# LOCAL STATE HISTORY AND CONTEMPORARY CONFLICT: EVIDENCE FROM SUB-SAHARAN AFRICA*

Emilio Depetris-Chauvin†

October 24, 2013

## JOB MARKET PAPER [Preliminary Version]

**Abstract**

I examine empirically the role of historical political centralization on the likelihood of modern civil conflict in Sub-Saharan Africa. I combine a wide variety of historical sources to construct an original measure of long-run exposure to statehood at the local level. I then exploit variation in this new measure along with geo-referenced conflict data to document a robust negative statistical relationship between local long-run exposure to statehood and contemporary conflict. I argue that locations with long histories of statehood are better equipped with mechanisms to establish and preserve order. I provide two pieces of evidence consistent with this hypothesis. First, locations with relatively long historical exposure to statehood are less prone to experience conflict when hit by a negative shock to the agricultural sector. Second, exploiting contemporary individual-level survey data for 18 Sub-Saharan countries, I show that within-country long historical experience with statehood is linked to people's positive attitudes toward state institutions and traditional leaders.

*I would like to thank David Weil, Pedro Dal Bó, and Stelios Michalopoulos for their guidance and encouragement on this project. Also, I wish to thank Oded Galor, Louis Putterman, Ola Olsson, Alejandro Molnar, Boris Gershman, participants of Macroeconomics Lunch and Macroeconomics Seminar at Brown University, and seminar participants at Universidad de San Andres for comments and helpful discussions. I am also grateful to Ömer Özak and Stelios Michalopoulos for sharing data, Nickolai Riabov and Lynn Carlson for computational assistance with ArcGIS and R, and Santiago Begueria for helpful discussions and suggestions regarding the Standardized Precipitation Evapotranspiration Index. All errors are my own.

†Brown University, Department of Economics, Graduate School, Robinson Hall, 64 Waterman Street, Providence, RI 02912, USA. email: Emilio_Depetris_Chauvin@Brown.edu

# 1 Introduction

Civil conflict imposes enormous costs on a society. On the humanitarian side, beyond the lives lost as a result of direct violent confrontations, there are negative consequences on health and social fragmentation. On the economic side, civil conflict may not only directly affect the short-run economic performance by disrupting markets but it may also affect long-run growth fundamentals such as human capital accumulation, income inequality, institutions, and culture. Not surprisingly, understanding the determinants of civil conflict through the lens of economics and its toolkit has been the aim of a growing body of economic literature.[1]

The particular case of Sub-Saharan Africa has received considerable attention from academics for the simple reason that civil conflict is very prevalent in this part of the world; over two thirds of Sub-Saharan african countries experienced at least one episode of conflict since 1980. Easterly and Levine (1997), Collier and Hoeffer (1998), and Fearon and Laitin (2003), among others, have pointed to civil conflict as a key factor holding back African economic development.

In this paper I explore the relationship between the prevalence of modern civil conflict and historical political centralization. Specifically, I uncover a within-country robust negative statistical relationship between long-run exposure to statehood and the prevalence of contemporary conflict. My approach of looking at a historical determinant of modern civil conflict is motivated by the empirical literature showing evidence on the importance of historical persistence for understanding current economic development (see Nunn 2013 and Spoloare and Wacziarg 2013 for extensive reviews). My paper draws on a particular strand of this literature which documents that traditional African institutions not only survived the colonial period but that they still play an important role in modern African devolopment (Gennaioli and Rainer, 2007; Nunn and Wantchekon, 2011; and Michalopoulos and Papaioannou, 2013, among others).

A key aspect of my approach is to exploit within country differences the prevalence of modern conflict and its correlates. I take my empirical analysis to a fine local scale for several reasons. First, conflict is a local phenomenon which does not usually expand to the whole territory of a country.[2] Second, there is arguably a large within-country heterogeneity in

[1] Blattman and Miguel (2010) provides an extensive review and discussion on the literature of civil conflict, including the theoretical approaches and most important empirical findings on the causes and consequences of civil conflict.

[2] Raleigh et al (2009) argue that civil conflict does not usually expand across more than a quarter of a

historical determinants of conflict. Given that modern borders in Sub-Saharan Africa were artificially drawn during colonial times without consideration of previous historical boundaries (Green, 2012), substantial heterogeneity in location histories and people characteristics persists today within those borders. Therefore, the country aggregation of these characteristics may artificially erase a rich source of heterogeneity. Third, other conflict determinants previously highlighted in the literature, such as weather anomalies or topography, are in fact geographical and location-specific. Fourth, exploiting within-country variation in deeeply-rooted institutions allows me to abstract from country-level covariates, such as national institutions or the identity of the colonial power that ruled the country.

Pre-colonial Sub-Saharan Africa comprised a large number of polities of different territorial size and varying degrees of history of political centralization (Murdock, 1967). At one extreme of the spectrum one can distinguish large states, such as Songhai in modern day Mali, which had professional armies, public servants and formal institutions (i.e., courts of law and diplomats). On the other extreme, one can find groups of nomadic hunter-gatherers with no formal political head such as the Bushmen of South Africa. Historical political centralization varies even within countries. Consider, for example, the case of Nigeria, where the Hausa, the Yoruba, and the Igbo represent almost 70 percent of the national population and have quite different histories of centralization. Unlike the Hausa and Yoruba, the Igbo had a very short history of state centralization in pre-colonial time despite having been settled in southern Nigeria for centuries. In order to account for this heterogeneity in historical state prevalence, I introduce an original measure which I refer to as the state history index at the local level. To do this, I combine a wide variety of historical sources to identify a comprehensive list of historical states, along with their boundaries and chronologies. In its simplest version, my index measures a given territory's fraction of years under indigenous state-like institutions over the time period 1000 - 1850 CE.

I then document a within-country strong negative correlation between my state history index and geo-referenced conflict data. My OLS results are robust to a battery of within modern countries controls ranging from contemporaneous conflict correlates and geographic factors to historical and deeply-rooted plausible determinants of modern social conflict such as slave trade prevalence, historical exposure to conflict, genetic diversity or ecological drivers of historical trade. These results are not driven by historically stateless locations, influential observations, heterogeneity across regions, or the way conflict is coded. Nonetheless, this robust statistical association does not neccesarily imply causality. Indeed, history is not a

country's territory.

random process in which long-run exposure to statehood has been randomly assigned across regions. The historical formation and evolution of states is a complex phenomenom. Factors underlying the emergence and persistence of states may still operate today. To the extent that some of those factors are unobserved, isolating the causal effect of long experience with statehood on conflict is a difficult task. I aim, however, to convince the reader here that local history has left its mark on the pattern of contemporaneous conflict. Although I cannot rule out the possibility that other hard-to-account-for factors are partially explaining this strong statistical association, I argue that it is unlikely that my results are fully driven by those omitted factors. Following Altonji, Elder, and Taber (2005)'s approach I show that the influence of unobservables would have to be considerable larger than the influence of observables to explain away the uncovered correlation.

Beyond selection on unobservables, there exists an additional potential source of bias in the OLS estimates. Given the obvious limitations to documenting historical boundaries in Sub-Saharan Africa, a high degree of measurement error is likely to be present in my index. To tackle this limitation, I follow an instrumental variable approach based on the proposed link between the neolithic revolution and the rise of complex political organization. Using archaeological data containing the date and location of the earliest evidence of crop domestication I construct a measure of the time elapsed since the neolithic revolution at a fine local level. My IV estimates suggest an even larger statistical association between local long-run exposure to statehood and the prevalence of contemporary conflict; this finding is consistent with the idea that measurement error in my state history index was introducing a sizeable bias toward zero. To address concerns regarding the validity of the exclusion restriction, I examine whether my IV estimates are affected by the inclusion of several controls such as ecological diversity, migratory distance to Addis Adaba (a strong predictor of genetic diversity), absolute latitude, soil suitability to grow the most relevant indigenous crops, and intertemporal climate variability. The addition of this rich set of covariates does not qualitatively affect my results. To further investigate the extent of the bias from measuremente error, I follow a second instrumental variable strategy and exploit a panel of large African cities for the time period 1000-1850 CE to construct an additional measure of the degree of influence from centralized states. I obtain similar results using this alternative measure which provides additional support to my main hypothesis.

Why would the long history of statehood matter for contemporary conflict? The accumulation of years of experience with state-like institutions may result in improved social cohesion over time. In this vein, I argue that locations with long histories of statehood are better equipped with mechanisms to establish and preserve order. These institutional capabilities can be manifested, inter alia, in the ability to negotiate compromises and allocate scarce resources, the existence of traditional collective organizations and legal courts to peacefully settle differences over local disputes, and a stronger presence of police force or even a more autocratic regime. I provide two additional pieces of evidence consistent with this hypothesis. First, I exploit panel data to show that locations with relatively long historical exposure to statehood are remarkably less prone to experience conflict when hit by a negative agricultural productivity shock. Second, I exploit contemporary individual-level survey data for 18 Sub-Saharan countries to present empirical evidence that a long history of statehood can be linked to people's positive attitudes towards state institutions. In this sense, I show that key state institutions are regarded as more legitimate and trustworthy by people living in district with long history of statehood. These results are not driven by unobservable ethnic characteristics (i.e; estimates are conditional on ethnic fixed effect), which constitutes a striking result and suggests that the institutional history of the location where people currently live matter for people's opinion about state institutions independently of the history of their ancestors.

This paper belongs to a vibrant and emerging strand of literature within economics tracing the historical roots of contemporary development. Specifically, my work is related to economic research on the relationship between institutional history and contemporary outcomes; a line of research which originates in Engerman and Sokoloff (1997), La Porta e al. (1999), and Acemoglu, Johnson, and Robinson (2001). In particular, this paper is related to the literature examining the developmental role of state history. It is methodologically related to Bockstette, Chanda, and Putterman (2002) which introduces a State Antiquity Index at the country level.[3] I contribute to the related literature by constructing an original measure at the local level. My work connects to Hariri (2012) who presents cross-country evidence of an autocratic legacy from early statehood.[4] In this vein, this autocratic legacy may arguably be one of the mechanisms underlying my findings since an enhanced local autocratic leader can have better abilities to establish and preserve order. In fact, I show suggestive evidence that history of statehood can be linked to popular support for traditional leaders.

Particularly in the context of Africa, my work is also related to works on the impact of

[3]Bockstette, Chanda, and Putterman (2002) introduces the State Antiquity Index and shows that it is correlated with indicators of institutional quality and political stability at the country level. Borcan, Olsson, and Putterman (2013) extends the original index back to 4th millenium BCE.

[4]Hariri (2012) argues that indigenous institutions limited the European diffusion of democracy and its related institutions.

pre-colonial political centralization on contemporary outcomes (Gennaioli and Rainer, 2007; Huillery, 2009, Michalopoulos and Papaioannou, 2013). More importantly, my work contributes to the line of research on how historical factors shaped pattern of conflict during the African post-colonial era (Michalopoulos and Papaioannou 2011, and Besley and Reynal-Querol 2012).[5] Of most relevance to my work is Wig (2013) which is, to the best of my knowledge, the first study looking at the relationship between pre-colonial political centralization and conflict. The author finds that ethnic groups with high pre-colonial political centralization and are not part of the national government are less likely to be involved in ethnic conflicts than groups without pre-colonial political centralization.[6] While attempting to address a similar question on how historical political centralization may prevent conflict, there are two main differences between Wig (2013) and my work. First, I do not restrict my analysis to ethnic conflict; rather, I study a more general definition of civil conflict. Second, unlike Wig (2013) who only focuses on ethnic political centralization recorded by ethnographers around the colonization period, I trace the history of statehood further back in time to account for differences on long-run exposure to statehood. In this vein, I argue that not only the extensive but also the intensive margin of prevalence of historical institutions matters crucially to understand contemporary conflict.

This paper also speaks to cross-country literature on the role of contemporary institutions to prevent conflict.[7] Methodologically, I depart from the cross-country empirical approach. Rather than focusing on contemporary institutional differences at the national level, I investigate the role of deeply-rooted institutional characteristics at the local level. In revealing how a deeeply-rooted factor relates to contemporary conflict, this paper also connects to recent work by Arbatli, Ashraf, and Galor (2013), which shows that genetic diversity strongly predicts social conflict.[8] Finally, my work is also methodologically related to recent litera-

[5]Michalopoulos and Papaioannou (2011) exploits a quasi-natural experiment to show that civil conflict is more prevalent in the historical homeland of ethnicities that were partitioned during the scramble for Africa. Besley and Reynal-Querol (2012) provides suggestive evidence of a legacy of historical conflict by documenting a positive empirical relationship between pattern of contemporary conflict and proximity to the location of recorded battles during the time period 1400 - 1700 CE.

[6]Based on bargaining theories of civil conflict, Wig (2013) hypothesizes that the aforementioned ethnic groups can rely on traditional institutions to strike credible non-violent bargains with the national government.

