas is the fact that it may not capture ethnic groups' characteristics as they were prior to European contact. The issue, therefore, is whether the ethnographic records contained in the Atlas already capture some form of colonial influence. To deal with this potential issue, I also control for the year when the ethnographic records took place, as indicated in the Atlas.

The third empirical challenge relates to the fact that the measure of political centralisation I employ, as constructed from the Ethnographic Atlas, is an imprecise measure of the ethnic groups' state history. This could lead to attenuation bias. That is because the Atlas' variable captures the presence of a pre-colonially centralised government, but not how long it existed for, and this measurement error may therefore bias the results towards zero. To

3

overcome this, I exploit data on the historical disease environment and instrument my measure of pre-colonial political centralisation with an environmental feature that potentially shaped historical state formation in Africa, namely the presence of the TseTse fly. As Alsan (2015) argues, the presence of the TseTse fly affected African pre-colonial agricultural practices, patterns of subsistence, population density and the probability of ethnic groups being politically centralised. I thus employ a TseTse Suitability Index of the historical ethnic homeland, obtained by combining data on the suitability of the environment to the TseTse fly with Murdock's (1959) ethnolinguistic map of Africa, as an instrument for the ethnic group's pre-colonial centralisation. By using one of the factors determining whether ethnic homelands had favourable conditions for early state formation, I aim to retrieve the true effect of state history. The instrumental variable strategy will also allow me to deal with any residual hard-to-measure factors that may cause the centralisation measure to be endogenous.

I find that individuals from ethnic groups that were pre-colonially centralised show significantly higher levels of trust in current-day state institutions. This result is confirmed by both OLS and IV estimates, although OLS produces estimates that are biased towards zero. Results also hold when I focus on the sub-sample of individuals living outside the historical ethnic homeland of their group, producing estimates that are qualitatively identical to the full sample ones. Overall, findings are consistent with the initial hypothesis that long-run exposure to statehood produces a legacy of confidence in state institutions, through the mechanism of intergenerational transmission.

I then pursue a number of strategies to determine whether the results I document are valid and imply a causal relationship. First, I undertake a falsification exercise to test whether the TseTse presence instrument is correlated with omitted factors from the colonial rule. Drawing from the literature on how colonisation strategies were affected by settlers' mortality, I employ a measure of the prevalence of malaria in the historical ethnic homeland as instrument for the group's pre-colonial political centralisation. Results show how, in the context of this analysis, TseTse and Malaria are two fundamentally different aspects of the disease environment. Second, I test the robustness of results by changing the classification of pre-colonial political centralisation. Using Murdock (1967) original classification of political institutions, I find that estimation results are essentially unaltered. Third, I assess robustness to alternative measures of pre-colonial state formation. Using data from Chandler (1987) on the location of African cities in 1800, I employ a proxy indicating the presence of a large city on the area inhabited by the ethnic group. Estimation results from using this alternative measure are comparatively similar to the ones obtained using the pre-colonial political centralisation variable. Fourth, to account for the potential selection of individuals into more developed areas, I employ a series of current location fixed effects allowing me to obtain a within-location comparison of

individuals. These estimates again confirm the previous findings. Fifth, to establish the role played by the presence of co-ethics on ethnicity-related attitudes, I employ an interaction of the pre-colonial centralisation measure with the share of district's population that belongs to the same ethnic group as the respondent. Results reveal a heterogeneous effect of co-ethnics presence.

With this research, I build on different strands of the literature. Most relevant is the economics literature looking at the relationship between historical factors and cultural norms (see Nunn, 2012 and Gershman, 2017 for reviews). In particular, this article adds to a large number of empirical studies arguing for the persistence of cultural norms over long periods of time and how these norms act as a mechanism through which historical factors shape modern differences in economic outcomes (Nunn and Wantchekon, 2011; Alesina et al., 2013; Giuliano and Nunn, 2013; Nunn, 2008; Guiso et al., 2016; Becker et al., 2016; Voigtländer and Voth, 2012; Michalopoulos et al., 2016). My paper contributes to this line of work by uncovering a relationship between long-run exposure to statehood and present confidence in state institutions. I also contribute to a group of studies measuring cultural differences across societies with different methodologies, such as using lab experiments (Henrich, 2004; Henrich et al., 2005) or studying natural settings where people from different cultural backgrounds face the same decision in the same environment (Giuliano, 2007; Algan and Cahuc, 2010; Fernandez, 2007; Fernandez and Fogli, 2009; Giavazzi et al., 2014). In this article I study individuals with very diverse ethnic backgrounds as fellow citizens of modern African countries. Another relevant body of research to which I relate is the one studying the relationship between culture and institutions (Greif, 1994; Tabellini, 2008a; Tabellini, 2008b; Bisin and Verdier, 2017; Tabellini, 2010; Alesina and Giuliano, 2015), with this paper examining the long-run relationship between culture and institutions both theoretically and empirically. On a broader scale, this work relates to the literature on the historical institutional origins of contemporary development (Diamond, 1997; Acemoglu et al., 2001, 2002; Bockstette et al., 2002; Dell, 2010; Gennaioli and Rainer, 2006, 2007; Michalopoulos and Papaioannou, 2013, 2014).

The remainder of the paper is organised as follows. Section II develops a theoretical model to aid the analysis. Section III describes the data. Section IV discusses the estimating framework. Section V presents estimation results. Section VI provides a summary and concluding remarks.

II. Theoretical Model

This model shows how reliance on tradition can cause beliefs to reflect experiences of past generations, highlighting the mechanism through which descendants from ethnicities with a "state history" will be more likely to have positive attitudes towards the state. There are three periods of time: period 0 (pre-past), period 1 (past) and period 2 (present). The model is constituted by Individuals living in Societies, with Nature randomly drawing Societies' initial conditions and Individuals' types. There is a continuum of Individuals in each Society, and each period a new generation is born while the old one dies.

II.A. *Nature*

Nature exogenously determines the initial conditions e (environment) of each Society, with e_s being the environment of Society s. The environment e is chosen from a given distribution with PDF $g(.)$ and support $[\underline{\omega}, \overline{\omega}]$, with values of e tending to $\underline{\omega}$ indicating an environment more favourable to state formation and values of e tending to $\overline{\omega}$ indicating an environment less favourable to state formation. Once e_s is chosen by Nature, it stays the same for all periods. Nature also exogenously decides the type θ of Individuals, according to a binomial distribution with PMF $b(.)$. Individuals can be of two types, traditionalists (T) and information-seekers (I), with both types having an equal probability of being born. Individuals' types are not correlated across generations.

II.B. *Societies*

STATEHOOD. A Society can have two statehood types μ. It can either be stateless (0) or have a state (1), such that:

$$\mu \in \{0; 1\}$$

During period 0, all Societies are stateless, such that $\mu_0 = \{0\}$ for every s. Societies in period 1 can either have a state (1) or be stateless (0), such that $\mu_1 \in \{0; 1\}$. The Society's environment e_s determines of which type Society s is going to be at the beginning of period 1. There is a threshold Δ such that if the value of e_s is below (above) this threshold, then Society s is going to have a state (be stateless), therefore:

$$\mu_1 = \begin{cases} 1 & if \ e_s < \Delta \\ 0 & if \ e_s \geq \Delta \end{cases}$$

In period 2, there is a shock happening for each Society (one possible interpretation being

6

colonialism), such that every Society receives a state, i.e. $\mu_2 = \{1\}$ for every s. An example of the possible evolution of two Societies' statehood μ is given by:

	μ of Society i	μ of Society j
Period 0	0	0
Period 1	1	0
Period 2	1	1

GOODNESS OF STATE.   There are two types η of state, good (G) and bad (B):

$$\eta \in \{G; B\}$$

At the beginning of period 1 both good and bad states can emerge, however only good states survive through the rest of the period. This is due to the fact that, outside of period 2, bad states are unsustainable and experience a collapse. In case a bad state emerges, the Society will therefore revert back to being stateless, and stay like that for the rest of period 1. Thus, in period 1 Societies can either be stateless ($\mu_1 = 0$) or have a good state ($\mu_1 = 1$, $\eta_1 = G$).[1]

During period 2 all Societies become states regardless of their type in the previous period. Differently from period 1, both good and bad states can exist in period 2, i.e. $\eta_2 \in \{G; B\}$.

	Period 0	*Period 1*		*Period 2*
Statehood (μ)	0	0	1	1
Goodness of state (η)	-	-	G	G/B

In case a Society was already a state in period 1 ($\mu_1 = 1$), then the goodness of the state in

---

[1]The period 1 assumption of bad states always collapsing and good states always surviving could be relaxed, for example with both types of state having a positive probability to collapse and bad states collapsing with a higher probability than good states. Given that, in the context of this model, relaxing the assumption would yield the same results intuition-wise, the current extreme assumption is chosen for the sake of simplicity.

7

period 2 (η_2) is correlated with the goodness of the state in period 1 (η_1) by a factor $\rho > 0$. In other words, since only good states can survive in period 1, a Society that is a state in both periods has a higher chance to be a good state in period 2, compared to a Society that was stateless in period 1, therefore:

$$Pr(\eta_2 = G \mid \mu_1 = 1, \mu_2 = 1) > Pr(\eta_2 = G \mid \mu_1 = 0, \mu_2 = 1)$$

II.C. *Individuals*

There are three generations of Individuals, with each generation living in one of the three periods. Individuals can have two kinds of beliefs σ, either that the state is good (P) or that the state is bad (N), such that:

$$\sigma \in \{P; N\}$$

There are two types θ of Individuals, traditionalists (T) and information-seekers (I), such that:

$$\theta \in \{T; I\}$$

The type θ of an Individual is independent of his parents' type in the previous generation, therefore it is possible for a traditionalist to descend from an info-seeker and vice-versa. Info-seekers form their beliefs based on observation and judgement, therefore based on the actual goodness of the state η. If no information is available about the goodness of the state, info-seeker will form a belief that the state is bad.[2] The different cases are:

i. Info-seekers ($\theta = I$) living in a good state ($\eta = G$) will form a belief that the state is good ($\sigma = P$);

ii. Info-seekers ($\theta = I$) living in a bad state ($\eta = B$) or stateless Society ($\mu = 0$) will form a belief that the state is bad ($\sigma = N$);

Traditionalists, on the other hand, are unable to acquire information. Instead, they employ a heuristic decision-making strategy where they imitate the belief formed by their previous generation. Their belief is independent of the existence or goodness of the state. Some examples are:

i. A traditionalist ($\theta = T$) living in a good state ($\eta = G$) but descending from someone with a bad belief of the state ($\sigma_{t-1} = N$) will also believe that the state is bad ($\sigma_t = N$);

---

[2]Similarly to the assumption regarding the collapse of states in period 1, one could relax the assumption on the Individuals' prior belief of the state, without changing the model's results. A more generalised version would have the prior belief being that the state is bad only with some positive probability.

ii. A traditionalist ($\theta = T$) living in a bad state ($\eta = B$) but descending from someone with a good belief of the state ($\sigma_{t-1} = P$) will also believe that the state is good ($\sigma_t = P$);

Individuals from the first generation, living in period 0, can be both info-seekers or traditionalists. However, regardless of their type θ all Individuals will have a belief that the state is bad ($\sigma = N$), given that during period 0 all Societies are stateless ($\mu_0 = 0$) and no information is available about states and their goodness. After the emergence of states is determined by the environment e at the beginning of period 1, Individuals from the second generation form their beliefs. If a Society remains stateless in period 1, info-seekers from that generation will believe that the state is bad. In the case a good state emerges and survives, info-seekers will update and form a belief that the state is good. Traditionalists, on the other hand, will not update their beliefs from the previous generation and keep believing that the state is bad, regardless of what kind of Society turns out in period 1, since they can only imitate bad beliefs about the state. In period 2, when all Societies become states, which can be either good or bad, Individuals from the third generation form their beliefs. Info-seekers will have a good (bad) belief if the Society they live in has a good (bad) state. Traditionalist, similarly to the ones from the previous period, form good or bad beliefs by imitating their parents.

II.D. *Mechanism and Prediction*

The outcomes for each period are:

Period 0: Both info-seekers ($\theta = I$) and traditionalists ($\theta = T$) believe that the state is bad ($\sigma = N$).

Period 1: Info-seekers ($\theta = I$) can either have good ($\sigma = P$) or bad ($\sigma = N$) beliefs about the state, depending on whether they live in a state ($\mu_1 = 1$) or not ($\mu_1 = 0$). Traditionalists ($\theta = T$), on the other hand, can only have bad beliefs ($\sigma = N$) about the state in this period.

Period 2: Info-seekers ($\theta = I$) can either have good ($\sigma = P$) or bad ($\sigma = N$) beliefs about the state, depending on whether they live in a good state ($\eta_2 = G$) or in a bad state ($\eta_2 = B$). Traditionalists ($\theta = T$) can either have good ($\sigma = P$) or bad ($\sigma = N$) beliefs about the state, depending on who they descend from ($\sigma_t = \sigma_{t-1}$). A period 2 traditionalist will believe the state is bad if descending from a period 1 traditionalist or from a period 1 info-seeker who lived in a stateless Society. A period 2 traditionalist will believe the state is good if and only if descending from a period 1 info-seeker who lived in a good state.

Based on the two assumptions that traditionalists have beliefs that reflect those of the

previous generation and that the cross-period correlation in goodness of state ρ is greater than zero, the model predicts that Individuals from Societies that were characterised by a state for longer will be more likely to have a good belief of the state. In other words, Individuals from Societies with a "state history", i.e. Societies characterised by a state in both periods 1 and 2, will be more likely to believe that the state is good in period 2, because of these two reasons:

1. Due to the cross-period correlation in goodness of state ρ, societies having states in both periods 1 and 2 will be more likely to be good states in period 2, making info-seekers more likely to have a good belief.

2. Due to the traditionalists' heuristic decision-making, period 2 traditionalists in societies with a "state history" will have a chance of descending from a period 1 info-seeker with a good belief, therefore making traditionalists more likely to have a good belief.

Based on the mechanism of tradition, I formulate a general hypothesis that I bring to the data.

Hypothesis *Reliance on tradition causes beliefs to reflect experiences of past generations. The presence of traditionalists makes descendants from societies with a "state history" more likely to believe that the state is good.*

The other mechanism explained above also allows me to highlight a challenge I will address in the analysis, i.e. the need to distinguish the effect working through the traditionalists' beliefs from the effect working through the cross-period correlation ρ. I now turn to the empirical analysis, where I test for the persistence of beliefs associated with the state.

III. Data Sources and Description

III.A. *Afrobarometer Surveys*

The individual-level data are from the third and fourth rounds of the Afrobarometer surveys. Afrobarometer is an independent and non-partisan research project conducted by the Centre for Democratic Development (CDD, Ghana), the Institute for Democracy in South Africa (IDASA), the Institute for Empirical Research in Political Economy (IREEP) with support from Michigan State University (MSU) and the University of Cape Town, Center of Social Science Research (UCT/CSSR).[3] Implemented by national partners, Afrobarometer measures economic conditions and the political atmosphere in African countries. The questionnaire is

---

[3]This information is taken from the Afrobarometer web site at Afrobarometer.org.

standardized to facilitate comparison between the countries covered. The surveys are based on interviews conducted in the local languages of a random sample of between 1,200 and 2,400 people per country.[4] I pool two round of the Afrobarometer surveys: the third (2005) and fourth (2008-2009).[5] The Afrobarometer third round of surveys covers the following 18 countries: Benin, Botswana, Cape Verde, Ghana, Kenya, Lesotho, Madagascar, Malawi, Mali, Mozambique, Namibia, Nigeria, Senegal, South Africa, Tanzania, Uganda, Zambia, Zimbabwe. The fourth round of surveys covers the same countries, as well as Burkina Faso and Liberia. In total, there is a potential sample of 53,110 respondents from the two pooled rounds. A map of the 20 countries included in the forth round of surveys is shown in Figure A.I.

For a number of individuals it is not possible to match the reported ethnicity to the historical data, the reason being either (i) respondents listed "other" as their ethnicity, (ii) respondents listed their country as their ethnicity, (iii) the ethnicity is not an indigenous African ethnicity, (iv) the ethnicity could not be matched to the historical sources or (v) the ethnicity could be matched but no historical data is available for that ethnic group. These observations were dropped from the analysis. Additionally, for Cape Verde Islands the ethnicity of the respondent is not recorded in the surveys. Finally, due to the lack of data on TseTse suitability (see section III.C) for Madagascar, this country is also dropped in order to keep the upcoming analysis balanced. The aforementioned data limitations leave a potential sample of 37,287 respondents.

In order to measure popular attitudes toward the state, I consider the reported trust with respect to four institutional figures: the president, the parliament, the police and courts of law. For these trust questions, respondents were asked "How much do you trust each of the following? The President / The Parliament / The Police / Courts of Law". The answer categories were (i) "Not at all", (ii) "Just a little", (iii) "Somewhat", and (iv) "A lot". Respondents had the option of answering "Don't know". Furthermore, for some respondents the answer is reported as "Missing". Respondents with "Don't know" or "Missing" answers are dropped from the sample. Figure I illustrates the distribution of responses to each question. By using these four variables the intention is to consider the different powers of the state. In order to capture a common underlying determinant, I aggregate these four measures by employing Principal Component Analysis (PCA) and extract the first principal component. I normalise the resulting variable to be in $[0,1]$ and label it as Institutional Trust Index (ITI). This index constitutes the dependent variable which I will use in the empirical analysis.

---

[4]The minimum sample of 1,200 people gives a margin of error of 3% and a degree of confidence of 95%.

[5]The two earlier rounds were excluded as they do not have information on the ethnicity of respondents making it impossible to match individuals to ethnic group data. Later rounds are in process of being added.

Appendix A provides all the details concerning this index. Due to missing values in the trust questions, the Institutional Trust Index is available for 33,419 respondents.
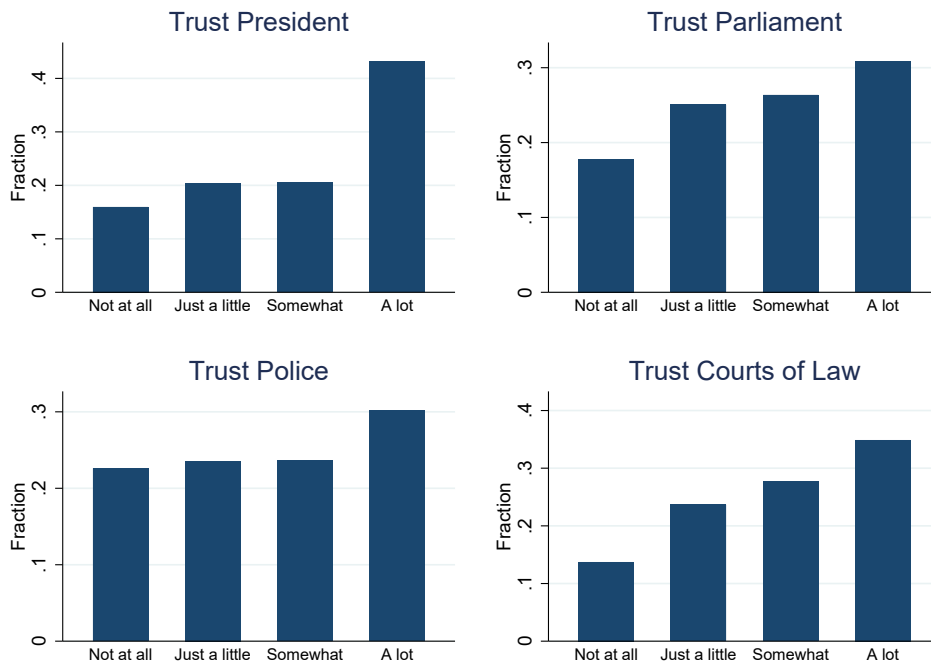
FIGURE I. Distribution of responses to trust questions from the Afrobarometer surveys

III.B. *Pre-Colonial Political Centralisation*

I employ historical data on African ethnic groups based on the work of the anthropologist George Peter Murdock. Figure II(A) shows Murdock's (1959) ethnolinguistic map of Africa, which portrays the spatial distribution of ethnicities across Africa at the beginning of European colonisation in the mid/late 19th century, including 843 historical ethnic homelands.

The publicly available data by Nunn and Wantchekon (2011) and Deconinck and Verpoorten (2013) allowed me to match the ethnicities as reported by respondents in the Afrobarometer surveys with the classification of ethnic groups constructed and mapped by Murdock (1959). In his work following the mapping of African ethnicities, Murdock (1967) produced an Ethnographic Atlas that coded approximately 60 variables, capturing cultural, geographical, and economic characteristics of 1,265 ethnicities around the world, of which 534 in Africa. Given that there is not a perfect match for all observations between the ethnolinguistic map and the Ethnographic Atlas, I use the algorithm developed by Fenske (2013), which

joins unmatched ethnic groups based on alternative name, supergroup or location.[6] The pre-colonial political centralisation measure I employ in the analysis is based on Murdock's (1967) "Jurisdictional Hierarchy Beyond the Local Community Level" index, which is the standard variable referred to in the literature as a proxy for the institutional development of historical ethnic groups (Gennaioli and Rainer, 2006; Michalopoulos and Papaioannou, 2013). This ordered variable classifies ethnic groups into five categories, ranging from 0 to 4, indicating the number of political jurisdictions beyond the local level. Jurisdictional hierarchy is coded so that a score of zero indicates pre-colonial stateless societies "lacking any form of centralised political organization", a score of one indicates petty chiefdoms, a score of two indicates large chiefdoms/petty states, and scores of three or four indicate large states.

(A)　　　　　　　　　　　　　　　　　　　　(B)
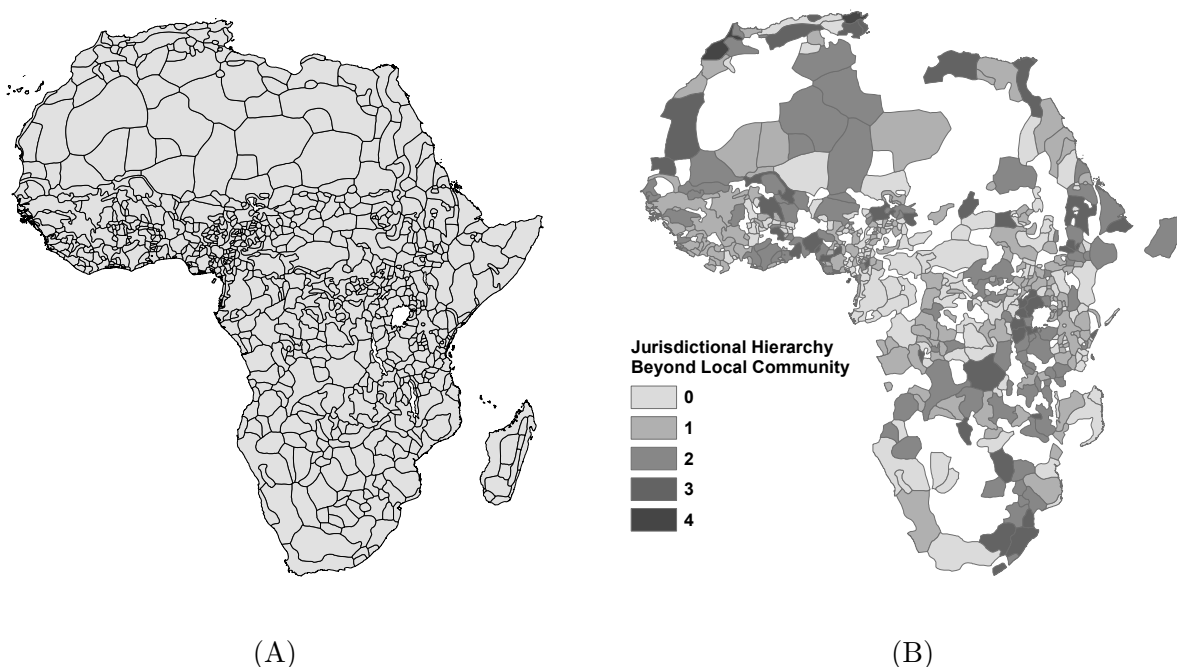
FIGURE II. (A) Historical ethnic homelands from Murdock's (1959) ethnolinguistic map. (B) Pre-colonial centralisation from Murdock's (1967) Ethnographic Atlas.

Figure II(B) provides a visual representation of the ethnic groups' jurisdictional hierarchies according to the classification in the Ethnographic Atlas. While Murdock assembled the Ethnographic Atlas by relying on the previous work of various ethnographers, which should prevent systematic bias arising from his own predispositions, the classification of political jurisdictions possibly suffers from some degree of subjectivity. Following the literature, in particular Gennaioli and Rainer (2006, 2007), I consider an ethnic group to be politically centralised if the jurisdictional hierarchy variable has a score greater than one. Therefore, I

---

[6]The file can be found in the Web Appendix of Fenske (2013), available at jamesfenske.com.

code pre-colonial political centralisation as a binary variable where a value of one indicates centralised ethnic groups and zero indicates non-centralised/stateless ethnic groups.