[7]For instance, the importance of state capacity (Fearon and Laitin 2003 and Besley and Persson 2008) and cohesive political institutions (Besley and Persson 2011, Collier, Hoeffler, and Soderbom 2008) has been emprically studied.

[8]To the extent that my main results are robust to the inclusion as a control of migratory distance from the location of human origin (a strong linear predictor of genetic diversity), my findings can be interpreted as consistent with and complementary to the empirical regularities documented in Arbatli, Ashraf, and Galor (2013).

ture in economics that takes a local approach to conflict (Besley and Reynal-Querol, 2012; Laferrara and Harari, 2012).

# 2 A New Index of State History at the Local Level

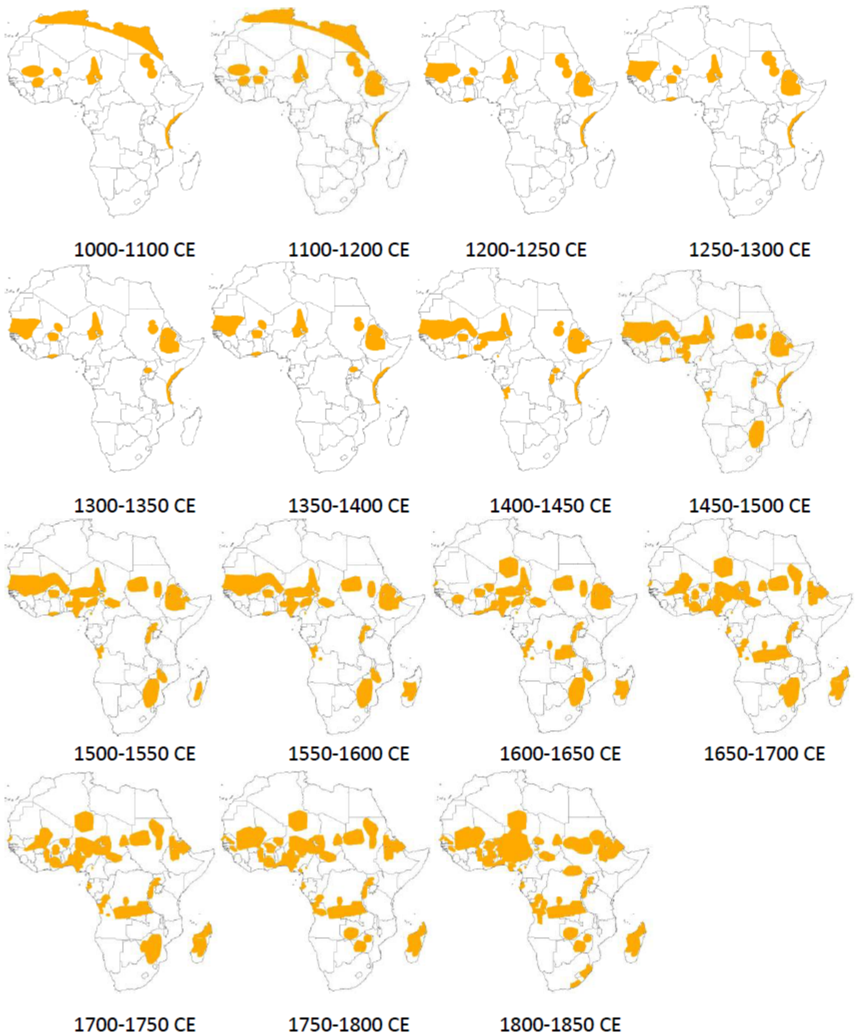
## 2.1 Overview of the Construction Procedure

In this section I present an overview of the construction procedure of my new index of state history at the local level. Two dimensions are relevant for my purpose; the time period to consider for the computation of the index and the definition of a geographical location for which the index is calculated. That is, I have to define the units of analysis that will determine the scope of both the extensive and the intensive margin of state history.

*Time period under analysis.* I focus on the period 1000-1850 CE for two reasons. First, the aim of my research is to examine the legacy of indigenous state history, thus I consider only pre-colonial times. I am not neglecting, however, the importance of the colonial and post-colonial periods to understand contemporary pattern of conflict. In fact, the persistence of most of the indigenous institutions during and after the colonial indirect rule experience represents an important part of the main argument in this paper. Second, I ignored years before 1000 CE due to the low quality of historical information and to the fact that no much known variation on historical states would have taken place in Sub-Saharan Africa before that period.[9] I then follow Bockstette, Chanda, and Putterman (2002), and divide the period 1000-1850 CE in 17 half-centuries. For each 50 years period I identify all the polities relevant for that period. I consider a polity to be relevant for a given half-century period if it existed for at least twenty six years during that fifty-years interval. Therefore, I construct seventeen cross sections of the historical boundaries previously identified in the pre-colonial Sub-Saharan Africa. Figure 1 displays the evolution of historical map boundaries over the period 1000-1850 CE.

[9]There would have been few cases of state formation before 1000 CE in Sub-Saharan Africa: the Aksum and Nubian Kingdoms (Nobadia and Alodia) in the Ethiopian Highland and along the Nile river, the Siwahalli City-States in East Africa, Kanem in Western Chad, and Ghana and Gao in the West African Sahel (Ehret 2002).

Figure 1: Evolution of Historical Map Boundaries (1000 - 1850 CE)




*Definition of geographic unit.* My empirical analysis focuses on two different definitions for local level (i.e: geographical unit of observation). First, I focus on grid cells; an artificial constructions of 2 by 2 degrees. Second, I also focus on African districts and thus construct a 1-degree radius buffer around the centroid of each district.[10] Given these different levels of aggregation to compute local state history, I start by constructing the index at a sufficient fine level. Therefore, I divide Sub-Saharan Africa in 0.1 by 0.1 degree pixels (0.1 degree is

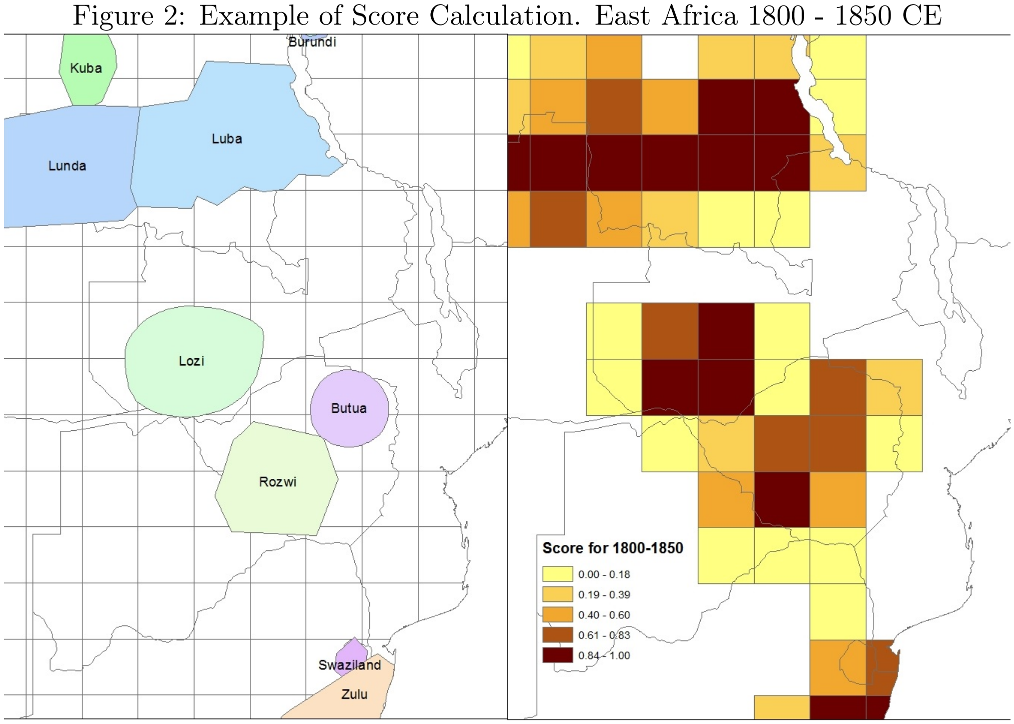
[10]A district is a second order administrative division with an intermediate level of disagregation between a region or province and a village.

approximately 11 kilometers at the equator). I then dissolve the compiled historical maps into 0.1 by 0.1 degree pixels taking the value 1 when an historical state intersects the pixel, and 0 otherwise.[11] For a given level of aggregation $i$, its local state history value would be determined by:

$$State\ History_i = \sum_{1000}^{1850} \beta \times S_{i,t}$$

where, $S_{i,t} = \frac{\sum \theta_{p,t}}{P}$ is the score of $i$ in period $t$, with $\theta_{p,t}$ taking the value 1 if the pixel $p$ is intersected by the map of an historical state in period $t$, 0 otherwise; and $P$ being the number of pixels in $i$.[12] The variable $\beta$ is the discount factor. Since I do not have any theoretical reason to pick a particular discount factor, I base most of my analysis in a discount factor of 1. Figure 2 shows an example of the calculation of the score in East Africa circa 1800 when the level of aggregation is a grid cell of 2 degree by 2 degree.

Figure 2: Example of Score Calculation. East Africa 1800 - 1850 CE




There are three crucial and challenging pieces in the construction of the index. First, my procedure requires the compilation of a comprehensive list of historical states. Second,

[11]Therefore, the pixel will take a value 1 even when an overlap of two historical states exists. That is, a pixel intersected multiple times is considered only once.

[12]Therefore, the score $S_{i,t}$ denotes what fraction of the territory of $i$ is under an historical state in the period $t$.

the boundaries of those historical states have to be identified, digitized and georeferenced. Third, an even more difficult task is to account for potential expansions and contractions of those boundaries over time.

*Identifying historical states.* I use a wide variety of sources to identify historical maps of states in pre-colonial Sub-Saharan Africa for the time period 1000-1850 CE.[13] Identifying what constituted a state in the remote past of Africa is not a easy task. Of course, historical records are incomplete and some time the demarcation between tribes and kingdoms was not that clear. Further, heterogeneity in political structures was indeed very large in pre-colonial Africa. Nonetheless, my operative definition of states includes city-states, kingdoms, and empires and it is built upon the conception of a centralized power exercising influence over some periphery. That is, a historical state is the result of the amalgamation of smaller settlement units in a relatively large unit of territory ruled by centralized political head. I consider the existence of an army as a necessary but not sufficient condition to constitute a state. For instance, the Galla people (also known as Oromo) in modern Ethiopia developed states only two hundred years after conquering ethiopian soil (Lewis, 1966). Before founding the five Gibe kingdoms, Galla people were governed at the village level. Although coordinated in the competition against neighboring kingdoms, each local independent group was under its own leader. Thus, I only considered the Galla's polities once the Gibe kingdoms were established in late eighteenth century. Note that my notion of state is not necessary a proxy for societal complexity.[14] Non-political centralized complex societies such as the Igbo in modern Nigeria, which had a complex system of calendars (Ògụafọ) and banking (Isusu), are not considered as historical states. In fact, only after conforming the trade confederacy in the year 1690, I consider the Aro, a subgroup of the Igbo, to be taken into account for the computation of my index.

The starting point then was to identify the historical states referenced in the version 3.1 of the State Antiquity Index introduced in Bockstette, Chanda, and Putterman (2002). I complement this initial list with a variety of additional sources (Ajayi and Crowler 1985, Barraclough 1979, Vansina 1969, McEvedy 1995, Murdock 1967, and Ehret 2002). Once I complete the list of all the polities to be taken into account in the computation of my state

[13]I define Sub-Saharan Africa to all the geography contained within the borders of the following countries: Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Congo DRC, Congo, Cote d'Ivoire, Ethiopia, Eritrea, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, Somalia, South Africa, Sudan, Swaziland, Tanzania, Togo, Uganda, Zambia, and Zimbabwe.

[14]Note also that stateless does not imply either absence of laws or existence of a small societies. The Nuer of the Souther Sudan and the Tiv of Nigeria serve as good examples.

history index, I document approximate dates of foundation and declination of each polity. Table A.1 in the appendix includes the complete list of polities (with their relevant dates) used in the computation of my index. Note that I only consider indigenous states in my analysis. Therefore, I do not consider foreign states such as the Portuguese colony in the coastal strip of Angola (present for more than four hundred years) or occupations such as Morocco's in Songhai's territory at the beginning of the seventeenth century.

*Compilation of historical maps.* The following task was to identify, digitize and georeference the maps of the historical states on the list. Some of the maps were already digitized and some of them were also georeferenced.[15] When a map of a given polity was available for more than one period of time, I took into account all of them for my analysis. This helps me to partially account for expansions and contractions of states' geographic influence over time.[16] Some judgment was needed when two sources disagreed in the way the boundaries of a historical state were recorded for a similar historical period. I kept the map I found more reliable.[17] I abstract for now from the difficulties (and consequences) of defining historical map boundaries; I discuss this issue below in more detail.