III.C. *TseTse Suitability Index*

As part of my identification strategy I employ data on the historical disease environment, one of the factors through which the natural environment shaped historical institutions (Engerman and Sokoloff, 2000). In particular, I exploit data on the historical presence of the TseTse fly, a blood sucking insect vector carrying a parasite that can infect humans and animals with Trypanosomiasis (sleeping sickness). Such sickness is fatal if untreated, and affects animals more than humans. In a recent paper, Alsan (2015) shows how the presence of the TseTse fly, by limiting the use of domesticated animals in agriculture, affected African pre-colonial agricultural practices, patterns of subsistence, population density and the probability of ethnic groups being politically centralised.
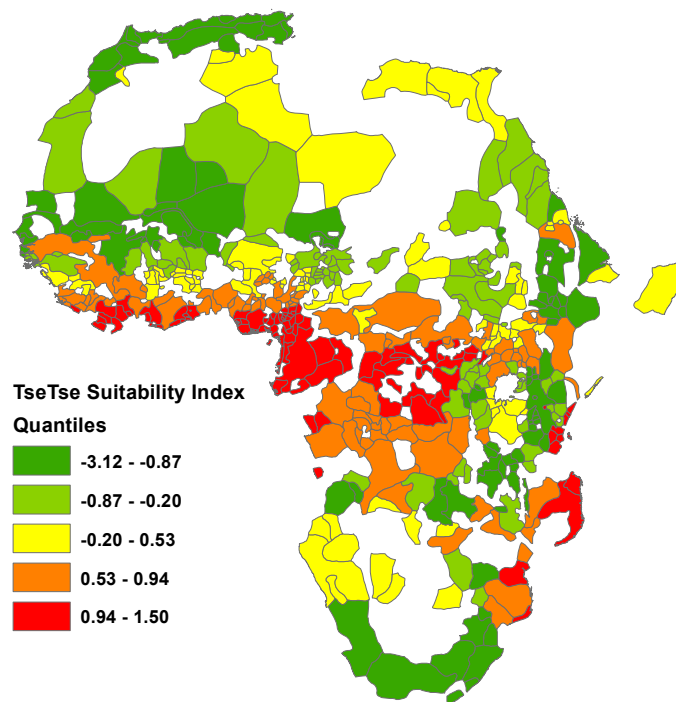

FIGURE III. TseTse Suitability Index from Alsan (2015)

Alsan (2015) employs climate variables on temperature and humidity recreated for the late 19th century to generate a TseTse Suitability Index (TSI) capturing the historical suitability

14

of the environment to the TseTse fly.[7] By combining the TSI data with Murdock's (1959) ethnolinguistic map of Africa, it is possible to construct a suitability index for each ethnic homeland. Figure III provides a visual representation of the TSI across the historical ethnic homelands for which data from the Ethnographic Atlas is also available.

IV. ESTIMATION FRAMEWORK

IV.A. *Main Empirical Model*

The primary interest of this paper is to study how state history shapes present attitudes towards state institutions through the mechanism of intergenerationally transmitted beliefs. To establish the importance of this mechanism, I regress present day individual attitudes on a measure of the state history of the individuals' ethnic groups. The analysis is based on the following main specification:

$$ITI_{i,e,l,c} = \alpha_c + \beta \, PC_e + X'_{i,e,l,c} \, \Gamma + Y'_{l,c} \, \Omega + Z'_e \, \Lambda + \varepsilon_{i,e,l,c} \tag{1}$$

where i indexes individuals, e ethnic groups, l locations (enumeration areas and districts) and c countries. The dependent variable $ITI_{i,e,l,c}$, which varies across individuals, denotes the Institutional Trust Index obtained from the four trust survey measures described in the data section. PC_e measures the level of political centralisation of ethnic group e in the pre-colonial period. The coefficient of interest is β, capturing the estimated relationship between the pre-colonial political centralisation of an individual's ethnic group and the current level of institutional trust of the individual.

Country fixed effects α_c are included to capture nation-specific factors that may affect individuals' institutional trust, such as institutions (Tabellini, 2008b), teaching practices (Algan et al., 2013) and government policies (Aghion et al., 2010).

The vector $X_{i,e,l,c}$ is a set of individual-level controls, including age, age squared, gender, an urban location indicator, education indicators, religion indicators and self-reported living conditions indicators. The education and living conditions indicators are meant as proxies of income, since a direct measure of the respondents' income is not available from the surveys.

The vector $Y_{l,c}$ is a set of enumeration area (village/town/city) and district controls.[8] The first components accounts for the presence of public goods in the individual's enumeration

---

[7]Data and in-depth explanation of how the index was constructed can found in the Web Appendix of Alsan (2015), available at people.stanford.edu/malsan

[8]Enumeration areas and districts are distinct geographical units, with districts encompassing multiple enumeration areas. The two were grouped together in order to simplify the notation.

area, since the lack of public goods is expected to influence the dependent variable. The controls are for the presence of schools, health clinics, electricity, piped water and a sewage system in the individual's enumeration area. The remaining controls include district ethnic fractionalization and the proportion of the sample population in the district which is of the same ethnic group as the individual. The former has been found to affect both income and trust (Easterly and Levine, 1997), while the latter is meant to account for the effect of being part of an ethnic minority.

The vector Z_e is a set of ethnicity-level controls capturing historical characteristics of the individual's ethnic group. These characteristics include the ethnic group's percent dependence on agriculture, practice of intensive agriculture, practice of indigenous slavery, settlement patterns, number of slaves taken from the group during the slave trade (Nunn and Wantchekon, 2011), latitude, longitude, presence of a coast, log land area and agricultural suitability of the group's historical homeland, cultural province and year of observation indicators from the Ethnographic Atlas.[9] I will address the importance of these controls in the next paragraphs.

IV.B. *Identification Strategy*

In order for my estimates to identify the causal effect of the ethnic groups' state history on present day individuals' institutional trust, three main empirical challenges need to be addressed. The first empirical challenge relates to the fact that historical state formation also influenced current economic and institutional development (Michalopoulos and Papaioannou, 2013), which in turn influence institutional trust. This is similar to how, in my theoretical model, goodness of the state across periods is correlated by a factor $\rho > 0$. While I am interested in the direct effect that historical state formation had on institutional trust through intergeneration transmission, I cannot exclude that there is an indirect effect working through the current environment. To deal with this confounding problem, I first control for country fixed effects. By comparing individuals with different ethnic backgrounds within the same country, I eliminate the concern of capturing the effect of better current day institutions. Nonetheless, there remains the issue that, even within the same country, ethnic homelands which historically had more developed institutions correspond to areas that are more developed today. Therefore, as a second step, I exploit the fact that when individuals migrate their beliefs and values migrate with them, but their external environment remains behind (Fernandez, 2007; Fernandez and Fogli, 2009). Murdock's (1959) ethnolinguistic map of Africa allows me to determine which area ethnic groups historically inhabited. Using data on the location of respondents in the Afrobarometer surveys, I can determine whether these

---

[9]Full description of the variables and their sources in Appendix E

individuals are still living in the historical homeland of the ethnic group they are part of.[10] I determine that about 50% of the individuals in the sample are no longer living in their ethnic group's historical homeland. Later in the analysis I am going to use these "migrants" (as I will call them from now on) to better identify intergenerational transmission, since the concern of the indirect effect of historical state formation working through the current external environment will be reduced. This comes from the fact that comparing the results using the full sample with the ones from the sub-sample composed of "migrants" will allow me determine to what extent the estimates change because of the indirect effect. A concern is that individuals in the "migrants" sub-sample are not comparable to the rest, which would make this strategy ineffective. In Appendix Table B.I I compare "migrants" with "non-migrants", i.e. the individuals who are still living in the area that was historically inhabited by their ancestors. The two sub-samples are comparatively similar in terms of both modern individual and historical ethnic group characteristics. What they differ in is that "migrants" are more likely to reside in urban areas, and live in locations that are more ethnically fractionalized and with less co-ethnics present. Given the similarity of individuals from the two parts of the sample, I will later estimate Equation (1) using only the "migrants" sample to determine the importance of portable ethnicity-related cultural traits whose influence expands even outside of the ethnic group's historical homeland (Michalopoulos et al., 2016; Nunn and Wantchekon, 2011).

The second empirical challenge I address relates to the colonial rule. In particular, while I can account for the country-level effect of colonialism through country fixed effects, I am unable to determine how colonisation interacted with the pre-existing ethnicity-level political structures I want to measure. While the Ethnographic Atlas is meant to describe living conditions and characteristics of ethnic groups prior to European contact, a problem would arise if these measures were to capture some form of colonial influence. This could be the case in the event of colonial rule already being in place by the time the ethnographers recorded the groups' features. Regarding this issue, I explore a variable in the Ethnographic Atlas indicating the "year of observation", i.e. when the ethnographic records took place. For the ethnic groups of interest, the earliest time of observation is 1830, with decade changes until the latest in 1960, and with the average being around 1916. The question is whether the groups who were recorded later are more likely to have been influenced by the colonisers, compared to the ones recorded earlier. Regarding the pre-colonial political centralisation measure I consider, the risk is that groups which were recorded later may result as politically developed due to colonial policies. Econometrically, if these colonial policies were to be correlated with some long-lasting effect that colonisation had on commonly-held views about

---

[10]See Appendix Figure B.I.

institutional figures, the issue would amount to my estimations being at risk of suffering from an omitted variable bias. To deal with this potential endogeneity, I am going to include year of observation fixed effects as ethnicity-level controls, which will allow me to conduct a within-period of observation comparison. Appendix C provides more details on this measure and its relationship with the pre-colonial political centralisation measure.

The third empirical challenge is that pre-colonial political centralisation as contained in Murdock's (1967) Ethnographic Atlas is an imprecise measure of the ethnic groups' state history. This is because Murdock's variable captures the presence of a pre-colonially centralised government, but not how long it existed for. More specifically, it may be that a centralised government existed for centuries during the pre-colonial period, or it could be that it emerged only a couple of decades before colonisation, but Murdock's variable would still classify it as the same. Further, while by the time it was measured no centralised government existed, it could be that one such government was in place for a long time in a preceding period. This measurement error may therefore cause an attenuation bias (Greene, 2003) and underestimate the importance of state history in the results.

To deal with the attenuation bias, I will instrument PC_e in Equation (1) with one of the factors that potentially shaped historical state formation in Africa, namely the presence of the TseTse fly. As Alsan (2015) shows, the presence of the TseTse fly affected African pre-colonial agricultural practices, patterns of subsistence, population density and the probability of ethnic groups being politically centralised. By circumscribing the use of domesticated animals as a source of draft power and the adoption of technologies complementary to draft power, the TseTse fly has been hypothesised to have hindered the generation of an agricultural surplus and the transportation of goods overland, negatively affecting the likelihood of groups adopting sedentary lifestyles and more complex organisational structures. I will employ the TseTse Suitability Index for the ethnic homeland, TSI_e, as an instrument for the ethnic group's pre-colonial political centralisation, PC_e. By using one of the factors determining whether the ethnic homeland had favourable conditions for early state formation, I aim to retrieve the true effect of state history. An important point to stress is that the presence of the TseTse fly affected different characteristics of the ethnic groups apart from their pre-colonial political centralisation. As I previously mentioned, Alsan (2015) found the effect of the fly to span over different aspects of the groups' organization. The set of ethnicity-level variables I include in Equation (1) are not only important controls meant to capture historical characteristics of the ethnic groups, but are also used to make sure that the exclusion restriction holds and TSI_e affects the Institutional Trust Index ($ITI_{i,e,l,c}$) only through the variable PC_e. The ethnic groups' percent dependence on agriculture and practice of intensive agriculture are included to capture the importance and type of agricultural practices. As previously mentioned, the

ability to generate an agricultural surplus was one of the aspects that the presence of TseTse fly affected the most and is therefore a necessary control. The practice of indigenous slavery was in some cases due to the lack of animals to be used as a source draft power, and therefore one of the possible consequences of the TseTse fly (Alsan, 2015). Moreover, indigenous slavery was in some cases a feature of autocratic governments (Galor and Klemp, 2017), which may have historically eroded the view of institutional figures, influencing current attitudes. The settlement patters indicators lists categories in order of increasing social and economic development, ranging from fully nomadic to complex settlements, and serves as a proxy of historical population density. The number of slaves taken from the ethnic group during the slave trade is also a necessary control, given that the more advanced and densely populated groups were the most impacted and may therefore show overall lower levels of trust (Nunn and Wantchekon, 2011). As ethnicity-level geographical controls, I include the ethnic homeland's latitude and longitude, and the presence of a coast on its boundaries. The ethnic homeland's log land area is intended as a proxy for the potential size of the ethnic group, which could have affected the likelihood to develop a large government structure. Given that the TseTse fly tended to proliferate in areas suitable for agriculture (Alsan, 2015), I also include an agricultural suitability index. This index refers to land suitability for rain-fed crops, and is used to capture historical agricultural conditions of the ethnic group's homeland. While the vector Z_e includes various ethnicity-level covariates, this list of controls is not necessarily exhaustive. In particular, I cannot exclude the existence of other ethnicity-level covariates that would prevent me from isolating the effect of TSI_e working through PC_e. Including ethnicity fixed effects would allow me to account for these covariates, however the fact that PC_e is measured at the ethnicity level prevents me from doing so. As a flexible alternative to the ethnicity fixed effects, I employ cultural province fixed effects. Cultural provinces are groupings devised by Murdock, capturing spatial and cultural/genealogical correlation of ethnic groups. Appendix Figure C.II shows the cultural provinces the ethnic groups in the Ethnographic Atlas sample belong to, with shading used to represent the different provinces.