*Cross-Sectional Variation.* Figure 3 displays the cross-sectional variation of my State History Index based on grid-cell aggregation (with a discount factor of 1). Sub-Saharan Africa is divided in 558 grid cells of 2 degree by 2 degree. Following Bockstette, Chanda, and Putterman (2002), I rescale the index by dividing all the values by the maximum possible value; therefore $State\, History_i \in [0, 1]$.

Roughly one third of Sub-Saharan Africa has no state history before 1850 CE. State history is more prevalent in the north, particularly in western part of Sahel, the highlands of Ethiopia, and the region along the Nile river. In this sense, proximity to water is a relevant factor to explain the historical presence of states. In particular, proximity to major rivers such as Niger, Benue, Senegal, Volta, Congo, and Zambezi; and great lakes such as Victoria, Tanganyika, Malawi, and Chad correlates with high values of the index. Almost no state

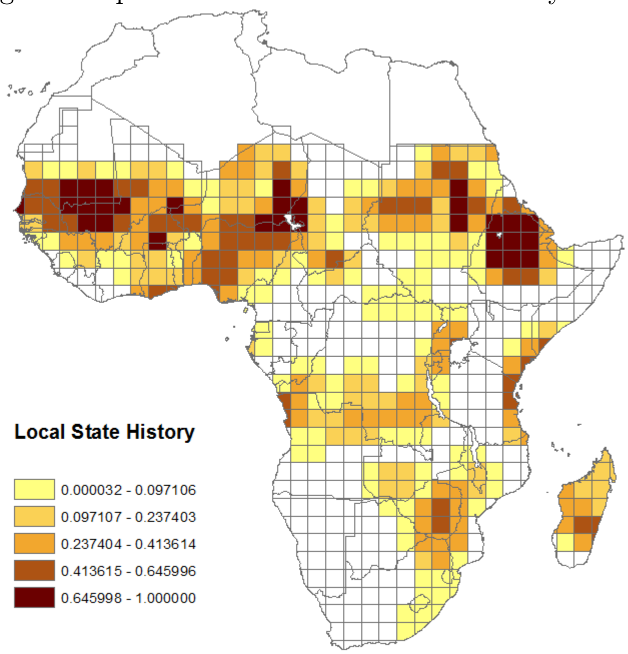
[15] For instance, some maps from McEvedy's (1995) Atlas of African History were already digitized and georeferenced by AfricaMap, a project developed by the Center for Geographic Analysis at Harvard. After checking for inconsistencies with original sources and correcting irregularities in border drawings, I also considered some maps digitized by the ThinkQuest Project of The Oracle Education Foundation.

[16] For instance, I was able to document how political influence of Songhai's people evolved over my period of analysis. Figure 1 includes the first Songhai polity (pre-imperial) during the time period c.1000-c.1350CE around the city of Gao, its expansion consistent with the establishment of the Songhai Empire from c.1350 CE to c.16000CE and the late formation of Dendi Kingdom as a result of the Morrocan invasion and declination of the empire in c.1600 CE.

[17] In some cases I made the decision based on the consistency with natural borders like majors rivers or elevations.

history is documented in the African rainforest and South-West Africa.

Figure 3: Spatial Distribution of State History Index




*Major sources of measurement error.* Any attempt to rigorously define state boundaries in pre-colonial Africa is doomed to imperfection for several reasons. Indigenous historical records are scarce in Sub-Saharan Africa; and most of the modern reconstruction of African history relies upon account by travelers, traders and missionaries (particularly during the nineteenth century), the transmission from oral history, or analysis of archaeological sites. Further, this scarcity of historical records exacerbates the farther south or away from the coast one looks. Most importantly perhaps, almost no indigenous map making existed in pre-colonial Africa (Herbst, 2000). Regardless of the problems due to lack of historical records, the extension of authority to the periphery in pre-colonial Africa was itself irregular, contested, and weak. As argued by Herbst (2000), boundaries were, in consequence, a reflection of this difficulty of broadcasting power from the center. Therefore, the lines of demarcation for boundaries of any historical state are, by construction, inevitably imperfect. As a matter of fact, I find different historical atlases displaying quite dissimilar maps for the same polity under similar period of time. Nevertheless, while bearing in mind the aforementioned caveat, documenting imperfect boundaries provides at least a useful starting point for my empirical analysis.

The aforementioned imperfection in the demarcation of boundaries represents a source of measurement error affecting my econometric analysis. There is little reason to believe that this particular measurement error is correlated with the true measure of state antiquity. Therefore, this would represent a case of classical errors-in-variables that would introduce an attenuation bias in the OLS estimates of the relationship between historical state prevalence and conflict.

An additional source of measurement errors in my state history variable will result from the introduction of an upper bound when computing the index. When considering only historical states starting 1000 CE, I am excluding many years of state history in region with long history of statehood. For instance, I am omitting more than 250 years of the Ghana empire in West Africa. Further, the Kingdom of Aksum, existing during the period 100-950 CE and located in modern day Eritrea and Ethiopia, was not considered in the computation of the state history index. Since locations with some history of state before 1000 CE tend to present high values of my index, the introduction of the bound in the period of analysis for its computation would tend to underestimate the long run exposure to statehood for some regions. Therefore, an additional upward bias in the OLS estimates is introduced. It is precisely for the sake of alleviating the resulting biases due to measurement error in my data what will provide a key motivation for the implementation of an instrumental approach later on.

# 3 Empirical Relationship between State History and Contemporary Conflict

## 3.1 Sources and Description of Conflict Data

In this paper I exploit georeferenced conflict event data to construct different measures of conflict prevalence at the local level. There are two leading georeferenced conflict datasets for Sub-Saharan Africa, the Uppsala Conflict Data Program Georeferenced Events Dataset (UCDP GED, from now on) and the Armed Conflict Location Events Dataset (ACLED, from now on). For reasons I detail below, the core of my analysis is based on UCDP GED. However, I show that the main results are not dependent on the choice of the conflict dataset.

The UCDP GED, version 1.5 (November 2012) provides geographically and temporally disaggregated data for conflict events in Africa (for a full description of the dataset, see Sundberg and Melander, 2013). Specifically, UCDP GED provides the date and location (in latitude and longitude) of all conflict events for the period 1989-2010. A conflict event is defined as "the incidence of the use of armed force by an organized actor against another organized actor, or against civilians, resulting in at least one direct death in either the best, low or high estimate categories at a specific location and for a specific temporal duration" (Sundberg et al, 2010). The dataset comprises of all the actors and conflicts found in the aggregated, annual UCDP data for the same period. UCDP GED traces all the conflict events of "all dyads and actors that have crossed the 25-deaths threshold in any year of the UCDP annual data" (Sundberg et al, 2010). Note that the 25-deaths threshold is the standard coding to define civil conflict and that the definition for dyad does not exclusively need to include the government of a state as a warring actor. Finally, also note that once a dyad crossed the 25-deaths threshold, all the events with at least one death are included in the dataset. That is, these events are included even when they occurred in a year where the 25-deaths threshold was not crossed and regardless of whether they occurred before the year in which the threshold was in fact crossed. The UCDP GED contains 21,858 events related to approximately 400 conflict dyads for the whole African continent. More than 50 percent of those events include the state as one of the warring actors (although only about 10 percent of conflict dyads included the state). For the best estimate category, the total fatality count is approximately 750,000 deaths (Sundberg and Melander, 2013).

I prefer UCDP GED over ACLED for several reasons. First, the definition of conflict event in UCDP GED is restricted to fatal events and it adheres to the general and well established definitions in UCDP–PRIO Armed Conflict Dataset, which has been extensively used in the conflict literature (see for example, Miguel et al, 2004, and Esteban et al, 2012). On the contrary, the definition of event in ACLED includes non-violent events such as troop movements, the establishment of rebel bases, arrests, and political protests. Moreover, the definitions of armed conflict and what constitutes an event in ACLED is not fully specified. This is indeed worrisome because it makes harder to understand the potential scopes of measurement errors in the conflict data. Nonetheless, ACLED data does allow the user to identify battle and other violent events. Second, UCDP GED provides an estimate of number of casualties per event that allows me to calculate an alternative measure of conflict intensity. Third, Eck (2012) argues that ACLED presents higher rates of miscoding. Fourth, the UCDP GED provides a larger temporal coverage (22 years vs 14 years in ACLED).

Despite of my aforementioned reasons to choose UCDP GED over ACLED, the latter has been recently used by economists (see, for instance, Harari and Laferrara 2013, Besley and Reynal-Querol 2012, and Michalopoulos and Papaioannou 2012). Therefore, I show as a robustness check exercise that using ACLED data does not qualitatively affect the main results of my empirical exercise.

### 3.2 Cross Sectional Evidence

I start my empirical analysis by looking at the statistical relationship between prevalence of conflict and state history at the 2 by 2 degree grid cell level. The key motivation to have an arbitrary construction (i.e., grid cell) as unit of observation, as opposed to subnational administrative units, is to mitigate concerns related to the potential endogeneity of the borders of those political units. In particular, political borders within modern countries may be a direct outcome of either patterns of contemporary conflict or any of its correlates (such as ethnic divisions).[18]

Table 1 presents summary statistics of the 558 grid cells in my sample. The average area of a grid cell in my sample is 42,400 square kilometer which represents approximately one tenth of the average size of a Sub-Saharan African country. A mean conflict prevalence of .189 implies that, during the period 1989-2010, an average grid cell experienced 4 years with at least one conflict event. Approximately one fourth of the grid cells had at least one conflict onset.[19]

[Insert Table 1 here]

I now turn to the analysis of the empirical relationship between local state history and contemporary conflict at the grid cell level. I begin by estimating the following baseline

[18] The determination of the spatial resolution (i.e., size and position of the unit of observation) may be subject to the modifiable areal unit problem (MAUP), which may affect the results due to the potential existence of an statistical bias resulting from the scaling and zoning methods (see Wrigley et al, 1996). Zoning does not appear to quite relevant in the study of conflict at the grid cell level (Hariri and Laferrara, 2013). I pragmatically centered the northwesternmost grid cell so it perfectly corresponds with the raster of gridded population data (originally in a resolution of 2.5 arc-minutes -aproximately 5km at the equator-). The election of the size of the unit of observation is a more delicate issue. Choosing a higher resolution facilitates the identification of local factors affecting the prevalence of conflict. However, a higher resolution may not only exacerbate measurement error but also make spatial dependence more relevant for the identification of local effects. On the one hand, higher resolution would make nearby observations more dependent of each other, thus introducing potential underestimation of standard errors of point estimates. This is an issue that can be addressed by implementing spatially robust or clustered robust estimation methods. On the other hand, spatial dependance in the dependent variable is more problematic since neglection of this dependence would bias the point estimates. Therefore, when choosing the size of my unit of observation, I attempt to balance the trade-off between masking subnational heterogeneity and introducing a potential bias due to spatial dependence. I acknowledge that implementing a 2 by 2 degrees approach does not completely overcome spatial dependence issues. In ongoing work, I explore some of these issues.

[19] Conflict onset is defined as the first event within a dyad.

equation:

$$Conflict_{i,c} = \alpha + \beta State\,History_i + G_i^{'}\Gamma + X_i^{'}\Delta + C_i^{'}Z + \lambda_c + \epsilon_{i,c} \qquad (1)$$

where $i$ and $c$ denote grid cells and countries respectively. The variable $Conflict_{i,c}$ is a measure of conflict prevalence and represents the fraction of years with at least one conflict event during the period 1989-2010 for the grid cell $i$ in country $c$. The variable $State\,History_i$ is my new index for state history at the local level $i$. Therefore, $\beta$ is the main coefficient of interest in this exercise. The vector $G_i^{'}$ denotes a set of geographic and location specific controls. The vector $X_i^{'}$ includes a set of controls for ecological diversity and a proxy for genetic diversity. $C_i^{'}$ is also a vector and includes potential confounding variables which may be also arguably outcomes of historical state formation. Thus, including these variables may result in a potential bad control problem (see Angrist and Pischke 2009, for discussion). Finally, $\lambda_c$ is country $c$ fixed effect included to account for time-invariant and country-specific factors, such as national institutions, that may affect the prevalence of conflict.[20]

**OLS Estimates**

Table 2 provides a first statistical test to document a strong negative correlation between state history and contemporary conflict at the local level. Below each estimation of my coefficient of interest I report four different standard errors. To start with and just for sake of comparison I report robust standard errors which are consistent with arbitrary forms of heterokedasticity. I also report standard errors adjusted for two-dimensional spatial autocorrelation for the cases of 5 degrees and 10 degrees cut-off distances.[21] I finally report standard errors adjusted for clustering at the country level. For all the specifications in Table 2 standard errors clustered at the country level are much larger than under the other alternative methods. This pattern holds for all the specifications presented in this paper. Therefore, clustering at the country level appears to be the most conservative approach to avoid over-rejection of the null hypothesis regarding the statistical significance of the coeffi-

[20] Each grid cell is assigned to exclusively one country when defining country dummies. When one grid cell crosses country borders it is assigned to the country with the largest share on the grid cell. Given the relevance of proximity to international borders as a correlate of conflict, for the remainder of the paper I will control for a variable indicating the number of countries intersected by each grid cell.