Finally, even though I control for a variety of contemporary and historical elements, the presence of hard-to-account for factors is highly likely given the long time span I am looking at. Assuming the validity conditions are satisfied by TSI_e, the instrumental variable approach will also allow me to deal with residual omitted factors correlated with both pre-colonial political centralisation and my outcome variable.

V. Results

V.A. *Main Results: OLS and IV Estimates*

Estimates of Equation (1) are reported in Table I. Panel A reports OLS estimates, while Panel B reports IV estimates. In columns (1)-(2) I employ the full sample, while in columns (3)-(4) I employ the "migrants" sub-sample. Given that the main regressor and many of the explanatory variables vary at the ethnicity level, I adjust standard errors for potential clustering, allowing for non-independence between respondents belonging to the same ethnic group. In Appendix Table D.II I also calculate standard errors adjusted for two-way clustering within ethnic groups and within district, within ethnic groups and within regions, and within ethnic groups and within countries. These methods all produce similar standard errors. For the rest of the paper I am going to report robust standard errors clustered at the ethnic group level.

Looking at the OLS estimates in Panel A, starting from column (1), the coefficient on pre-colonial political centralisation is significant, although weakly, but small. As anticipated in the estimation framework section, I suspect these estimates to suffer from both an attenuation bias and omitted variable bias. To deal with these biases, I employ an identification strategy that makes use of instrumental variables and a battery of modern and, most importantly, historical controls. Panel B of Table I reports IV estimates, where pre-colonial political centralisation of the ethnic group, PC_e, is instrumented with the TseTse Suitability Index of the historical ethnic homeland, TSI_e. For this instrumental variable strategy to work, the TSI of the ethnic homeland needs to affect the current institutional trust of individuals from ethnic group e only through the group's pre-colonial political centralisation. In all specifications I include individual controls from $X_{i,e,l,c}$, location controls from $Y_{l,c}$, country fixed effects and ethnicity-level historical controls from Z_e, while in even-numbered columns I also include year of observation fixed effects. The first stage estimates in Panel B show that a higher TseTse suitability is significantly and negatively correlated with pre-colonial political centralisation as in Alsan (2015). The second stage estimates report a positive and significant coefficient on PC_e, meaning that present individuals from ethnic groups that were pre-colonially centralised show significantly higher levels of institutional trust. To assess the magnitude of the coefficients, consider that the Institutional Trust Index is normalised to be in [0,1], and that PC_e is a binary variable taking on the value of 1 if the ethnic group was politically centralised in the pre-colonial period. Therefore, the reported coefficients on PC_e can be interpreted as percentage changes in the Institutional Trust Index. Taking the second stage estimates in columns (1) as a reference, the coefficient implies that being from an ethnic group that was politically centralised in the pre-colonial period is associated with a

Dependent variable:	Full sample		"Migrants" only	
Institutional Trust Index ($ITI_{i,e,l,c}$)	(1)	(2)	(3)	(4)
	PANEL A: OLS ESTIMATES			
Pre-colonial political centralisation (PC_e)	0.0185* (0.0110)	0.0318*** (0.0104)	0.0198* (0.0112)	0.0280*** (0.0106)
	PANEL B: IV ESTIMATES			
Second stage				
Pre-colonial political centralisation (PC_e)	0.0985*** (0.0268)	0.0608** (0.0252)	0.1321*** (0.0382)	0.0606** (0.0255)
Endogeneity test (p-value)	<0.01	0.19	<0.01	0.14
First stage				
TseTse suitability index (TSI_e)	-0.1964*** (0.0370)	-0.2517*** (0.0497)	-0.1997*** (0.0401)	-0.2748*** (0.0516)
F-stat of excluded instrument	28.24	25.67	24.78	28.41
Observations	29,550	29,550	15,597	15,597
Individual controls	✓	✓	✓	✓
Location controls	✓	✓	✓	✓
Country FE	✓	✓	✓	✓
Ethnicity-level historical controls	✓	✓	✓	✓
Year of observation FE	✗	✓	✗	✓

Notes: Panel A reports OLS estimates. Panel B reports IV estimates. The top of Panel B reports second stage estimates while the bottom reports first stage estimates. The unit of observation is an individual. Columns (1)-(2) employ the full sample, while columns (3)-(4) employ the "migrants" sub-sample. "Migrants" are defined as individuals not living in their historical ethnic homeland. All specifications include country and round fixed effects (constant not reported). Individual controls are for age, age squared, a gender indicator variable, 5 self-reported living conditions indicators, 10 education indicators, 54 religion indicators and an indicator for whether the respondent lives in an urban location. Location controls include dummies for the presence of public goods in the individual's enumeration area, district's ethnic fractionalization and proportion of the district's population that is of the same ethnic group as the individual. Ethnicity-level controls include pre-colonial era percent dependence on agriculture, practice of intensive agriculture, practice of indigenous slavery, settlement patterns, number of slaves taken from the group during the slave trade, latitude, longitude, presence of a coast, log land area, agricultural suitability of the group's historical homeland and cultural province fixed effects. Even-numbered columns include year of observation fixed effects. Below the estimates are reported in parentheses robust standard errors clustered at the ethnic group level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

9.85 percent higher Institutional Trust Index in the present.

Based on the Hausman test, I reject the null hypothesis of consistency of the OLS estimates in column (1) at the 1 percent level, implying that there is considerable omitted variable bias. Moreover, in column (1) the IV coefficient on PC_e is significantly larger when compared with the OLS coefficient. If the instrumental variable strategy is valid, this would indicate that the OLS coefficient is being biased downwards, underestimating the true effect of pre-colonial political centralisation. In column (2) of Table I, I include year of observation fixed effects. As explained in the previous section, these fixed effects indicate the year the ethnographic record of the ethnic group's characteristics took place, and are included to take into account a potential colonial influence on the group's political development by the time of the record. There are three changes to notice once I include these year of observation fixed effects in column (2). First, the OLS coefficient on PC_e increases in size and becomes highly significant. Second, the IV coefficient on PC_e decreases by slightly less than 40 percent, while still retaining its significance. Third, based on the Hausman test, the OLS estimates in column (2) no longer result as suffering from endogeneity. These results thus imply that the problem of omitted variables was biasing the OLS estimates downwards and the IV estimates upwards. While dealing with this bias brings OLS and IV coefficients closer in terms of size and significance, the results imply that an attenuation bias is still present, given that the IV coefficient on PC_e is still two times bigger than the OLS one.

To address the fact that the pre-colonial political centralisation of an ethnic group also influenced its current economic and institutional environment, in columns (3)-(4) I replicate columns (1)-(2) while considering only individuals who no longer reside in the ethnic homeland of their group. This strategy results in a smaller sample of 15,597 observations. Same as when using the full sample, the Hausman test in column (3) indicates PC_e as being endogenous, suggesting that omitted variables are still an issue even when only "migrants" are considered. Columns (3)-(4) also show a strong first stage and, similarly to before, the inclusion of the year of observation fixed effects in column (4) brings IV and OLS coefficients closer in terms of size and significance. Overall, the estimates obtained from the "migrants" sub-sample are remarkably similar to the full sample ones, implying that even for individuals living outside of their ethnic homeland, the influence of their group's state history is still strong. Also, the indirect effect working through current institutional and economic development may not be as problematic as initially considered.

V.B. *Further Results*

The positive correlation between ethnic group's state history and institutional trust of modern individuals I documented in the previous section is consistent with the hypothesised mechanism of intergenerationally transmitted belief. This correlation, however, could also be the result of how I measure my variables, an invalid identification strategy or because of unaccounted for selection and confounding effects of additional unobserved factors. In this section, I undertake a number of strategies to determine whether the results I document are valid and imply a causal relationship. First, I undertake a falsification test where I instrument the ethnic group's pre-colonial political centralisation with the ethnic homeland's Malaria Ecology Index instead of the TseTse Suitability Index. Second, I employ the original classification of institutional development contained in the Ethnographic Atlas as a measure of pre-colonial political centralisation. Third, I use data on the location of African cities in 1800 as an alternative proxy for pre-colonial state formation. Fourth, to account for the potential selection of "migrant" individuals into more developed areas, I introduce a series of current location fixed effects allowing me to conduct a within-location comparison of these individuals. Fifth, to establish the role played by the presence of co-ethnics in preserving ethnicity-related beliefs, I interact the pre-colonial centralisation measure with the share of the district's population that belongs to the same ethnicity as the respondent.

V.B.1 *Malaria Ecology Index Falsification*

My instrumental variable strategy rests on the assumption that the historical presence of the TseTse fly affects modern institutional trust only through pre-colonial political centralisation. To make sure the exclusion restriction holds, I introduced a battery of modern and historical controls. I dedicate particular attention to the colonial rule, which, as pointed out in estimation framework section, is possibly the most important omitted factor. In the event that the colonisers' strategy was also affected by the initial disease environment I consider, then it may possible that the exclusion restriction is not satisfied. Specifically, the TSI instrument may not affect the Institutional Trust Index solely through the ethnic group's pre-colonial political centralisation, but also through the colonisation strategy implemented by Europeans. Given that Malaria was found to be one of the major factors affecting the deadliness of the environment for European settlers (Acemoglu et al., 2001, 2002), I undertake a falsification exercise where I employ as instrument a measure of the prevalence of malaria in the homelands historically inhabited by African ethnic groups. This exercise will allow me to compare the results obtained while using the TSI instrument with those obtained while instrumenting with an aspect of the disease environment which the literature has found to be

23

correlated with the colonisers' strategy. The approximation of malaria disease environment in the historical ethnic homelands is obtained by combining the Malaria Ecology Index by Kiszewski et al. (2004) with Murdock's (1959) ethnolinguistic map. Appendix Table D.III reports IV estimates mirroring those of Table I, with the Malaria Ecology Index being used as instrument. First state estimates in columns (1) and (3) report a significantly negative effect of the Malaria Ecology Index on pre-colonial political centralisation, although coefficients are about ten times smaller than those on TSI_e. On the other hand, once the year of observation fixed effects are included in columns (2) and (4), the first stage coefficient on the Malaria Ecology Index becomes small and statistically insignificant. Across all specifications, the F-statistic of excluded instruments is low and the second stage coefficient on PC_e not significant. This falsification test does not provide direct support for the validity of the TSI instrument. Nonetheless, the stark difference in results obtained while instrumenting with the Malaria Ecology Index shows how these two aspects of the disease environment are fundamentally different in the context of this analysis. The concern of TseTse presence capturing omitted colonial rule factors, while not solved, is mitigated to an extent.

V.B.2 *Different Classification of Political Centralisation*

Up to this point the analysis employed a binary measure of pre-political political centralisation, based on Murdock's (1967) "Jurisdictional Hierarchy Beyond the Local Community Level" index. This variable, ranging from 0 to 4, indicates the number of jurisdictional levels above the local level, where 0 stands for stateless societies, 1 and 2 stand for petty and large chiefdoms, while 3 and 4 stand for states and large states. Following Gennaioli and Rainer (2006, 2007), I aggregated the index into a binary measure to account for the measurement error potentially deriving from subjectivity in the classification of political jurisdictions by ethnographers. To check whether the results are sensitive to this aggregation, I replicate the estimates of Table I using the original "Jurisdictional Hierarchy Beyond the Local Community Level" index. Appendix Table D.IV reports OLS and IV estimates. Overall, results using the original index are similar in terms of both size and significance to those obtained with the binary measure. Worth noticing, however, is that compared to those reported in Table I, these estimates perform more poorly in terms of excluded instrument and result as endogenous even when the full set of controls is included.