[21] I follow Conley (1999)'s methodology in which the asymptotic covariance matrix is estimated as a weighted average of spatial autocovariances where the weights are a product of kernel functions in North-South and East-West dimensions. These weights are zero beyond an specified cutoff distance. I consider 3 cutoffs distances, namely 3, 5, and 10 degrees.

cient of interest. For the remainder of this paper, I report standard errors and statistics of the hypothesis test that are robust to within-country correlation in the error term.

I now turn to the analysis of the estimates in Table 2. For the first column I only focus on the statistical relationship between state history and conflict after controlling for country dummies. The point estimate for $\beta$ suggests a negative (albeit statistically insignificant at standard levels of confidence when clustering standard errors at the cuntry level - p-value = .14) correlation between state history and conflict at the local level.

In column 2 I add a vector of geo-strategic controls that may also correlate with historical prevalence of states.[22] Distances to the ocean and the capital of the country are intended to proxy the peripheral location of the grid cell. To further account for the possibility of within-country variation in national state penetration, I also control for terrain's characteristics (i.e: elevation and ruggedness) that were highlighted in previous literature (see, for example, Fearon and Laitin 2003, and Cederman 2008). Distance to a major river, total length of major roads, and a capital city dummy are also included to account for their geo-political relevance as main targets for conflict actors.[23] Total area of the grid cell is also included among the controls as well as an indicator of the number of countries intersecting each grid cell. The latter accounts for the fact that conflict is more prevalent near international borders (see, for instance, Michalopoulos and Papaioannou, 2012) whereas the former accounts for the smaller size of coastal grid cells. A positive correlation between income from natural resources and conflict has been extensively documented (see, for example, Fearon, 2003, Collier and Hoeffler, 2004, and Fearon and Latin, 2005). Thus I add a dummy variable taking the value one if at least one natural resource site (i.e: gems, diamond, gas or oil) is located in the grid cell. It is worth noting that most of these controls also help to explain within-country variation in economic development.

All the point estimates (not shown) for the geo-strategic controls present same sign as previously documented in conflict literature (see, in particular, Harari and Laferrara 2013 for a cross-sectional analysis based on grid cells). More importantly, the point estimate for $\beta$ suggests an statistically significant negative relationship between state history and

[22]By geo-strategic dimension I refer to geographical or geo-political characteristic of the grid cell that may affect the likelihood of conflict through its effect on either the capabilities of central government to fight insurgency or the benefits for any of the warring actors (such as seizing the capital or controlling major roads). See appendix to detailed description of all the variables.

[23]One may argue that the location of the modern capital city could be an outcome of state history and thus may constitute a case of "bad control". Nonetheless, note that most of the location of modern capital cities in Sub-Saharan Africa followed decisions made by colonizers to service their needs and did not necessarily overlap with the preexisting polities (Herbst, 2000). None of the results in this paper are driven by the inclusion of this vector of geo-strategic controls.

contemporary conflict. Since the standard deviation for the dependent variable (0.232) is very similar to the standard deviation of my state history index (0.227), the interpretation of the coefficient estimates for $\beta$ in terms of standard deviations is straightforward. One standard deviation increase in local state history is associated with 0.2 standard deviation reduction in the prevalence of conflict during the period of analysis (roughly one year in the sample period or one fourth of the mean prevalence of conflict).

I now consider the potential effects of land endowment and the disease environment. Early state development has been influenced by the geographic, climatic, demographic and disease environment (Diamond 1997, Reid 2012, and Alsan 2013). I first include, in column 3, a measure of soil suitability to grow cereal crops which not only positively correlates with early statehood but also it correlates with modern population density, an important driver of conflict.[24] Then, in column 4, I introduce two measures accounting for the ecology of malaria and the suitability for the tsetse fly. The former weakly correlate with my index of state history whereas the latter is strongly negatively correlated with it. In addition, Cervellati, Sunde, and Valmori (2012) find that persistent exposure to diseases affects the likelihood of conflict by affecting the opportunity cost of engaging in violence. In column 5 I include together the two set of controls. The point estimate for $\beta$ remains unaltered.

[insert Table 2 here]

*Potential confounding effects of genetic and ecological diversity.* In Table 3 I explore whether the main correlation of interest documented so far may partially account for the effect of genetic diversity on conflict. Ashraf and Galor (2013a, 2013b) argue that genetic diversity had a long-lasting effect on the pattern of economic development and ethnolinguistic heterogeneity (including fractionalization among other measures). Even more importantly, Arbatli, Ashraf, and Galor (2013) show that genetic diversity strongly correlates with several measures of social conflict. Unfortunately, no data on genetic diversity at the grid cell level exists. To tackle this problem, I use the fact that migratory distance from the location of human origin (i.e: Addis Adaba in Ethiopia) is a strong linear predictor of the degree of genetic diversity in the populations (Ramachandran et al. 2005, Liu et al. 2006, and Ashraf and Galor 2013a).[25] Results in column 1 shows that migratory distance to Addis

[24]Data on soil suitability for growing cereal comes from the Food and Agriculture Organization (FAO)'s Global Agro-Ecological Zones (GAEZ) database. The suitability of the soil is calculated based on the physical environment (soil moisture conditions, radiation, and temperature) relevant for each crop under rain-fed conditions and low use of inputs. The suitability measure ranges between 0 (not suitable) to 1 (very suitable).

[25]Migratory distance from each grid cell's centroid to Addis Adaba is constructed based on Özak (2012a,

Adaba enters with the expected sign suggesting that genetic diversity has a positive impact on conflict.[26] Nevertheless, the point estimate for $\beta$ is affected remarkably little (albeit it slightly decreases in size).

Fenske (2012) shows that ecological diversity is strongly related to the presence of pre-colonial states in Sub-Saharan Africa. Diversity in ecology correlates with potential drivers of conflict such as linguistic or cultural diversity (Michalopoulos 2012, and Moore et al, 2002) and population density (Fenske 2012, Osafo-Kwaako and Robinson 2013). In addition, herders cope with climate limitations by moving between ecological zones which potentially leads to land-related conflicts with farmers (a well-documented phenomenon in conflict literature, in particular for the Sahel region -see Benjaminsen et al 2012). To account for this potential bias, I follow Fenske (2012) and measure ecological diversity as a Herfindahl index constructed from the shares of each grid's area that is occupied by each ecological type on White's (1983) vegetation map of Africa.[27] Point estimates in column 2 of Table 3 show that ecological diversity presents indeed a statistically significant and positive correlation with contemporary conflict. The negative association between local state history and conflict remains statistically strong. Further, I obtain a similar point estimate when controlling for both ecological and genetic diversity in column 3. Figure 4 depicts the scatter plot and partial regression line for the statistical relationship between contemporary conflict and state history from the last specification in Table 3 (labels corresponds to the country ISO codes).

[Insert Table 3 here]

### Robustness Checks

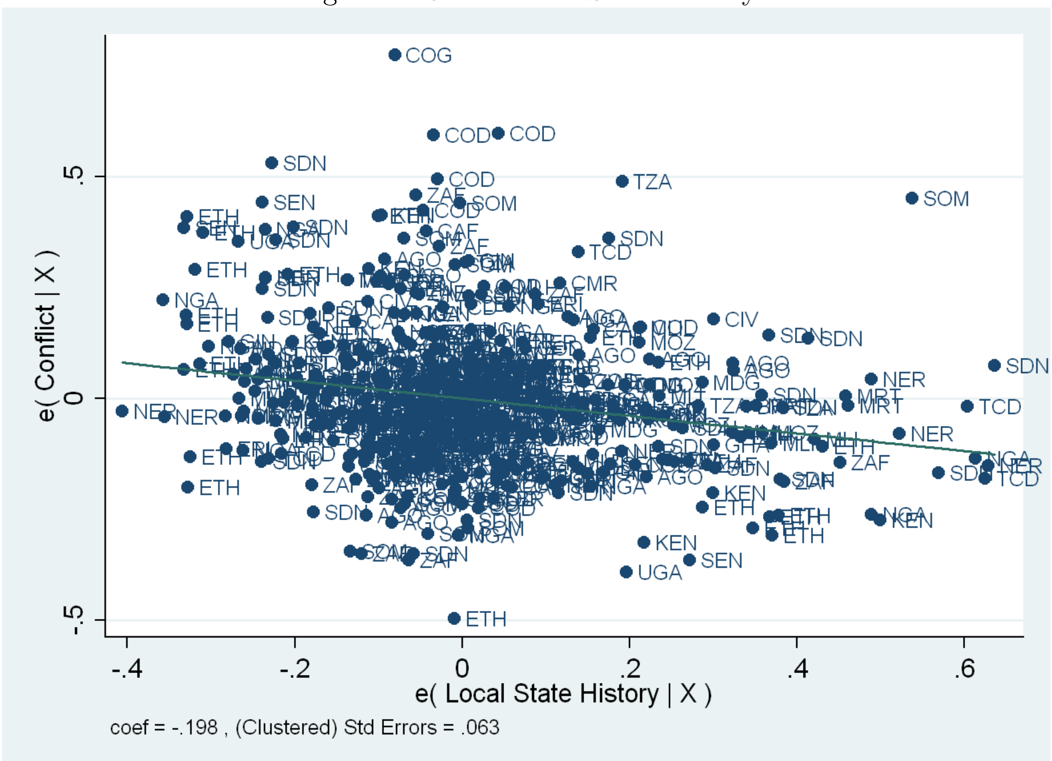
*Considering potential "bad controls" and potential mediating channels.* There are certainly others contemporaneous and historical confounding factors for my analysis. I next show how the point estimate for my variable of interest is affected by the inclusion of additional

2012b), who calculated the walking time cost (in weeks) of crossing every square kilometer on land. The algorithm implemented takes into account topographic, climatic, and terrain conditions, as well as human biological abilities (Özak 2012a).

[26] Controlling for distance and its square (to account for the fact that genetic diversity has been shown to have a hump-shaped relationship with economics development) does not affect the results.

[27] They are 18 major ecological types in White's (1983) map: altimontaine, anthropic, azonal, bushland and thicket, bushland and thicket mosaic, cape shrubland, desert, edaphic grassland mosaic, forest, forest transition and mosaic, grassland, grassy shrubland, secondary wooded grassland, semi-desert, transitional scrubland, water, woodland, woodland mosaics, and transitions. See appendix.

Figure 4: Conflict and State History




controls which can be arguably considered outcomes of a long-run exposure to centralized polities. While not conclusive, changes in my main point estimate when including these controls may be suggestive of the existence of mediating channels through which state history impacts modern conflict. I focus on pre-colonial economic prosperity, population density, ethnic fractionalization, slave trade prevalence, proximity to historical trade routes and historical conflict sites, and contemporary development (proxied by light density at nights obtained from satellite images). I start with pre-colonial ethnic controls accounting for historical levels of prosperity and economic sophistication.[28] I focus on two sets of ethnicity level variables. First, I consider the subsistence income shares derived from hunting, fishing, animal husbandry, and agriculture (variables v2 to v5 from Ethnographic Atlas).[29] Second, I consider a variable describing the pattern of settlement. This variable (v30 from Ethno-

[28]I construct pre-colonial ethnographic measures at the grid cell level based on information from the Ethnographic Atlas (Murdock, 1967) and the spatial distribution of ethnic groups from Murdock's (1959) map. All these measures are 1960 population-weighted averages of traits of ethnic groups whose historical homelands intersect a given grid cell. I basically follow the procedure described in Alesina, Giuliano, and Nunn (2013). See appendix for details.

[29] I omit the category share of income from gathering activities to avoid multicollinearity.

graphic Atlas) is coded in order of increasing settlement sophistication taking values from 1 (nomadic) to 8 (complex settlement). Overall, my point estimate for $\beta$ does not change (albeit its precision is improved) with the addition of these controls in column 1.

Next I analyze the confounding effect of population density.[30] Unfortunately no detailed historical data on population density exists at my level of analysis to the best of knowledge. I only observe population density in 1960 instead and, to the extend that population density may have a persistent effect over time, I use it to proxy for within-country variation of population density in pre-colonial times.[31] Further, using population figures from 1960 alleviates concerns of reverse causality from contemporary conflict to population distributions. The point estimates for $\beta$ increases almost 10% and remains strongly statistically significance. I next construct an ethnic fractionalization variable based on the index introduced in Alesina et al (2003).[32] For similar aforementioned reasons I compute a fractionalization index based on grid population in 1960.[33] The point estimates for $\beta$ remains unaltered when including ethnic fractionalization as control.[34] I next consider slave trade.[35] I construct population-weighted averages of slave trade prevalence at the grid cell level using

[30]Population density is positively correlated with the prevalence of conflict (see, among others, Buhaug and Rød, 2006; Raleigh and Hegre, 2009, and Sundberg and Melander, 2013). It has been argued that low population density was one of the main obstacles for state formation in the pre-colonial Sub-Saharan Africa (see, among others, Bates 1983, Diamond 1997, and Herbst 2000). This hypothesis is, however, contested in a recent work by Philip Osafo-Kwaako and James Robinson (2013). On the other hand, high population density in the past may have also negatively affected ethnic diversity by reducing isolation (Ahlerup and Olsson, 2012).