V.B.3 *Alternative Proxy for Pre-Colonial State Formation*

To assess the robustness of results to alternative measures of pre-colonial state formation, I employ a proxy based on the data from Chandler (1987) on the location of African cities

in 1800. Under the assumption that states tended to have a large city as political center, I construct an indicator variable equalling one if there was a city in 1800 with more than 20,000 inhabitants on the area of the ethnic homeland. OLS and IV estimates employing this alternative measure are reported in Appendix Table D.V. OLS estimates produce results similar to those obtained with the pre-colonial political centralisation measure, both in terms of size and significance. Regarding the IV estimates, the first-stage coefficient on TSI_e is highly significant, although the size is two to three times smaller compared to Table I. Second-stage estimates report a highly significant and positive coefficient on the city measure, showing an effect on the Institutional Trust Index which is even stronger than the one obtained with PC_e. About this last point, the number of ethnic groups with a city on its homeland in 1800 is limited compared to the groups that result as politically centralised in the pre-colonial period, therefore it is possible that the estimates are being biased upwards.

V.B.4 *Accounting for Selection*

As part of the identification strategy, I focused on the "migrants" sub-sample as a way to deal with the indirect effect of state history working through the current environment. A potential issue with this strategy is that there may be a selection of where these individuals chose to live. For example, it could be that these individuals, or a previous generation, migrated to places that are more developed or with better institutions. Alternatively, these individuals could have migrated to a place that is similar to their homeland of origin or inhabited by co-ethnics. As shown in Appendix Table B.I, "migrants" are more likely to reside in urban areas, consistently with a general migration pattern of individuals moving to urban centers in search of better opportunities. The reported comparative statistics also show that, on average, less co-ethnics are present where "migrants" live. On a different note, it may be that these individuals assimilated the beliefs and values of where they migrated. While intergenerationally transmitted beliefs are the focus of this work, current environment and norms of where individuals live also have an influence on their beliefs and values.

To deal with these concerns, I introduce a series of current homeland fixed effects corresponding to the ethnic homeland where individuals currently results to be located, based on Murdock's (1959) ethnolinguistic map and the location of respondents in the Afrobarometer surveys. Adding these fixed effects in a regression considering only "migrants" will absorb time-invariant characteristics related to the geographic, economic and institutional environment of where they currently reside, allowing for a within-location comparison of these individuals. Furthermore, corresponding to the location where ethnic groups historically resided, these fixed effects will also allow me to take into account cultural and historical

features of the location.[11] The current ethnic homeland strategy introduces a considerable number of fixed effects, substantially reducing the amount of within-group variation to exploit in the estimation. Due to the cultural province and year of observation fixed effects being at the ethnicity level like the current homeland fixed effects, including all three together would leave very little within-group variation, lessening the meaningfulness of the results. Therefore, for the sake of this exercise, I will not include the cultural province and year of observation fixed effects I previously employed. Appendix Table D.VI reports IV estimates of baseline Equation (1) and alternative specifications introduced until now, with current homeland fixed effects. Overall, the estimates are very similar to the ones I obtained before and confirm the previous findings.

V.B.5 *Proportion of Co-Ethnics*

So far, the analysis has found that descending from an ethnic group with a history of political centralisation is, on average, positively correlated with current individuals' institutional trust. I will now consider a source of heterogeneity that may affect the resilience of ethnicity-related beliefs and other ethnic traits across generations, that is the presence of co-ethnics. In particular, I ask whether interacting with people of the same ethnic group, or more generally living in areas predominantly inhabited by co-ethnics, is determinant in preserving these intergenerationally transmitted beliefs. For this purpose, I consider a variable previously included as a control, the proportion of co-ethnics in the district. This measure, whose construction is based on the sample population of the Afrobarometer surveys, indicates the share of the district's population that is of the same ethnicity as the respondent. Studying the distribution of this variable reveals that about 50% of the individuals in my sample live in a district with more than 75% of its population being composed of co-ethnics. This finding is consistent with what I previously determined regarding the number of individuals still living in the historical homeland of their ethnic group. To study the heterogeneous effect of co-ethnics presence, I create four dummy variables indicating, respectively, a proportion of co-ethnics in the district between 0% and 25%, 25% and 50%, 50% and 75%, or 75% and 100%. Interacting these variables with the pre-colonial political centralisation measure will allow me to determine the differential effect of the varying presence of co-ethnics. Appendix Table D.VII reports OLS and IV estimates employing these interaction terms. Across all specification I include the full set of controls introduced in Table I. For both OLS and IV estimates I report four coefficients: the one on the (centered) proportion of co-ethnics in the district, the one on the ethnic group's pre-colonial political centralisation (PC_e), the one on

---

[11]Running such a regression with the full sample would absorb all the variation of ethnicity-related variables for individuals still living in their ethnic homeland.

the interaction with the quartile proportion dummy (*Proportion X PC_e*), and the one on the linear combination of the previous two (*PC_e + Proportion X PC_e*). OLS estimates show that for individuals living in a district with less than 25% of its population being composed of co-ethnics, the interaction term shows a significantly negative effect. Moreover, the linear combination results as not being statistically significant, indicating that, for individuals whose ethnic group is scarcely present in the area, the effect of PC_e is not significantly different from zero. On the other hand, for individuals with 25-50% and 50-75% shares, the coefficient on the interaction term is trivial, and the linear combinations report coefficients whose size and significance is equivalent to those of PC_e. In other words, for individuals whose ethnic group is averagely present in the district, there is no differential effect. Finally, for individuals living in districts with more than 75% of its population being made of co-ethics, the interaction term reports a significant and positive effect. Additionally, the linear combination shows a significantly positive coefficient, with the effect of PC_e for these individuals being higher than the one obtained on average in the OLS estimates of Table I. Living in districts homogeneously composed of co-ethnics is associated with a stronger impact of the group's history. Turning to the IV estimates, adding an interaction with the instrumented pre-colonial centralisation variable requires me to obtain an additional instrument. I proceed with a commonly chosen approach, adding the interaction of TSI_e with the proportion dummy (*Proportion X TSI_e*) to the list of instruments. While IV results match the OLS ones, although with overall reduced significance, the two employed instruments do not emerge as jointly excluded, which prevents me from attaining proper identification for these estimates. Nonetheless, Table I shows how, when the full set of controls is included, OLS estimates do not result as suffering from endogeneity. Therefore, whereas the IV results in Appendix Table D.VII cannot be considered as properly identified, it is still possible to obtain informative evidence from the OLS estimates, although biased towards zero because of the attenuation bias. Overall, the heterogeneity analysis provides evidence that the presence of co-ethnics matters for the resilience of these ethnicity-related beliefs, with limited presence leading to a statistically insignificant effect, whereas ethnically homogeneous communities lead to their strengthening.

VI. Conclusions

At the start of European colonisation, Africa's numerous ethnic groups were characterised by a remarkable heterogeneity in terms of subsistence economy and state development. Generations later, descendants of those ethnic groups live together in the same countries and settings. Various studies have documented how historical factors can shape cultural traits, and how these cultural differences can persist for long periods of time. Is the perception of

27

state institutions one of those persistent traits? Does one's ancestry in terms of pre-colonial state development matter for present attitudes?

In this paper, I address these questions both theoretically and empirically. I develop a model showing how, by heuristically relying on tradition, beliefs can reflect experiences of past generations. It highlights a mechanism through which individuals from groups with a history of political centralisation will be more likely to form a positive belief of the state. I then test my hypothesis empirically, by combining contemporary individual-level survey data on trust in institutional figures with historical data on pre-colonial political centralisation by ethnic group from the Ethnographic Atlas. To identify a causal effect, I employ two main identification strategies. First, I use data on the historical disease environment, namely TseTse fly presence, as an instrument for the level of state development. Second, to isolate the channel of intergenerational transmission and determine the importance of portable ethnicity-related cultural traits, I consider survey individuals living outside of their ethnic group's historical homeland. Overall, results show that individuals descending from ethnic groups that were pre-colonially centralised show significantly higher levels of institutional trust. I then perform a number of tests to assess the validity of my identification strategy and results. Specifically, I perform a falsification test by instrumenting with a different measure of the disease environment; test robustness by employing a different classification of political centralisation and an alternative measure of pre-colonial state formation based on data on the location of African cities in 1800; and account for selection of individuals into more developed areas. These tests confirm previous the results. I then study the role played by the presence of co-ethnics, finding a heterogeneous effect on the resilience of ethnicity-related beliefs.

This paper adds to a vibrant new literature in economics, studying how persistent cultural traits can act as a mechanism through which historical factors shape current outcomes. Results call for future research. Particularly, it highlights a promising research agenda on the intergenerational transmission of values and beliefs. Moreover, a next natural step would be to study how these ethnic legacies interact with national policies.

Bibliography

Acemoglu, Daron and Johnson, Simon and Robinson, James A (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

Acemoglu, Daron and Johnson, Simon and Robinson, James A (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly journal of economics*, 117(4):1231–1294.

Aghion, Philippe and Algan, Yann and Cahuc, Pierre and Shleifer, Andrei (2010). Regulation and distrust. *The Quarterly Journal of Economics*, 125(3):1015–1049.

Alesina, Alberto and Giuliano, Paola (2015). Culture and institutions. *Journal of Economic Literature*, 53(4):898–944.

Alesina, Alberto and Giuliano, Paola and Nunn, Nathan (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2):469–530.

Algan, Yann and Cahuc, Pierre (2010). Inherited trust and growth. *The American Economic Review*, 100(5):2060–2092.

Algan, Yann and Cahuc, Pierre and Shleifer, Andrei (2013). Teaching practices and social capital. *American Economic Journal: Applied Economics*, 5(3):189–210.

Almond, Gabriel and Verba, Sidney (1963). The civic culture: political attitudes and democracy in five countries. *Princeton University Press*.

Alsan, Marcella (2015). The effect of the tsetse fly on african development. *American Economic Review*, 105(1):382–410.

Becker, Sascha O and Boeckh, Katrin and Hainz, Christa and Woessmann, Ludger (2016). The empire is dead, long live the empire! long-run persistence of trust and corruption in the bureaucracy. *The Economic Journal*, 126(590):40–74.

Bisin, Alberto and Verdier, Thierry (2017). On the joint evolution of culture and institutions. *NBER Working Paper*.

Bockstette, Valerie and Chanda, Areendam and Putterman, Louis (2002). States and markets: The advantage of an early start. *Journal of Economic Growth*, 7(4):347–369.

Boyd, Robert and Richerson, Peter J (1985). *Culture and the evolutionary process.* Chicago: University of Chicago Press.

Boyd, Robert and Richerson, Peter J (1995). Why does culture increase human adaptability? *Ethology and sociobiology*, 16(2):125–143.

Boyd, Robert and Richerson, Peter J (2005). *The origin and evolution of cultures.*

New York: Oxford University Press.

CHANDLER, TERTIUS (1987). *Four thousand years of urban growth: an historical census.* Lewiston New York St. Davids University Press.

DECONINCK, KOEN AND VERPOORTEN, MARIJKE (2013). Narrow and scientific replication of 'the slave trade and the origins of mistrust in africa'. *Journal of Applied Econometrics*, 28(1):166–169.

DELL, MELISSA (2010). The persistent effects of peru's mining mita. *Econometrica*, 78(6):1863–1903.

DIAMOND, JARED (1997). *Germs, guns, and steel: the fates of human societies.* New York: Norton.

EASTERLY, WILLIAM AND LEVINE, ROSS (1997). Africa's growth tragedy: policies and ethnic divisions. *The Quarterly Journal of Economics*, 112(4):1203–1250.

EASTON, DAVID (1965). *A systems analysis of political life.* New York: Wiley.

ENGERMAN, STANLEY L AND SOKOLOFF, KENNETH L (2000). History lessons: Institutions, factors endowments, and paths of development in the new world. *The Journal of Economic Perspectives*, 14(3):217–232.

FENSKE, JAMES (2013). Does land abundance explain african institutions? *The Economic Journal*, 123(573):1363–1390.

FERNANDEZ, RAQUEL (2007). Women, work, and culture. *Journal of the European Economic Association*, 5(2-3):305–332.

FERNANDEZ, RAQUEL AND FOGLI, ALESSANDRA (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–177.