[31]The use of this proxy can help to illustrate the importance of the bias when including a bad control. Consider for simplicity that conflict ($C$) is only related to state history ($S$) and historical population density ($P$), then the true model I would like to estimate is: $C_i = \beta_0+\beta_1 S_i+\beta_2 P_i+u_i$ . However, I only have data on population density in 1960 ($P^{1960}$ ) which is a function of both $S$ and $P$ : $P^{1960} = \gamma_0+\gamma_1 S_i+\gamma_2 P_i+\epsilon_i$ . When regressing $C$ on $S$ and $P_{1960}$, I am estimating $C_i = \left[\beta_0 - \beta_2 \frac{\gamma_0}{\gamma_1}\right] + \left[\beta_1 - \beta_2 \frac{\gamma_2}{\gamma_1}\right] S_i + \frac{\beta_2}{\gamma_1} P_i^{1960} + \left(u_i - \beta_2 \frac{\epsilon_i}{\gamma_1}\right)$. Since it is apparent that $\beta_2 > 0, \gamma_2 > 0$, and $\gamma_1 > 0$, the inclusion of population density in 1960 would overestimate the negative impact of state history on conflict.

[32]Ethnic fractionalization denotes the probability that two individuals randomly selected from a grid cell will be from different ethnic groups. In order to be consistent throughout this paper my definition of ethnic group is based on Murdock (1959). Therefore, I construct shares of ethnic population using gridded population and the spatial distribution of ethnic groups in Murdock's map. See appendix for details.

[33]Ethnic heterogeneity is a commonly stressed determinant of conflict (see, among others, Easterly and Levine 1997 and Collier, 1998) and it is likely to be correlated with state history ( see Bockstette et al, 2002; and Ahlerup and Olsson, 2012).

[34]I obtain almost identical results (not shown) if I use ethnolinguistic fractionalization (i.e: using ethnologue to compute linguistic distances between pair of ehtnic groups within a grid) instead of ehtnic fractionalization.

[35]Why would slave trade be important for contemporary conflict? First, Nunn (2008) finds that slave trade resulted in long-run underdevelopment within Africa. More importantly, historical slave trade has been shown to have an effect on ethnic fragmentation (Whatley and Gillezeau, 2011b) and individual's mistrust (Nunn and Wantchekon, 2011), which are both arguably potential drivers of social conflict.

Nathan Nunn's data. The expected correlation between slave trade prevalence and state history is ex ante ambiguous.[36] Results in column 4 show that the introduction of slave trade prevalence as a determinant of contemporary conflict does not affect the estimation of $\beta$. The inclusion of shortest distance to historical trade routes in column 5 does not affect the results. I next add the distance to the closest historical battle during the period 1400-1700 CE. This variable is constructed upon information recorded and georeferenced by Besley and Reynal-Querol (2012) who find a robust correlation between proximity to the location of historical battles and contemporary conflict.[37] Results in column 6 are in line with Besley and Reynal-Querol's (2012) main finding. As expected, the point estimate of my variable of interest slightly increases and remains statistically significant. One-standard deviation increase in state history is statistically related to a reduction of the prevalence of conflict of 1/4 of its standard deviation. Neither the inclusion of (ln of) light density, as measured in Michalopoulos and Papaioannou (2013), or the inclusion of the previous variables all together affect the statistical significance of my main finding. Therefore, if anything, the inclusion of these potential confounders makes the negative statistical association between state history and contemporary conflict stronger.

[Insert Table 4 here]

*Robustness to the choice of conflict measure (dataset, incidence, onset, and intensity).* I next show that the main results are robust to the election of the georeferenced conflict dataset and the way conflict is coded. For column 1 of Table 5, I construct the conflict measure using ACLED. Therefore, the dependent variable accounts for the fraction of years with at least one broadly defined conflict event in ACLED. In column 2 I only consider battle events recorded in ACLED. For column 3 I consider any violent event (i.e: battles and violence

[36] On the one hand, Nunn (2008) suggests that slave trade could have been an impediment for pre-colonial state development in Africa. In the same direction, Whatley and Gillezeau (2011a) argues that increasing international demand for slaves might have reduced the incentive to state creation (relative to slave raiding) by driving the marginal value of people as slaves above their marginal value as tax payers. On the other hand, there exist several historical accounts linking the rise of some African kingdoms to the slave trade (see, for example, Law 1977 for the case of the Oyo Empire, and Reid 2012). For instance, while analyzing the role of warfare, slavery and slave-taking in Yoruba state-building, Ejiogu (2011) documents slave-taking campaigns of Oyo against neighboring Nupe (note that Oyo -part of Yoruba - and Nupe share territories within grid cells).

[37] They also show that proximity to historical conflict site correlates with mistrust, stronger ethnic identity, and weaker sense of national identity. Provided this documented long-lasting effect and considering that violent conflict between and within historical African states was part of the state-building processes in the past (see, among others, Lewis, 1966; Ben-Amos Girshick and Thornton, 2001; Ejiogu 2011; Reid 2012 and Bates 2013), the omission of this control would underestimate the effect of state history on contemporary conflict.

against civilians). In column 4 I focus on riots. For all the conflict indicators but riots I find the same pattern: a strong negative statistical relationship between conflict and local state history (for the case of conflict the p-value for $\beta$ is slighty above 0.1).

In column 5 I focus on a measure of conflict intensity. The dependent variable is the (log of) number of casualties due to conflict (best estimate in UCDP-GED). The point estimate for $\beta$ reaffirms the hypothesized negative effect of state history on conflict. My conflict measure under the baseline specification represents the prevalence of conflict violence. It does not make distinction between onset and incidence of violence. That is, this measure does not distinguish a violent event that represents the onset of a new conflict within a dyad from an event that is the continuation of previous confrontations. In column 6 I consider a measure of prevalence of conflict onset (i.e: first confrontation within a dyad). I identify all conflict onsets in the period of analysis and code 1 a grid cell - year observation if at least one onset occurs. As a result, my conflict measure in column 6 represents the fraction of years with at least one conflict onset in the grid cell. Only 149 grid cells experienced at least one conflict onset (out of 417 different conflicts in the period 1989-2010). The point estimate suggests that the onset of conflict is strongly and negatively related to long history of statehood.

[Insert Table 5 here]

*Heterogeneity across regions.* I next explore whether the uncovered relationship between state history and contemporary conflict holds within Sub-Saharan Africa regions. I estimate the specification in column 3 of Table 3 for two regions, namely West-Central Africa and East-South Africa.[38] Results in column 1 and 2 of Table 6 are in line with previous results. In column 3 I exclude Western Africa to show that this particular region is not driving the main results.

[Insert Table 6 here]

*Excluding countries and influential observations.* I sequentially estimate the specification in column 3 of Table 3 by excluding one country at a time. Figure 5 depicts thirty seven different point estimates with their associated t-statistics; the excluded country is labeled in the x-axis. All the coefficients fall in the interval [-0.15, -0.22]. All coefficients but the one for the specification excluding Sudan are statistically significant at the 1 percent level (for the

[38] I follow UN to classify each country in one of the 5 UN regions (North, South, East, West, and Central). Only one country belongs to North (Sudan) and it is assigned to East. I group the original regions in only 2 regions to balance the number of observations in each sample.

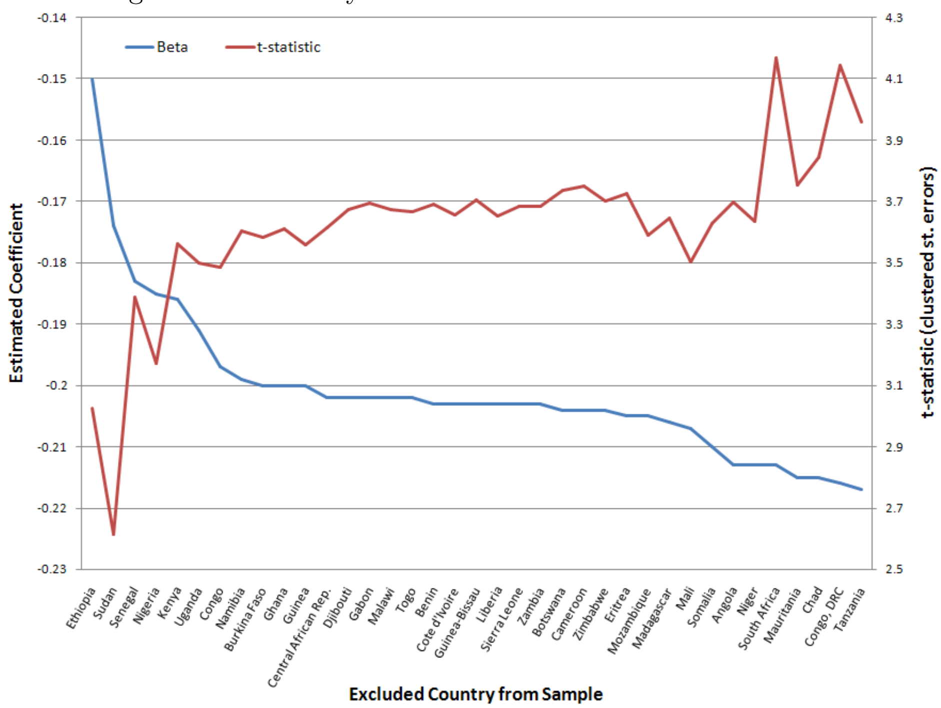
case of Sudan, the p-value is 0.013). There are two reasons for the somehow relative weaker result when excluding Sudan: (1) sample size drops by 10 percent increasing the standard errors (The statistical significances and standard errors of other covariates are also affected -results not shown-), and (2) Sudan presents some locations with very high values of the state history index; locations for which my state history may be underestimating their true long-run exposure to statehood.[39] Excluding those locations reduce the upward bias in the OLS estimate due to the measurement error from bounding the period of analysis above the year 1000 CE.[40] The same pattern arises when excluding another country with long history of states before 1000 CE (i.e: Ethiopia). Finally, the strong negative statistical association between state history and conflict persists when excluding influential observations. In this vein, I follow the standard practice of estimating $\beta$ when excluding all the observations for which $\mid DFBETA_i \mid > 2/\sqrt{N}$, where $N$ is the number of observations and $DFBETA_i$ is the difference between the estimate of $\beta$ when the observation $i$ is excluded and included (scaled by standard error calculated when this observation is excluded). The point estimate is -0.16 (statistically significant at the 1 percent level -results not shown-).

*On the discount factor and long-run exposure.* I next explore how my OLS estimates are affected by the election of different discount factors to compute the state history index. I report in columns 1 to 4 of Table 7 results for four different specifications with discount rates of 5, 10, 25, and 50 percent. For the sake of comparison, I report both the point estimates and the beta standardized coefficients. All the specifications include the full set of controls as in Table 3. Only when a discount rate of 50 percent is applied, my coefficient of interest is slightly statistically insignificant under the conventional levels of confidence. Two facts are worths to note. First, the higher the discount rate, the lower the statistical significance of the coefficient for the corresponding state history measure. Second, the beta standardized coefficient is also decreasing on the discount rate suggesting indeed that history has an influence on conflict. For instance, the beta standardized coefficient when the discount rate is 0 (i.e., -0.20, not show in Table 7) is more than 50 percent larger that for the case in which the discount rate is 25 percent. For columns 5, 6, and 7 I break my period of analysis in two sub-periods, namely before and after 1500 CE when external influence became more relevant for Africa due to the prevalence of slave trade and early european colonialism. In

[39]The Nubian Kingdoms (northern Sudan) were founded several centuries before 1000 CE.

[40]I estimated additional specifications in which I excluded all the observations with some exposure to states during the period 1000-1100 CE. Consistently with the existence of measurement error from bounding the analysis above 1000 CE introducing an upward bias, the beta standardized coefficient slightly decreased about 10 percent (albeit they remained strongly statistically significant -results not shown-) when I excluded those observations.

Figure 5: Sensitivity of Estimates to Exclusion of Countries




columns 5 I only consider the accumulation of state exposure from 1500 CE to 1850 CE. Albeit statistically and economically weaker, there is still a negative statistical association between state history and modern conflict. In columns 6 I only consider the period 1000 -1500 CE and the coefficient of interest is strongly significant and of the similar magnitude when compared with the estimation from the specification using my original measure of state history. When including both measures only the one considering the accumulation from 1000 CE to 15000 CE is strongly statistically and economically significant. This result suggests that the state history that matters the most is the one accumulated before 1500 CE.

[Insert Table 7 here]

*Intensive versus extensive margin.* To argue that what matter the most is the intensive margin of exposure to state institutions (long history) rather than the extensive margin (any state vs. no state at all right before the Scramble for Africa), I estimate a new specification in column 1 of Table 8 for which the state history variable is the state history score the last period considered in the computation of my index (i.e., 1800 - 1850 CE).

The coefficient estimate, albeit negative, is statistically insignificant. Further, I construct a 1960 population-weighted average of the degree of ethnic centralization in the grid cell using the Ethnographic Atlas's variable "Juridisctional Hierarchy beyond the Local Community" which ranges from 1 (no juridisction above village level) to 4 (large state). This variable has been used to document the importance of political centralization for current pattern of development (Gennaioli and Rainer 2007a, and Michalopoulos and Papaioannou 2013). Result in column 2 shows that the correlation between late pre-colonial ethnic centralization and the prevalence of modern conflict is not statistically significant. This result is quantitatively very similar to the point estimates in column 1. One can still argue that it is not the long history of state but its complete absence what explains the uncovered statistical association. In this sense, it may be the case that locations with no history of state whatsoever are located in remote and unpopulated regions with little national state penetration where rebel groups can easily operate. In the specification of column 3 I exclude all the observations with no history of state whatsoever (223 grid cells) and show that my main results are not driven by those locations. The point estimate is very similar and strongly statistically significant. If I restrict the sample even more and consider only locations with at least 100 years of state history (Thus, I exclude 329 grid cells) I obtain even stronger results (column 4).

[Insert Table 8 here]

*Assessing the extent of bias from unobservables.* The point estimates reported so far may still be biased due to unobservable factors correlated with both contemporaneous conflict and long-run exposure to states. How large would this selection on unobservables need to be (relative to selection on observables) to attribute the entire OLS estimates previously reported to a unobservable selection effect? I follow the intuitive heuristic in Nunn and Wantchekon (2011) based on Altonji, Elder, and Taber (2005) to assess the degree of omitted variables bias by studying stability of the estimates for $\beta$. The underlying idea is that, under the assumption that selection on observables is proportional to selection on unobservables, a coefficient not changing much as one adds controls would be suggesting that there is little remaining bias. I thus compare the point estimate in the last specification in Table 3 which includes a full set of controls ($\hat{\beta}_1 = -.198$) with the point estimate when only a basic set of controls (i.e., country fixed effect and geographical controls) is included ($\hat{\beta}_2 = -.191$). The ratio between $\hat{\beta}_1$ and $\hat{\beta}_1 - \hat{\beta}_2$ (the selection on observables) suggests that selection on unobservables would have to be more than 20 times the selection on observables to explain away the entire statistically relationship between state history and contemporaneous conflict.

### Instrumental Variable Approach

I have already documented a strong negative statistical relationship between history of statehood and contemporary conflict at the local level. This historical link is robust to a battery of within modern countries controls ranging from contemporaneous conflict correlates, geographic factors, to historical and deeply rooted determinants of social conflict. Unfortunately, history is not a random process. Even in a close to ideal and hypothetical quasi-random historical event determining the geographic assignment of long-run exposure to state capacities within Sub-Saharan Africa, the challenge of isolating the causal effect of state history on contemporary conflict is particularly difficult. Although one could argue that reverse causality from conflict to historical exposure to state-like institutions is not a source of concern for the identification of my parameter of interest, there may still be omitted variables correlated with both state history and contemporary conflict which may be driving the uncovered statistical association. I aim, however, to convince the reader here that hard-to-account-for factors manifested in differences in long-run exposure to centralized institutions matter crucially to understand contemporary patterns of conflict within Sub-Saharan Africa. In addition to a potential omitted variable bias, measurement error may also affect my point estimates in an ex-ante ambiguous direction depending on the true structure of the relevant measurement error. It is precisely for the sake of alleviating the potential bias from measurement error in my state history index what primarily motivates the introduction of an instrumental variable strategy. I follow two strategies. First, I exploit variation across locations on the timing elapsed since the neolithic revolution. Second, I construct an alternative measure to account for the degree of influence from politically centralized states by exploiting variation in the proximity to historical cities for the time period 1000 - 1800 CE.

### Time Elapsed since Neolithic Revolution

Several studies have established a link between the neolithic revolution and the rise of complex political organization. For instance, Diamond (1997) argues that the transition from hunter-gathered societies to settled agricultural communities is an essential factor to explain the rise of proto-states, and subsequently the formation of states. Agriculture allowed nomadic societies to settle, generate food surpluses, and shorten birth intervals (Ashraf and Michalopoulos 2010; Diamond 1997), which in turn resulted in denser populations. Further, the storage of those food surpluses allowed the emergence of non-food-producing sectors, hence economic specialization and also social stratification manifested not only in the existence of a labor force involved in the production process (both, foods and non-foods production) but also a labor force designated to provide public services (including armies). By implication, the possibility of taxation and the emergence of political organization, facilitates the rise of states.

A strong positive correlation between the timing elapsed since the neolithic revolution and Bockstette, Chanda, and Putterman (2002)'s state antiquity index has been empirically documented at the country level (see, for instance, Hariri 2012, and Petersen and Skaaning, 2010). Further, substantial variation in the timing of the transition to agriculture exists within the Sub-Saharan Africa. I therefore exploit Neolithic archeological sites information to construct the time elapsed since neolithic revolution at a finer geographical level and use this to instrument for my state antiquity index at the local level. I discuss the construction of the instrument in the appendix. Figure 6 depicts the geographical distribution of the time elapsed since the neolithic revolution.

Figure 6: Time Elapsed Since Neolithic Revolution

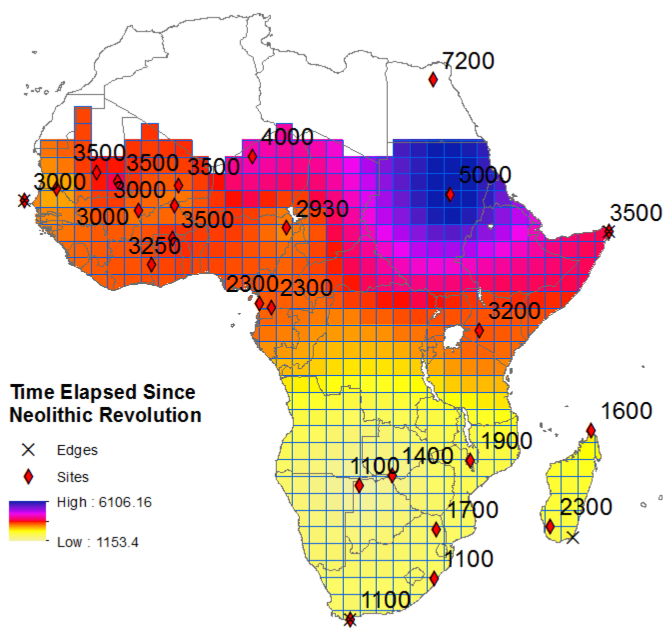



Table 9 presents point estimates for different specifications of the first-stage. All the specifications include the set of controls from my prefered specification in Table 3 (column 3). I report 3 different standard errors (i.e: clustered at the country level and adjusted for spatial autocorrelation for cut-off distances of 5 and 10 degrees). Again, clustering at the country level appears to be the most conservative approach. Point estimate in column 1 suggests a statistically significant association between the time elapsed since the neolithic revolution and my state history index. However, this association is only statistically significant at the 3 percent level due to a weaker association within the Central Africa region. Indeed, my constructed measure of the timing since the first adoption of agriculture is statistically insignificant within that region (column 2). In particular, observations from four countries explain this weaker statistical performance: Gabon, Congo, Congo DR, and Angola. It is well stablished that a group of Bantu people migrated south from modern-day Cameroon to modern-day Namibia across the rainforest. Although this Bantu migration route crossed the territory of modern-day Gabon, Congo, Congo DR, and Angola, no archaelogical information has been recorded on agricutural adoption (Putterman, 2006). Therefore the interpolated values for most of the observation from these four countries mostly depend on archaelogical evidence from two sites in Cameroon and the Namibia-Botswana border. When I exclude these countries from my sample, the first-stage is much stronger (column 3). An alternative way to statistically improve the performance of the first-stage is to consider the square of the instrument rather than its level. Indeed, the statistical association between the square of my instrument and my index of state history is stronger (column 4). To avoid reducing my sample size by 20 percent due to the exclusion of these countries for which the linear relationship is weaker, I use the square of the time elapsed since the neolithic revolution as an instrument for state history. I provide below evidence that the results are not driven by this particular specification.

[Insert Table 9 here]

In Table 10 I report IV results. The specification in column 1 includes the set of controls from specification in column 3 of Table 3. The point estimate is roughly 3 times larger that the previously reported OLS estimates; finding that is consistent with the idea that measurement error in my local state history variable was indeed introducing a sizeable bias toward zero (albeit it is also consistent with my instrument picking up the effect of other confounding factors, issue I discuss below). The point estimate for $\beta$ in column 1 suggests that one-standard deviation increase in local state history implies a .71-standard deviation reduction in the prevalence of contemporary conflict. This magnitude is equivalent to roughly 4 years of conflict in my period of analysis. My IV estimate may be still biased due to the omission of variables that could plausibly correlate with both contemporary conflict and the timing of the transition to agriculture. Therefore, I now focus on biogeographical variables which had been shown to correlate with my instrument. In column 2 I add absolute latitude. As Diamond (1997) argues, technologies and institutions have historically spread more easily at similar latitudes where climate and day duration were not drastically different. This is particularly true for the spread of agriculture within Africa. Regardless of the discussion of the ultimate underlying mechanisms, the high correlation between absolute latitude and development has been widely documented in the economic growth and development literature (see, for example, Spoloare and Wacziarg, 2013). Absolute latitude enters negatively and statistically significant in this specification; reducing 15 percent the size of $\beta$ which remains strongly statistically significant.

Haber (2012) argue that variation in biological (and technological) characteristics of crops had a long-run effect on institution and development. In addition, Sub-Saharan centralized military states were historically more prevalent in areas with soils suitable for the generation of agriculture surpluses to maintain armies (Reid 2012). Given their ease of storage and transport, cereals had a natural advantage over other crops, such as tubers and tree crops, to produce those necessary surpluses.[41] For the particular case of Sub-Saharan Africa, the most important indigenous crop for this matter are sorghum and millet. Further, it was precisely the domestication of sorghum and millet what played a crucial role in the transition to the Neolithic in this region. Hence, in column 3 I hold constant biogeographical factors affecting these crops by adding the principal component of the soil suitability to grow millet and sorghum.[42] The point estimate for $\beta$ remain statistically significant and economically sizeable (the size of coefficient varies little with respect to the baseline specification in column 1).

I next consider the potential confounding effect of climate variability. Ashraf and Michalopoulos (2013) show that historical climatic volatility has a non-monotonic effect on the timing of the adoption of agriculture. On the other hand, Durante (2009) show that, within Europe, variation in social trust is driven by historical variation in climate. When I include intertemporal temperature volatility, its square, and historical mean temperature, the size of my point estimates decreases by 30 percent (albeit it remains statistically significant).[43] This fact is consistent with the possibility that my hypothesized mitigation effect of state

[41] The tubers and tree crops have relative low levels of storability (compared with cereals) since they typically perish within days or weeks of harvesting.

[42] Data on soil suitability for growing millet and sorghum comes from the Food and Agriculture Organization (FAO)'s Global Agro-Ecological Zones (GAEZ) database.

[43] I use variation in modern data to proxy historical climatic variation. Ashraf and Michalopoulos (2013) show that spatial variation in temperature volatility remains largely stable over long periods of time; thus contemporary climate data can be meaningfully employed as informative proxies for prehistoric ones.

history on contemporary conflict may partially confound with higher levels of social trust induced by historical climate variability. The addition of the set of pre-colonial ethnic characteristics in column 5 reduces the point estimates by almost 20 percent (albeit it remains strongly statistically significant). This result represents suggestive evidence that pre-colonial differences in economic prosperity (proxied by the degree of the sophistication of settlement patterns and the type of economic activities) may represent a mediating channel through which long history of statehood of a given location may result in lower levels of modern conflict. In column 6 I include all the aforementioned controls together. Even under this challenging horse race the point estimates for $\beta$ remains negative and statistically significant at the 5 percent level. The IV point estimates doubles my prefered specification in the OLS case. The Kleibergen-Paap rk Wald F statistic is above the rule of thumb thus weak instruments would not be a concern. In the last column of Table 10 I exclude the four Central African countries for which the linear relationship between the neolithinc instrument and my state history index was statistically weak and show that instrumenting state history with the time elapsed since the neolithic revolution in levels lead to very similar point estimates compared to the case with the squared instrument in the full sample.