FUKUYAMA, FRANCIS (1995). *Trust: The social virtues and the creation of prosperity.* Free Press.

GALOR, ODED AND KLEMP, MARC (2017). Roots of autocracy. *NBER Working Paper.*

GENNAIOLI, NICOLA AND RAINER, ILIA (2006). Precolonial centralization and institutional quality in africa. *Institutions and norms in economic development*, pages 21–46.

GENNAIOLI, NICOLA AND RAINER, ILIA (2007). The modern impact of precolonial centralization in africa. *Journal of Economic Growth*, 12(3):185–234.

GERSHMAN, BORIS (2017). Long-run development and the new cultural economics. *Demographic Change and Long-Run Development*, page 221.

GIAVAZZI, FRANCESCO AND PETKOV, IVAN AND SCHIANTARELLI, FABIO (2014). Culture: Persistence and evolution. *NBER Working Paper.*

GIULIANO, PAOLA (2007). Living arrangements in western europe: Does cultural origin

matter? *Journal of the European Economic Association*, 5(5):927–952.

GIULIANO, PAOLA AND NUNN, NATHAN (2013). The transmission of democracy: From the village to the nation-state. *American Economic Review*, 103(3):86–92.

GIULIANO, PAOLA AND NUNN, NATHAN (2017). Understanding cultural persistence and change. *NBER Working Paper*.

GREENE, WILLIAM H (2003). *Econometric analysis*. Pearson.

GREIF, AVNER (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of political economy*, 102(5):912–950.

GUISO, LUIGI AND SAPIENZA, PAOLA AND ZINGALES, LUIGI (2016). Long-term persistence. *Journal of the European Economic Association*, 14(6):1401–1436.

HENRICH, JOSEPH AND BOYD, ROBERT AND BOWLES, SAMUEL AND CAMERER, COLIN AND FEHR, ERNST AND GINTIS, HERBERT AND MCELREATH, RICHARD AND ALVARD, MICHAEL AND BARR, ABIGAIL AND ENSMINGER, JEAN AND OTHERS (2005). "economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and brain sciences*, 28(6):795–815.

HENRICH, JOSEPH PATRICK (2004). *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies*. Oxford University Press.

KISZEWSKI, ANTHONY AND MELLINGER, ANDREW AND SPIELMAN, ANDREW AND MALANEY, PIA AND SACHS, SONIA EHRLICH AND SACHS, JEFFREY (2004). A global index representing the stability of malaria transmission. *The American journal of tropical medicine and hygiene*, 70(5):486–498.

LIPSET, SEYMOUR MARTIN (1959). Some social requisites of democracy: Economic development and political legitimacy. *American political science review*, 53(1):69–105.

MICHALOPOULOS, STELIOS AND PAPAIOANNOU, ELIAS (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica*, 81(1):113–152.

MICHALOPOULOS, STELIOS AND PAPAIOANNOU, ELIAS (2014). National institutions and subnational development in africa. *The Quarterly journal of economics*, 129(1):151–213.

MICHALOPOULOS, STELIOS AND PUTTERMAN, LOUIS AND WEIL, DAVID N (2016). The influence of ancestral lifeways on individual economic outcomes in sub-saharan africa. *NBER Working Paper*.

MURDOCK, GEORGE P (1959). *Africa: its peoples and their culture history*. New York: McGraw-Hill.

MURDOCK, GEORGE PETER (1967). *Ethnographic atlas*. Pittsburgh: University of Pittsburgh

Press.

NUNN, NATHAN (2008). The long-term effects of africa's slave trades. *The Quarterly Journal of Economics*, 123(1):139–176.

NUNN, NATHAN (2012). Culture and the historical process. *Economic History of Developing Regions*, 27(sup1):S108–S126.

NUNN, NATHAN AND WANTCHEKON, LEONARD (2011). The slave trade and the origins of mistrust in africa. *American Economic Review*, 101:3221–3252.

PUTNAM, ROBERT D (1993). *Making democracy work: Civic traditions in modern Italy.* Princeton University Press.

PUTNAM, ROBERT D (2000). *Bowling alone: The collapse and revival of American community.* New York: Simon and Schuster.

TABELLINI, GUIDO (2008a). Institutions and culture. *Journal of the European Economic Association*, 6(2-3):255–294.

TABELLINI, GUIDO (2008b). The scope of cooperation: Values and incentives. *The Quarterly journal of economics*, 123(3):905–950.

TABELLINI, GUIDO (2010). Culture and institutions: economic development in the regions of europe. *Journal of the European Economic Association*, 8(4):677–716.

VOIGTLÄNDER, NICO AND VOTH, HANS-JOACHIM (2012). Persecution perpetuated: the medieval origins of anti-semitic violence in nazi germany. *The Quarterly Journal of Economics*, 127(3):1339–1392.

Appendix

A. Institutional Trust Index

This section describes the construction of the Institutional Trust Index in detail. The index is constructed by aggregating via PCA the answers to four questions from the Afrobarometer surveys, being the reported trust with respect to: the president, the parliament, the police and courts of law. For each of these variables, the answer categories were (i) "Not at all", (ii) "Just a little", (iii) "Somewhat", and (iv) "A lot". An overview of the responses to these trust questions is provided in Table A.I.

TABLE A.I. Overview of responses to the trust questions

Response	How much do you trust each of the following:			
	President	Parliament	Police	Courts of law
Not at all	15.97%	17.77%	22.62%	13.74%
Just a little	20.36%	25.06%	23.56%	23.67%
Somewhat	20.54%	26.37%	23.66%	27.7%
A lot	43.13%	30.81%	30.17%	34.89%
Total	35,932	35,007	36,232	35,386

Notes: The table reports summary statistics for four measures of trust from the third and fourth rounds of Afrobarometer surveys. The unit of observation is an individual.

Given that the answers are categorical, I convert them into a variable assigning a number to each category. The resulting variables take on values going from 0 to 3, where 0 corresponds to "Not at all", 1 to "Just a little", 2 to "Somewhat" and 3 to "A lot".

The Institutional Trust Index is the first principal component of these four trust questions, normalised to be in $[0, 1]$. The first principal component explains 0.62 of total variance and has an eigenvalue of 2.49, while the second component has an eigenvalue of 0.75. The Institutional Trust Index loads positively on trust of the president (0.5006), positively on trust of the parliament (0.5107), positively on trust of the police (0.4956) and positively on

trust of courts of law (0.4929). Figure A.II illustrates the distribution of the index across the countries in the sample.

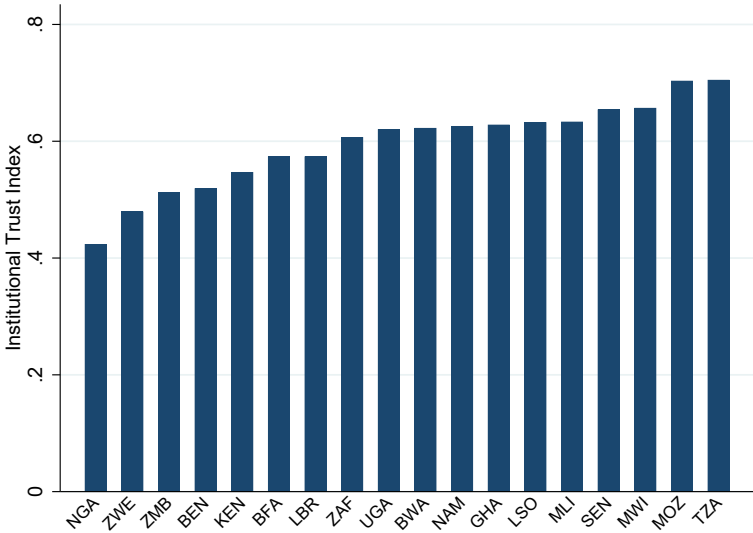FIGURE A.I. Countries included in the fourth round of the Afrobarometer surveys

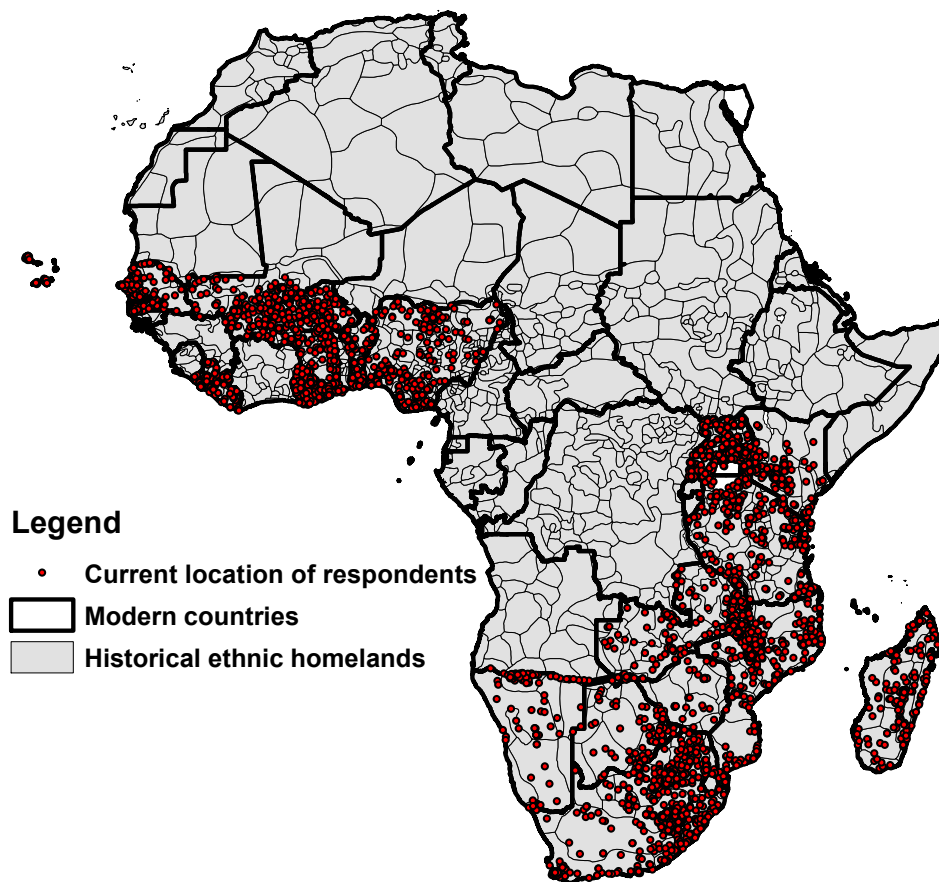FIGURE A.II. Institutional Trust Index by country

FIGURE B.I. Historical ethnic homelands from Murdock's (1959) ethnolinguistic map and current location of respondents in the Afrobarometer surveys.

TABLE B.I. COMPARATIVE STATISTICS

| | "Migrants" | | "Non-Migrants" | |
	Mean	(St.Dev.)	Mean	(St.Dev.)
Afrobarometer surveys				
Institutional Trust Index	0.58	(0.29)	0.60	(0.29)
Age	36	(14.23)	37	(14.84)
Male	0.50	(0.50)	0.50	(0.50)
Urban area	0.42	(0.49)	0.31	(0.46)
Education	3.26	(2.04)	2.96	(2.00)
Living conditions	2.59	(1.22)	2.56	(1.21)
District-level ethnic fractionalization	0.45	(0.29)	0.31	(0.29)
Proportion of ethnic group in district	0.52	(0.35)	0.74	(0.30)
Historical variables				
Pre-colonial political centralisation	0.41	(0.49)	0.45	(0.50)
Agricultural dependence	5.78	(1.49)	5.74	(1.45)
Intensive agriculture	0.28	(0.45)	0.27	(0.45)
Indigenous slavery	0.85	(0.36)	0.83	(0.38)
Slave exports (ln(1+exports/area))	0.31	(0.71)	0.31	(0.67)
Settlement patterns	5.79	(1.78)	5.83	(1.72)
Observations	20,040		17,247	

Notes: The table compares summary statistics between "migrants" and "non-migrants". The former is defined as the sub-sample of individuals who are not living in their ethnic group's historical homeland. The latter is defined as the sub-sample of individuals who are living in the area that was historically inhabited by their ancestors. For explanation and details of variables see Appendix E.