[Insert Table 10 here]

*A note about hypothesized effects of the neolithic revolution on contemporary outcomes.* The onset of the neolithic revolution was certainly one of the most important historical events for humankind. The economic literature has provided mixing evidence regarding the existence of a direct association between this historical event and contemporaneous outcomes. Some works have emphasized the long-lasting effect of the neolithic revolution on pre-industrial era's outcomes (see, among others, Ashraf and Galor 2011). Other works have shown that countries which experienced early transition to agriculture tend to have higher level of per capita income today (particularly after "ancestry adjustment", see Spoloare and Wacziarg, 2013). This line of argument emphasizes that the neolithic revolution may be responsable for cross-sectional differences in human capital and technologies in the pre-industrial era and, to the extent that those differences may be persistent, may still have an effect on current outcomes. In fact, the reduction on the size of my IV point estimates when including pre-colonial prosperity measures is consistent with this view although it does not explain away the statistical association of interest. In addition, the inclusion of light density at nights to proxy for contemporary levels of development does not affect my point estimates (results not shown). Using individual-level data and exploiting variation in my instrument for about 1,600 districts, I do not find evidence that my instrument significantly correlates with education levels (results not shown).

On the other hand, a line of research argues that neolithic still exert a negative effect on contemporary outcomes. In particular, Olsson and Paik (2013) argue that an early transition to agriculture might have directly shaped institutional trajectories by promoting autocracy (similar argument is presented in Hariri, 2012) and thus facilitating extractive institutions which turn in lower levels of development. In addition, Paik (2011) shows that, within Europe, early adoption of agriculture positively correlates with strong preference for obedience. This line of argument is based on the idea that precisely the experience with early political organization is the mediating channel through which the neolithic revolution affects contemporary outcomes. Bearing in mind that measuring culture traits is a difficult task, I later show that the time elapsed since the neolithic revolution can be linked to trust in state institutions.

In sum, although I cannot rule out the possibility that other hard-to-account factors may be driven the uncovered statisticaly association, IV estimates exploiting information on the timing elapsed since the neolithic revolution provides additional evidence consistent with my hypothesized negative effect of long-run exposure to statehood on modern pattern of conflict. This negative statistical association estimated from variation in the timing since the earliest date of domestication of plants is not driven by biogeographical factors such as land quality to grow cereals (cereals in general or sorghum and millet in particular), proximity to water (rivers and oceans), elevation, intertemporal climate variability, differences in disease environment (malaria and tse-tse), ecological and genetic diversity or absolute distance to the equator. I find evidence that pre-colonial differences in prosperity may constitute a mediating channel underlying my reduced-form relationship.

### Proximity to Historical Cities

I construct an independent imperfect measure of state history by exploiting information on the location and evolution of above sixty large African cities (of which thirty five were located in Sub-Saharan Africa) during the period 1000 - 1800 CE.[44] To the extent that kingdoms and empires tended to have a large city as political center, I consider proximity to a large city as an indicator of the degree of influence from a centralized power. I introduce this new measure for several reasons. First, to show that the negative statistical association

[44] I define a city to be large if it has more than ten thousand inhabitants. The list of cities comes from Chandler (1987) and Eggiman (2000).

uncovered in the OLS case still hold when using an alternative measure. Second, this new measure will overcome a potential caveat in my original measure of state history which assumes an homegeneous effect of centralization within the boundaries of a historical polity. This assumption had two implications: (1) the introduction of a sharp discontinuity at the border of the boundary, and (2) inconsistency with the idea that broadcasting power strenght may depend on the distance from the political center. Third, this new measure can be used to instrument the original state history measure and mitigates the attenuation bias from measurement error.

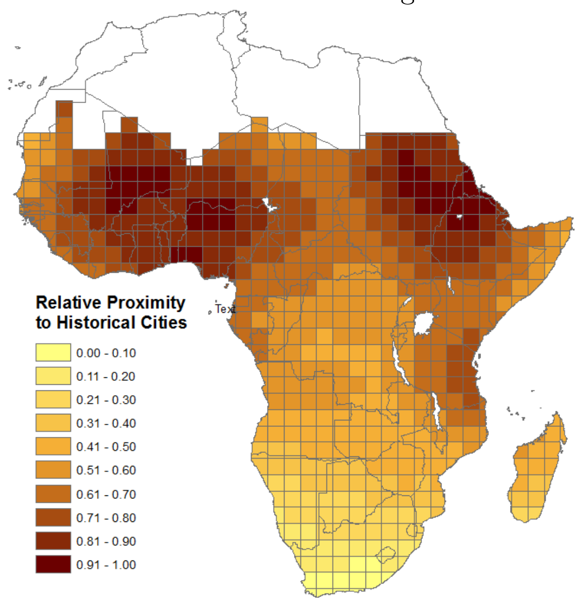
*Construction.* This measure exploits time-varying proximity to large cities. Therefore, some cities exert influence to their periphery only for particular time intervals. For instance, Djenne, in modern Mali, only enters in my panel of cities for the period 1300 - 1600 CE (See Table A.3 in appendix for the list of georeferenced cities). For each hundred years period I calculate the shortest distance to closest city from the centroid of each grid cell. I then calculate within-grid average of the distances for the whole period of analysis and map them into a 0 to 1 interval so the grid cell with the minimum average distance takes the value 1. Figure 7 displays the cross-sectional variation of this alternative measure.[45] As in the case of my original State History Index, this new measure presents higher values in western part of the Sahel, the highlands of Ethiopia, and the region along the Nile river.

In column 1 of Table 11 I present the OLS estimate for the reduced-form conflict and historical proximity to cities. I find the same the pattern as before. Historical proximity to cities for the time period 1000 - 1800 CE is negatively and strongly statistically associated to prevalence of modern conflict. In column 2 I present the IV estimate for my state history index when using historical proximity to cities as a instrument. I find results which are similar to the IV case using the time elapsed since the neolithic as an instrument. The point estimate is slighlty larger; fact that is consistent with the possibility that historical proximity to cities is picking up the effect of other omitted variables on conflict. Finally, in column 3 I include both instruments and find similar results. With this overidentified model at hands I test the exogeneity of the instruments. The Hansen J statistics suggests no rejection of the null hypothesis that my set of instruments are exogenous. Note that all the specifications where including not only the set of controls from Table 3 but also account for the joint effects of absolute latitude, suitability to grow sorghum and millet, intertemporal climate variability

[Insert Table 11 here]

[45] By construction, due to its dependance on distances, this measure presents a more continuous support.

Figure 7: Alternative Measure Using Historical Cities




### 3.3 Panel Data Evidence: Weather Induced-Agricultural Productivity Shock, State History, and Conflict

In a comprehensive synthesis of the climate-conflict literature, Burke, Hsiang, and Miguel (2013) argue that there is strong causal evidence linking climatic events to conflict. The existence of an income mechanism underlying this causal link has been proposed repeatedly times in the conflict literature but it has not been definitively identified yet. Harari and Laferrara (2013) present convincing evidence that what drives the observed empirical relationship between weather shocks and conflict in Africa is weather anomalies occurring within the growing season of the main local crops. In addition, Schenkler and Lobell (2010) shows that crop yields are indeed affected by growing season precipitation and temperature. Moreover, Brown et al (2011) argue that persistent drought conditions is the most significant climate influence on GDP per capita growth in Africa. Given the high dependance of Sub-Saharan Africa economies on rainfed agriculture, these results provide strong evidence consistent with the existence of an income mechanism. Therefore, I draw upon Harari and Laferrara (2013) to construct weather-induced agricultural shock by exploiting information on spatial distribution of crops, planting and harvesting calendars, and variability on water balance anomalies across space and time.[46] I hypothesize that locations with long history

[46]I discuss the construction of the weather-induced agricultural shock in the appendix.

of statehood should be better equipped of mechanisms to mitigate the negative effects of weather shocks. To support my hypothesis, I exploit panel data variation (over the time period 1989-2010) in the prevalence of conflict, weather-induced productivity shocks, and the interaction of my state history index with those shocks to estimate the following equation:

$$Conflict_{i,c,t} = \alpha + \beta State\,History_i + \gamma Shock_{i,t} + \delta State\,History_i \times Shock_{i,t}$$

$$+G_i^{'}\Gamma + X_i^{'}\Delta + C_i^{'}Z + W_{i,t}^{'}\Pi + \lambda_c + \mu_i + \nu_t + \epsilon_{i,c,t} \quad (2)$$

Where $t$ indexes year. The variable $Conflict_{i,c,t}$ takes the value 1 if at least one conflict event occurs in the grid cell $i$ in year $t$, and 0 otherwise. The variables $State\,History_i$, $G_i^{'}$, $X_i^{'}$, and $C_i^{'}$ are the same defined for equation (1). The vector $W_{i,t}^{'}$ includes year averages of monthly precipitation and temperature deviation from historical monthly means to account for any independent effect that these variables may have on conflict outside of the growing season. The variable $\nu_t$ denotes a year fixed effect whereas $\mu_i$ is a collection of grid cell fixed effect which is included only in some specifications (when included I cannot identify $\beta$, $\Gamma$,$\Delta$, $Z$, and $\lambda_c$). The main coefficient of interest in this exercise is $\delta$. Standard errors are clustered at the grid cell level.

In column 1 of Table 12 I present OLS estimates an specification of equation (2) for which yearly weather variables and grid fixed effect are not included. The point estimates suggest: a) an statistically significant negative correlation between conflict and state history, b) an statistically significant positive impact of negative weather shocks on conflict, and c) a negative correlation between the interaction term of the two aforementioned variables and conflict, which is consistent with my hypothesized mitigating effect of state history when a location is hit by a shock. In the following columns I present IV estimates. Column 2 shows the results for the same specification as in column 1. One-standard deviation increase in the shock measure statistically relates with 5 percent increase in the likelihood of experiencing at least one conflict event in a grid without history of state (The unconditional probability of having conflict is 0.18). The estimated coefficient for $\delta$ suggests that for the grid with the mean value of state history (i.e: 0.16), the negative impact of a weather shock on conflict is quite smaller than for the case of no previous state history. That is, one-standard deviation increase in the weather shock implies only 1.2 percent increase in the likelihood of conflict. In column 3 I add grid cell fixed effect, thus I cannot identify $\beta$, $E$, $\Gamma$, $\Delta$, and $\Lambda$. The point estimate for the direct effect of weather shock on conflict remains almost unaltered whereas for the case of the interaction term it is slightly smaller, albeit statistically significant and representing economically meaningful mitigation effect of state history. The addition of yearly measures of precipitation and temperature deviation in column 4 does not substantially affect the previous result. In the last specification in column 5, I include conflict prevalence lagged one period to account for conflict dynamics. Indeed, results in column 5 show that conflict in $t-1$ strongly predicts conflict in $t$. Albeit slightly smaller in size, the point estimates for the weather shock variable and its interaction with local state history remain both statistically significant.

[Insert Table 12 here]

*Other interaction effects.* I next consider a set of different cross-sectional characteristics that when interacted with weather shocks may partially account for the result previously documented, namely that locations with relatively long history of state are less prone to experience conflict when hit by a shock. This set of characteristics includes light density at nights (proxy of regional development), soil suitability for cultivating cereals, pre-colonial agricultural dependence, and historical temperature volatility. All the specifications in Table 13 include both year and grid fixed effects. The first column serves as a comparison. The point estimates from columns 2 to 5 suggest that locations with higher light density at night, better cereal suitability, and higher pre-colonial dependence on agriculture are more prone to experience conflict when hit by a shock. On the contrary, locations with higher temperature volatility are less prone to have conflict. The inclusion of these interaction terms separately or jointly (in column 6) does not wash away the statistical significance of the negative coefficient for the interaction term state history and weather shock.

[Insert Table 13 here]

# 4 Identifying Potential Mechanisms at Work: State History and Attitudes Towards State Institutions

It has been stressed that the lack of state legitimacy represents an underlying cause of the prevalence of civil conflict in Sub-Saharan Africa. It is argued that the lack of legitimacy undermines the institutional capacity and authority of a state to rule by consent rather than by coercion. Not surprisingly, Rotberg (2004) argues that the lack of legitimacy causes, in fact, state fragility; a concept that it is partially defined in terms of a society's probability to face major conflicts.[47] States with low levels of legitimacy tend to devote more resources towards retaining power rather than towards effective governance, which undermines even more its popular support and increases the likelihood of political turnover (Gilley, 2006). On the contrary, citizens that consider a government to be legitimate are less likely to rebel. Does state history at the local level shape individual's perception of state legitimacy? I argue that attitudes towards state institutions such as tax department, police force, and court of laws provide an informative way to measure views of legitimacy of the state. I thus show a strongly positive and robust correlation between individual's belief regarding the legitimacy of these key state institutions and local level of state history. I also present additional results suggesting the presence of a legacy of state history on individuals' trust in state institutions. Moreover, I present suggestive empirical evidence that state history can be linked to modern support for traditional leaders.