C. Further Ethnographic Atlas Controls

Year of Observation

The year of observation variable from the Ethnographic Atlas indicates the year when the ethnographic record took place for each ethnic group. Figure C.I provides an illustration of the distribution of the ethnographic records over the years.
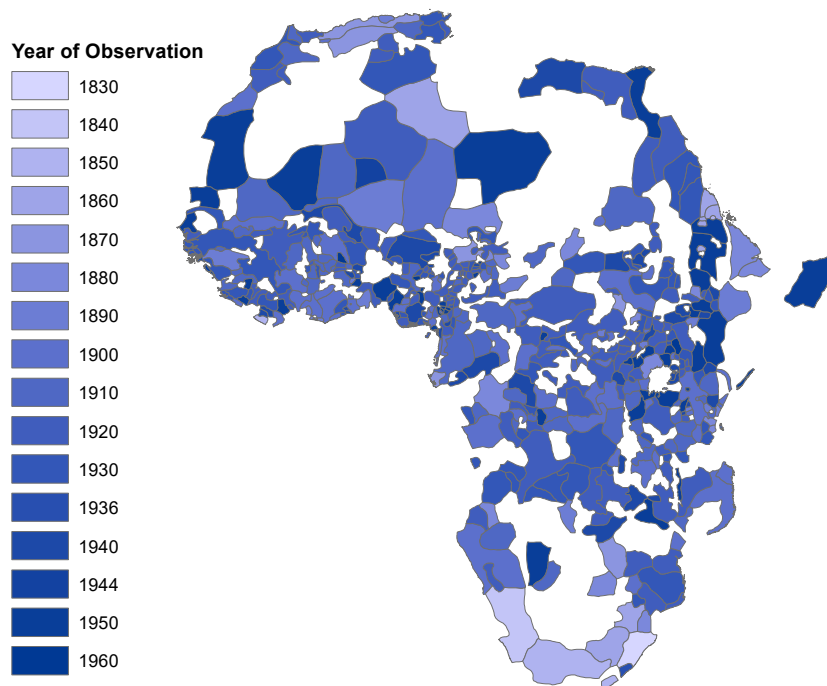
FIGURE C.I. Year of observation from Murdock's (1967) Ethnographic Atlas.

A concern is whether the groups that were recorded later resulted as being politically developed because they were more likely to have been influenced by the colonisers by the time the record took place. Regarding this concern, it would be useful to have information on how much time passed between the time an ethnic group first came in contact with Europeans and the time the ethnographic record took place. While this information is not available, looking at the correlation between the year of observation and the measure of pre-colonial political centralisation reveals a statistically significant and negative relationship ($\rho = -0.2390$). While this negative correlation does not confirm a lack of colonial influence, it implies that the more institutionally developed groups tended to be recorded earlier. Moreover, the year of observation is significantly and negatively correlated with the presence of a coast on the

ethnic homeland's boundaries ($\rho = -0.1703$), coherently with the historical narrative. On the other hand, pre-colonial political centralisation is not significantly correlated with the presence of a coast($\rho = 0.0479$). Overall, these simple correlations attenuate, to a certain extent, the concern of the pre-colonial political centralisation measure capturing the effect of colonial policies.

Cultural provinces

Cultural provinces are groupings devised my Murdock, capturing spatial and cultural/genealogical correlation of ethnic groups.
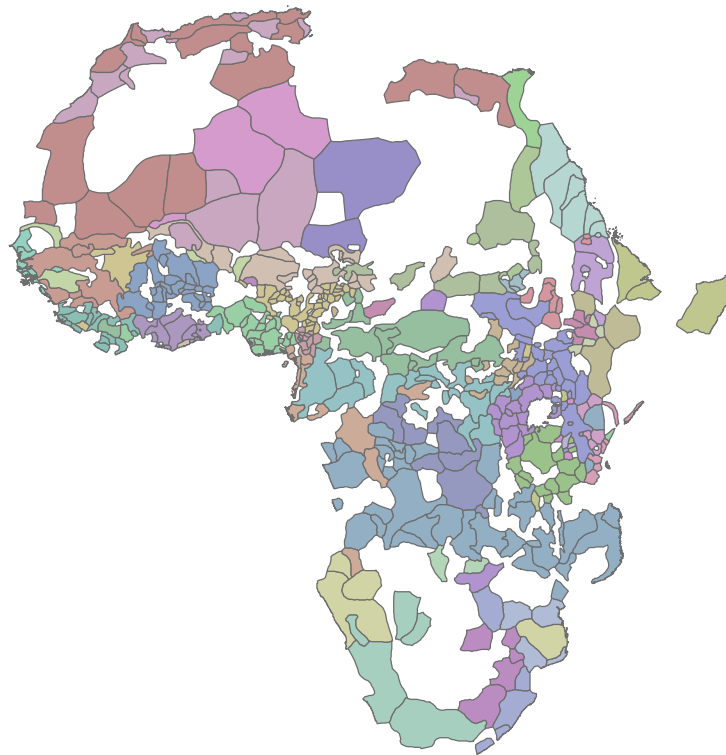

FIGURE C.II. Cultural provinces in Murdock's (1967) Ethnographic Atlas, with shading used to represent the 44 provinces.

D. Additional Tables

Table D.I. Summary Statistics

	Mean	St.Dev.	Min	Max
Afrobarometer surveys				
Institutional Trust Index	0.59	0.29	0	1
Age	36	14.52	18	99
Male	0.50	0.50	0	1
Urban area	0.37	0.48	0	1
Education	3.12	2.03	0	9
Living conditions	2.57	1.21	1	5
District-level ethnic fractionalization	0.38	0.30	0	0.91
Proportion of ethnic group in district	0.62	0.35	0	1
School	0.83	0.38	0	1
Electricity	0.53	0.50	0	1
Piped water	0.48	0.50	0	1
Sewage	0.22	0.41	0	1
Health clinic	0.55	0.50	0	1
Historical variables				
Pre-colonial political centralisation	0.40	0.49	0	1
Jurisdictional hierarchy beyond local community	1.35	0.92	0	3
TseTse suitability index	-0.01	0.99	-3.12	1.45
Agricultural dependence	5.78	1.48	0	9
Intensive agriculture	0.28	0.45	0	1
Indigenous slavery	0.84	0.37	0	1
Slave exports (ln(1+exports/area))	0.31	0.70	0	3.66
Settlement patterns	5.97	1.47	1	8
Latitude	-0.02	11.61	-33	28
Longitude	17.35	17.84	-17	48
Log land area	23.72	1.13	21.09	26.48
Coast	0.18	0.39	0	1
Agricultural suitability index	0.55	0.17	0.09	0.82
Year of observation	1916	23	1830	1960

Notes: For explanation and details of variables see Appendix E.

Dependent variable:	Full sample		"Migrants" only	
Institutional Trust Index ($ITI_{i,e,l,c}$)	(1)	(2)	(3)	(4)

	PANEL A: OLS ESTIMATES			
Pre-colonial pol. central. (PC_e)	0.0185	0.0318	0.0198	0.0280
Ethnicity-District double-cluster s.e.	(0.0110)*	(0.0101)***	(0.0105)*	(0.0101)***
Ethnicity-Region double-cluster s.e.	[0.0116]	[0.0107]***	[0.0116]*	[0.0114]**
Ethnicity-Country double-cluster s.e.	{0.0113}	{0.0096}***	{0.0127}	{0.0106}***

	PANEL B: IV ESTIMATES			
Second stage				
Pre-colonial pol. central. (PC_e)	0.0985	0.0608	0.1321	0.0606
Ethnicity-District double-cluster s.e.	(0.0281)***	(0.0252)**	(0.0375)***	(0.0252)**
Ethnicity-Region double-cluster s.e.	[0.0200]***	[0.0158]***	[0.0393]***	[0.0221]***
Ethnicity-Country double-cluster s.e.	{0.0275}***	{0.0297}**	{0.0487}***	{0.0306}**
First stage				
TseTse suitability index (TSI_e)	-0.1964	-0.2517	-0.1997	-0.2748
Ethnicity-District double-cluster s.e.	(0.0384)***	(0.0498)***	(0.0417)***	(0.0519)***
Ethnicity-Region double-cluster s.e.	[0.0377]***	[0.0504]***	[0.0422]***	[0.0534]***
Ethnicity-Country double-cluster s.e.	{0.0408}***	{0.0520}***	{0.0516}***	{0.0586}***
Observations	29,550	29,550	15,597	15,597
Individual controls	✓	✓	✓	✓
Location controls	✓	✓	✓	✓
Country FE	✓	✓	✓	✓
Ethnicity-level historical controls	✓	✓	✓	✓
Year of observation FE	✗	✓	✗	✓

Notes: Panel A reports OLS estimates. Panel B reports IV estimates. The top of Panel B reports second stage estimates while the bottom reports first stage estimates. The unit of observation is an individual. Columns (1)-(2) employ the full sample, while columns (3)-(4) employ the "migrants" sub-sample. "Migrants" are defined as individuals not living in their historical ethnic homeland. Individuals controls, location controls and ethnicity-level controls are the same as in Table I. All specifications include country and round fixed effects (constant not reported). Even-numbered columns include year of observation fixed effects. Below each coefficient three standard errors are reported. In parentheses are robust standard errors adjusted for two-way clustering within ethnic groups and within districts. In square brackets are robust standard errors adjusted for two-way clustering within ethnic groups and within regions. In curly brackets are robust standard errors adjusted for two-way clustering within ethnic groups and within countries. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

TABLE D.III. IV ESTIMATES OF THE EFFECT OF STATE HISTORY ON INSTITUTIONAL TRUST, MALARIA ECOLOGY INDEX FALSIFICATION TEST

Dependent variable:	Full sample		"Migrants" only	
Institutional Trust Index ($ITI_{i,e,l,c}$)	(1)	(2)	(3)	(4)
Second stage				
Pre-colonial political	-0.0018	0.0363	-0.0133	-0.6086
centralisation (PC_e)	(0.0588)	(0.7497)	(0.0722)	(1.5186)
Endogeneity test (p-value)	0.72	0.99	0.61	0.09
First stage				
Malaria ecology index	-0.0219**	-0.0020	-0.0216**	-0.0040
	(0.0097)	(0.0089)	(0.0098)	(0.0095)
F-stat of excluded instrument	5.07	0.05	4.91	0.18
Observations	29,561	29,561	15,610	15,610
Individual controls	✓	✓	✓	✓
Location controls	✓	✓	✓	✓
Country FE	✓	✓	✓	✓
Ethnicity-level historical controls	✓	✓	✓	✓
Year of observation FE	✗	✓	✗	✓

Notes: The table reports IV estimates. The top panel reports second stage estimates while the bottom panel reports first stage estimates. The unit of observation is an individual. Columns (1)-(2) employ the full sample, while columns (3)-(4) employ the "migrants" sub-sample. "Migrants" are defined as individuals not living in their historical ethnic homeland. Individuals controls, location controls and ethnicity-level controls are the same as in Table I. All specifications include country and round fixed effects (constant not reported). Even-numbered columns include year of observation fixed effects. Below the estimates are reported in parentheses robust standard errors clustered at the ethnic group level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

TABLE D.IV. OLS AND IV ESTIMATES OF THE EFFECT OF STATE HISTORY ON
INSTITUTIONAL TRUST, ROBUSTNESS CHECK WITH ALTERNATIVE CLASSIFICATION OF
PRE-COLONIAL POLITICAL CENTRALISATION

Dependent variable:	Full sample		"Migrants" only	
Institutional Trust Index ($ITI_{i,e,l,c}$)	(1)	(2)	(3)	(4)
	PANEL A: OLS ESTIMATES			
Jurisdictional hierarchies beyond local community	0.0065 (0.0052)	0.0114** (0.0046)	0.0095 (0.0059)	0.0122** (0.0049)
	PANEL B: IV ESTIMATES			
Second stage				
Jurisdictional hierarchies beyond local community	0.0599*** (0.0190)	0.0473** (0.0220)	0.0892*** (0.0290)	0.0494** (0.0224)
Endogeneity test (p-value)	<0.01	0.05	<0.01	0.05
First stage				
TseTse suitability index (TSI_e)	-0.3233*** (0.0849)	-0.3252*** (0.1113)	-0.2962*** (0.0840)	-0.3400*** (0.1098)
F-stat of excluded instrument	14.51	8.54	12.45	9.58
Observations	29,566	29,566	15,615	15,615
Individual controls	✓	✓	✓	✓
Location controls	✓	✓	✓	✓
Country FE	✓	✓	✓	✓
Ethnicity-level historical controls	✓	✓	✓	✓
Year of observation FE	✗	✓	✗	✓

Notes: Panel A reports OLS estimates. Panel B reports IV estimates. The top of Panel B reports second stage estimates while the bottom reports first stage estimates. The unit of observation is an individual. Columns (1)-(2) employ the full sample, while columns (3)-(4) employ the "migrants" sub-sample. "Migrants" are defined as individuals not living in their historical ethnic homeland. Individuals controls, location controls and ethnicity-level controls are the same as in Table I. All specifications include country and round fixed effects (constant not reported). Even-numbered columns include year of observation fixed effects. Below the estimates are reported in parentheses robust standard errors clustered at the ethnic group level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Dependent variable:	Full sample		"Migrants" only	
Institutional Trust Index ($ITI_{i,e,l,c}$)	(1)	(2)	(3)	(4)
	PANEL A: OLS ESTIMATES			
City in 1800	0.0359***	0.0296**	0.0373**	0.0372**
	(0.0136)	(0.0141)	(0.0153)	(0.0159)
	PANEL B: IV ESTIMATES			
Second stage				
City in 1800	0.2213***	0.0904**	0.3141***	0.1026**
	(0.0732)	(0.0378)	(0.1140)	(0.0406)
Endogeneity test (p-value)	<0.01	0.08	<0.01	0.08
First stage				
TseTse suitability index (TSI_e)	-0.0870***	-0.1678***	-0.0838***	-0.1609***
	(0.0271)	(0.0373)	(0.0275)	(0.0366)
F-stat of excluded instrument	10.32	20.26	9.27	19.36
Observations	29,581	29,581	15,630	15,630
Individual controls	✓	✓	✓	✓
Location controls	✓	✓	✓	✓
Country FE	✓	✓	✓	✓
Ethnicity-level historical controls	✓	✓	✓	✓
Year of observation FE	✗	✓	✗	✓

Notes: Panel A reports OLS estimates. Panel B reports IV estimates. The top of Panel B reports second stage estimates while the bottom reports first stage estimates. The unit of observation is an individual. Columns (1)-(2) employ the full sample, while columns (3)-(4) employ the "migrants" sub-sample. "Migrants" are defined as individuals not living in their historical ethnic homeland. Individuals controls, location controls and ethnicity-level controls are the same as in Table I. All specifications include country and round fixed effects (constant not reported). Even-numbered columns include year of observation fixed effects. Below the estimates are reported in parentheses robust standard errors clustered at the ethnic group level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Dependent variable: Institutional Trust Index ($ITI_{i,e,l,c}$)	(1)	(2)	(3)	(4)
		"Migrants" only		
Second stage				
Pre-colonial pol. central. (PC_e)	0.0668**	0.0085		
	(0.0262)	(0.0380)		
Juris. hier. beyond local comm.			0.0450**	
			(0.0176)	
City in 1800				0.1479**
				(0.0600)
Endogeneity test (p-value)	<0.01	0.84	<0.01	<0.01
First stage				
TseTse suitability index (TSI_e)	-0.1925***		-0.2862***	-0.0865***
	(0.0393)		(0.0789)	(0.0235)
Malaria ecology index		-0.0217***		
		(0.0077)		
F-stat of excluded instrument	24.01	5.04	13.16	13.51
Observations	15,260	15,260	15,265	15,280
Individual controls	✓	✓	✓	✓
Location controls	✓	✓	✓	✓
Country FE	✓	✓	✓	✓
Ethnicity-level historical controls	✓	✓	✓	✓
Current homeland FE	✓	✓	✓	✓

Notes: The table reports IV estimates. The top panel reports second stage estimates while the bottom panel reports first stage estimates. The unit of observation is an individual. All specifications employ the "migrants" sub-sample. "Migrants" are defined as individuals not living in their historical ethnic homeland. Individuals controls, location controls and ethnicity-level controls are the same as in Table I, with the exclusion of cultural province fixed effects. All specifications include country, round and current homeland fixed effects (constant not reported). The current homeland fixed effects are dummy variables corresponding to the historical ethnic homeland where individuals are currently located, based on the location from the Afrobarometer surveys and Murdock's (1959) ethnolinguistic map. Below the estimates are reported in parentheses robust standard errors clustered at the ethnic group level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Dependent variable:	*0-25%*	*25-50%*	*50-75%*	*75-100%*
Institutional Trust Index ($ITI_{i,e,l,c}$)	(1)	(2)	(3)	(4)

	PANEL A: OLS ESTIMATES			
Pre-colonial pol. central. (PC_e)	0.0378***	0.0319***	0.0322***	0.0260**
	(0.0109)	(0.0106)	(0.0104)	(0.0106)
Proportion of co-ethnics	0.0051	0.0232**	0.0241**	0.0187*
	(0.0133)	(0.0093)	(0.0094)	(0.0102)
Proportion X PC_e	-0.0194**	-0.0009	-0.0018	0.0152*
	(0.0098)	(0.0105)	(0.0083)	(0.0090)
PC_e + *Proportion X PC_e*	0.0184	0.0310**	0.0303**	0.0412***
	(0.0121)	(0.0128)	(0.0127)	(0.0122)

	PANEL B: IV ESTIMATES			
Second stage				
Pre-colonial pol. central. (PC_e)	0.0776**	0.0598**	0.0600**	0.0480
	(0.0303)	(0.0252)	(0.0251)	(0.0366)
Proportion of co-ethnics	-0.0196	0.0214**	0.0189	0.0134
	(0.0364)	(0.0095)	(0.0195)	(0.0229)
Proportion X PC_e	-0.0455	0.0103	0.0069	0.0317
	(0.0368)	(0.0213)	(0.0351)	(0.0657)
PC_e + *Proportion X PC_e*	0.0321	0.0701**	0.0669	0.0797*
	(0.0323)	(0.0321)	(0.0419)	(0.0470)
Endogeneity test (*p*-value)	0.33	0.39	0.42	0.43
F-stats of excluded instruments:				
TSI_e	13.03	13.93	13.54	13.10
Proportion X TSI_e	16.78	6.71	2.49	6.14
TSI_e + *Proportion X TSI_e*	4.97	4.75	2.27	1.64
Observations	29,547	29,547	29,547	29,547
Individual controls	✓	✓	✓	✓
Location controls	✓	✓	✓	✓
Country FE	✓	✓	✓	✓
Ethnicity-level historical controls	✓	✓	✓	✓
Year of observation FE	✓	✓	✓	✓

Notes: Panel A reports OLS estimates. Panel B reports IV estimates. The unit of observation is an individual. Columns (1), (2), (3) and (4) add the interactions of PC_e (in OLS and second-stage IV) and TSI_e (in first-stage IV) with a dummy equal to 1 if 0-25%, 25-50%, 50-75% and 75-100% of the district's sample population is of the same ethnicity as the respondent, respectively. Individuals controls, location controls and ethnicity-level controls are the same as in Table I. All specifications include country, round and year of observation fixed effects (constant not reported). Below the estimates are reported in parentheses robust standard errors clustered at the ethnic group level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

E. Data Description

Institutional Trust Index: See Appendix A. *Source: Afrobarometer. Data can be found at afrobarometer.org.*

Living Conditions: Variable ranging from 1 to 5, based on an Afrobarometer survey question where respondents were asked to rate their living conditions. The answer categories were (i) "Very bad", (ii) "Fairly bad", (iii) "Neither good nor bad", (iv) "Fairly good", and (v) "Very good". *Source: Afrobarometer. Data can be found at afrobarometer.org.*

Education: Variable ranging from 0 to 9, based on an Afrobarometer survey question where respondents were asked what is their education level. The answer categories ranged from the lowest "No formal schooling" to the highest "Post-graduate". *Source: Afrobarometer. Data can be found at afrobarometer.org.*

Ethnic Fractionalisation: Ethnic fractionalisation of the respondent's district. Measure constructed from the district sample population using Afrobarometer survey questions. *Source: Afrobarometer. Data can be found at afrobarometer.org.*

District's Population of same Ethnicity: Share of the district's population that is of the same ethnicity as the respondent. Measure constructed from the district sample population using Afrobarometer survey questions. *Source: Afrobarometer. Data can be found at afrobarometer.org.*

Public Goods Indicators: Variables constructed from Afrobarometer survey questions asking whether electricity, piped water, sewage, health clinics, and schools are available in the respondent's enumeration area (village/town/city). *Source: Afrobarometer. Data can be found at afrobarometer.org.*

Jurisdictional Hierarchy beyond Local Community: Variable ranging from 0 to 4 indicating the number of jurisdictional levels above the local level. A 0 indicates stateless societies, 1 and 2 indicate petty and large chiefdoms, 3 and 4 indicate large states. *Source: Murdock (1967). Original variable code in the Ethnographic Atlas v33.*

Pre-colonial Political centralisation: Binary variable taking the value of 0 if the Jurisdictional Hierarchy beyond Local Community variable equals 0 or 1. The variable takes

on the value 1 if the Jurisdictional Hierarchy Beyond Local Community variable equals 2, 3 or 4. *Source: Murdock (1967). Recoding follows Gennaioli and Rainer (2006, 2007).*

Agricultural Dependence: Variable ranging from 0 to 9, with 0 indicating 0-5% dependence and 9 indicating 86-100% dependence. *Source: Murdock (1967). Original variable code in the Ethnographic Atlas v5.*

Intensive Agriculture Indicator variable equal to one if group practices intensive or intensive irrigated agriculture. The original variable (Agriculture Type) is an index ranging from 0 to 4 reflecting the type of agriculture practised. The index equals 0 when there is "no agriculture", 1 when there is "causal agriculture", 2 when there is "extensive or shifting agriculture", 3 when there is "intensive agriculture", and 4 when there is "intensive irrigated agriculture". *Source: Murdock (1967). Original variable code in the Ethnographic Atlas v28.*

Indigenous Slavery : Indicator that equals one when some type of slavery (hereditary, incipient, or significant) is present and zero when there it is absent or near absent. *Source: Murdock (1967). Original variable code in the Ethnographic Atlas v70.*

Settlement Patterns: Variable ranging from 1 to 8 classifying "settlement pattern of each group". 1 indicates fully nomadic (migratory) groups, 2 indicates semi-nomadic groups, 3 indicates semi-sedentary groups, 4 identifies groups that live in compact and impermanent settlements, 5 indicates societies those in neighborhoods of dispersed family homes, 6 indicates for groups in separated hamlets forming a single community, 7 indicates societies living in compact and relatively permanent settlements, and 8 denotes the groups residing in complex settlements. *Source: Murdock (1967). Original variable code in the Ethnographic Atlas v30.*

Slave Exports: Measure of the number of slaves taken from the ethnic group during the slave trade. Export intensity is measured as the logarithm of 1 plus the total slaves exported, normalized by historical ethnic homeland land area. *Source: Nunn and Wantchekon (2011). Dataset can be found at scholar.harvard.edu/nunn.*

City in 1800: Indicator variable constructed from city location geospatial data for the year 1800, combined with Murdock's (1959) ethnolinguistic map. Variable equals one if there was a city with more than 20,000 inhabitants located on the land inhabited by each ethnic group. *Source: Chandler (1987).*

Malaria Ecology Index: Measure of the malaria disease environment of each historical ethnic homeland. Constructed by combining the index by Kiszewski et al. (2004) with Murdock's (1959) ethnolinguistic map. *Source: Nunn and Wantchekon (2011). Data can be found at scholar.harvard.edu/nunn.*

TseTse Suitability Index: Suitability to the TseTse fly of area corresponding to the historical ethnic homelands from Murdock's (1959) ethnolinguistic map. *Source: Alsan (2015). Data can be found at people.stanford.edu/malsan.*

Agricultural Suitability Index: Suitability for rain-fed crops, normalised to range from 0 to 1, with higher values indicating greater suitability. Calculated for the historical ethnic homelands from Murdock's (1959) ethnolinguistic map. *Source: FAO's Global Agro-Ecological Zones. Data can be found at fao.org/nr/gaez.*

Log Land Area: Logarithm of land area of historical ethnic homeland based on Murdock's (1959) ethnolinguistic map, measured in km^2. *Source: Calculated using ArcGis.*

Latitude: Latitude of historical ethnic homeland. *Source: Murdock (1967). Original variable code in the Ethnographic Atlas v104.*

Longitude: Longitude of historical ethnic homeland. *Source: Murdock (1967). Original variable code in the Ethnographic Atlas v106.*

Coast: Indicator equalling one if the historical ethnic homeland's boundaries included a coast, based on Murdock's (1959) ethnolinguistic map. *Source: Calculated using ArcGis.*

Cultural Province: Grouping capturing spatial and cultural/genealogical correlation of ethnic groups. See Appendix C for details. *Source: Murdock (1967). Chapter headings of Ethnographic Atlas.*

Year of Observation: Year when the characteristics of the ethnic group were recorded by ethnographers. See Appendix C for details. *Source: Murdock (1967). Original variable code in the Ethnographic Atlas v102.*