### Sources and Description of Individual-Level Data

In this section I exploit individual-level survey data to document the aforementioned potential mechanisms at work. My analysis is based on the Round 4 of Afrobarometer in 2008 and 2009 (Afrobarometer 4, from now on). The Afrobarometer 4 is a collection of comparative series of nationally representative surveys for twenty countries in Sub-Saharan Africa: Benin, Botswana, Burkina Faso, Cape Verde, Ghana, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mozambique, Namibia, Nigeria, Senegal, South Africa, Tanzania, Uganda, Zambia, and Zimbabwe. These countries have undergone some degree of political and economic liberalization during the last 20 years (Logan, 2013). In addition, the Afrobarometer 4 sample does not include countries under authoritarian regimes or civil wars (Afrobarometer, 2007). Nonetheless, all the countries in the Afrobarometer 4 sample but Benin, Burkina Faso, Cape Verde, and Malawi experienced violent conflict events during the period 1989-2010.[48] They also present high heterogeneity in key variables such as state history, historical conflict, and other correlates of civil conflict.

The Afrobarometer 4 relies on personal interviews conducted in local languages where the questions are standardized so responses can be compared across countries (Afrobarometer,

[47] For instance, a legitimacy score accounts for almost 50% of the State Fragility Index computed by the Center for Systemic Peace. Moreover, the operational definition of fragility in the index is associated with state capacity to manage conflict.

[48] The list of countries in Afrobarometer 4 with at least one conflict with more than 25 deaths during the period 1989-2010 is (event counts in parenthesis): Botswana (1), Ghana (34), Kenya (307), Lesotho (5), Liberia (510), Madagascar (39), Mali (98), Mozambique (261), Namibia (21), Nigeria (319), Senegal (187), South Africa (2624), Tanzania (9), Uganda (1549), Zambia (10), and Zimbabwe (45).

2007). These questions asses, among other topics, individuals attitudes toward democracy, markets, and civil society. In particular, I exploit information regarding individuals attitude toward state institutions and trust in politicians, public servants, and other individuals in general. I also benefit from a module of questions on local traditional authority. As described in detail below, the information in Afrobarometer 4 also allows me to construct controls at the level of village (i.e: enumeration area) and district.

The original sample size in Afrobarometer 4 is over 26,000 respondents. Cape Verde and Lesotho are not included in my analysis.[49] In addition, districts that I was not able to georeference, as well as individuals who could not be matched with ethnic names in Murdock's (1959) map were removed from the sample.[50] The final sample consists of 22,527 respondents from 1,625 districts and 221 different ethnic groups under Murdock's (1959) classification.[51]

### State History and State Legitimacy

I construct a measure of individual's attitude toward the legitimacy of state institutions based on the individual's level of agreement with the following statements: (1) the courts have the right to make decisions that people always have to abide by; (2) the police always have the right to make people obey the law; and (3) the tax department always has the right to make people pay taxes. The possible responses, coded from 1 to 5, are: "strongly disagree", "disagree", "neither agree nor disagree or don't know", "agree" and "strongly agree". Thus, my measure of state legitimacy is the first principal component of the responses to these questions about the legitimacy of the court decisions, police enforcement, and the tax department. I examine the statistical relationship between state legitimacy and state history by estimating different specifications of the following equation:

$$State\,Legitimacy_{i,e,a,d,c} = \alpha + \beta State\,History_{d,c} + I^{'}_{i,e,a,d,c}\Gamma + A^{'}_{a,d,c}\Delta + D^{'}_{d,c}E$$

$$+X^{'}_{d,c}H + \eta_c + \theta_e + \epsilon_{i,e,a,d,c} \tag{3}$$

[49] I exclude Lesotho and Cape Verde from my analysis for several reasons. I exclude Cape Verde because it was not taken into account in the original computation of my state antiquity index, no question on traditional leaders were asked during the round 4 of Afrobarometer, and difficulties to match the ethnicities of the respondents with Murdock's data. I exclude Lesotho due to difficulties to match ethnicities.

[50] My georeferencing work was built upon a previous work by Stelios Michalopoulos.

[51] 320 ethnicities are originally self-reported in my sample. Appendix includes the list of ethnicities and their match with names in Murdock's (1959) map

where $i$, $e$, $a$ , $d$, and $c$ index individuals, ethnicity, enumeration area (village), district, and country, respectively. The variable $State\,History_{d,c}$ represents the state antiquity index for a buffer with a 1-degree radius (approximately 100 kilometers) and centroid located at the coordinates of the district. The vector $I^{'}_{i,e,a,d,c}$ denotes a set of the respondent's characteristics such as age, age squared, education level, living conditions, unemployment status, and gender.[52]

The vector $A^{'}_{a,d,c}$ denotes a set of enumeration area-level covariates including a rural dummy and a subset of variables designed to capture the prevalence of public good provision and proxy for the quality of local government.[53] In a study for South Africa, Carter (2011) shows that individual who are more satisfied with the quality of public good provision tend to see the state as legitimate. In addition, Gennaioli and Rainer (2007b) argue that history of state centralization had an impact on the quality of local government public provision.[54] This hypothesis is contested by Bandyopadhyay and Green (2012) who find no correlation between pre-colonial centralization and local public good provision in Uganda.[55] Nonetheless, it worths to note that the introduction of the public good provision dummies has little impact on the estimation of the main coefficient of interest.

The vector $D^{'}_{d,c}$ is a set of district-level variables accounting for differences in development, which includes distance to the capital city, infant mortality, and per capita light density (in logs).[56] The $X^{'}_{d,c}$ denotes a vectors of district-level covariates, respectively; which are included in different specifications of equation (3) and are discussed below. Finally, $\eta_c$ and $\theta_e$ are country and ethnicity fixed effect. Since the main variable of interest, i.e: $State\,History_{d,c}$, varies at the district level, I adjust the standard errors for potential clustering at the district level.[57]

[52] The education variable takes value from 0 (no formal schooling) to 10 (post-graduate). The living condition variable is a self assessment of the respondent and takes values from 1 (very bad) to 5 (very good). Unemployment and gender are dummies variables taking value 1 if the respondent is unemployed and male, respectively. The Afrobarometer 4 does not include information on occupation of the respondent.

[53] I introduce 6 dummies indicating the presence of police, school, electricity, piped water, sewage system, and health clinic. Note that an enumeration area or village is the lowest order administrative unit available in Afrobarometer 4.

[54] Although robust to different specifications, the evidence in Gennaioli and Rainer (2007b) is arguably far from being conclusive due to pitfalls of aggregation of ethnographic data at the country level (and the number of countries being small).

[55] Although testing Gennaioli and Rainer's (2007b) main hypothesis at the local level represents an improvement from the original work, it is unclear that results in Bandyopadhyay and Green (2012) can be regarded as representatives of the whole Sub-Saharan Africa.

[56] Bandyopadhyay and Green (2012) also show that ethnic pre-colonial centralization positively correlates with level of development at the subnational and individual levels in Uganda.

[57] See appendix for all the details regarding definitions of the variables included in my analysis.

*Basic OLS and IV results.* The set of controls $I'_{i,e,a,d,c}$, $A'_{a,d,c}$, and $D'_{d,c}$ is included in all the specifications.To capture those ethnic-specific factors that may both affect the state legitimacy and also correlate with my state antiquity index at the district level, I include ethnic fixed effects. It is worths to mention that I am able to identify $\beta$, even after the introduction of ethnic fixed effect, because almost half of the individuals in my sample are not currently living in the historical homeland of their ancestors. Thus, it is also important to emphasize that the estimated coefficient for $\beta$ would be be representing the average statistical relationship between state history of the district and attitude toward legitimacy of state institutions for those individuals living outside the historical homeland of their ethnic groups. The OLS result suggests that there is no statistical relationship between state history at the district level and individual's opinion about the legitimacy of the state. On the contrary, the IV estimate suggests an statistically strong and positive correlation between these two variables. Therefore, the history of the place where people live, outside of the tradition of the people living in that place, is relevant to shape people's beliefs regarding the legitimacy of the state institutions. The positive coefficient of interest is not only strongly statistically significant, but it is also economically meaningful: one-standard deviation increase in state history (i.e.; 0.26) is associated with more than one-standard deviation increase (i.e; 1.2) in the state legitimacy index. Note that neither Gabon, Congo, Congo DR, nor Angola are included in the Afrobarometer 4. As a result, the first-stage specification using the years elapsed since the neolithic revolution in levels (instead of its square) produces a similar fit. I thus present next the results using the instrument in levels.

I next consider additional district-level controls in a validity exercise which is similar to the one implemented above for the study of the relationship state history and conflict. The introduction of migratory distance, ecological diversity, soil suitability for millet and sorghum, absolute latitude and the set of variables accounting for intertemporal temperature volatility (intertemporal temperature volatility, its square, and historical mean temperature) in column 3 of Table 14 reduces the size of $\beta$ by 15 percent (albeit it remains statistically significant at the 5 percent level).[58]

*Further district-level controls.* I next consider a subset of district-level controls included in $X'_{d,c}$. I consider the potential confounding effect of prevalence of historical conflict,[59]

[58]Adding these controls separateley lead to similar results.

[59]Using Afrobarometer data, Besley and Reynal-Querol (2012) show that the prevalence of historical conflict at the country level is correlated with less trust, stronger sense of ethnic identity and a weaker sense of national identity.

ethnic fractionzalition,[60] and slave trade.[61] I measure prevalence of historical conflict with an indicator variable taking the value 1 if at least one historical battle in the period 1400-1700 took place less than 100 kilometers away from the centroid of the district. I construct a fractionalization measure at the district level using information of the ethnicities of the respondents. I construct two measures of slave trade prevalence at the district level. First, I follow Nunn and Wantchekon (2011) and calculate the historical slave trade exposure for the ethnic group that historically inhabited the location (district) where the respondent currently lives. Second, I construct the weighted average slave trade prevalence of the district based on the slave trade exposure of all the ethnic groups reported in the survey for that district. The addition of these controls slightly increase the size of point estimate for $\beta$.[62]

Note that the Kleibergen-Paap rk Wald F statistics are slightly below the "rule of thumb" generally applied to identify a weak instruments problem.[63] Nonetheless, under a weak instrument problem my IV estimate would be biased toward OLS which is close to zero. Further, I also report point estimates based on Fuller 1 estimator; a biased corrected limited information maximum likelihood estimator. It has been argued that this type of k-class estimator have a better finite-sample performance than 2SLS when instruments are potentially weak (Baum, Schaffer, and Stillman 2007, and Stock, Wright, and Yogo 2002). Point estimates from Fuller 1 estimator are remarkably similar and statistically significant.

[Insert Table 14 here]

*Internal vs External Cultural Norms.* I attempt to distinguish whether it is the state history of the place where people live versus the state history of the ancestors of the people living

[60] For the particular case of rural western Kenya, Miguel and Gugerty (2005) show that ethnic diversity is associated with lower provision of public good at the local level. They argue that this collective action failure follows from the inability to impose social sanctions in highly diverse communities. Although I already control for the degree of public provisions, this inability could also be related to low levels of trust and trustworthiness. In fact, Barr (2003) argues that low levels of trust is related to ethnic heterogeneity (in Zimbabwe).

[61] Nunn and Wantchekon (2011) show that individuals from ethnic groups that were strongly affected by the slave trade in the past are less trusting today. In particular, those individuals trust less on the local councils. I argued above that the relationship between the history of state formation and slave trade prevalence is ambiguous. Nonetheless, if any relationship exists (regardless of its direction), omitting the impact of slave trade would introduce a bias in the estimation of my coefficient of interest. In line with Nunn and Watchekon's (2011) hypothesis, the prevalence of slave trade at the ethnicity level has indeed a negative impact (and strongly statistically significant) on the legitimacy of the state.

[62] Adding these controls separately lead to similar results.

[63] Note that I do not use the Stock and Yogo's critical values to evaluate the strength of the instrument. Baum, Schaffer, and Stillman (2007) suggest to apply caution on using Stock and Yogo's critical values (which were compiled for an i.i.d. case) in cluster robust specifications. For that reason, I still use the Staiger and Stock (1997)'s rule that the F-statistic should be at least 10 for weak identification not to be considered a problem.

in that place what matters for people's opinion about state legitimacy. For that purpose I also construct the average state history of each respondent's ethnic groups based on the historical distribution of ethnic homelands (from Murdock 1959). The first specification in column 1 of Table 15 includes ethnic fixed effect and suggests that people living in districts with long history of statehood remarkably regard state institutions as more legitimate. In column 2 I do not include ethnic fixed effect and focus on the average state history of the ethnic group of the respondent. I do not find a statistically significant association between long history of statehood at the ethnic level and views of legitimacy of state institutions. When I introduce district-fixed effect the point estimate for state history of the respondent's ethnic group increases in size but remains statistically insignificant. These are indeed striking results since ethnicity is arguably one of the most relevant vehicles for cultural norms at the individual level. Therefore, the strong positive impact of state history at the district level on legitimacy (even when holding ethnic characteristic fixed) and the apparent nonexistent statistically association between the ethnic-based state history measure (when holding district characteristics fixed) strongly suggests that it is the long run exposure to statehood of the location, rather of the history of the ancestors of the people living in that location, what determines individual's belief about state legitimacy.

[Insert Table 15 here]

**State History, Trust in Institutions and Traditional Leaders**

Does history impose a legacy of confidence in state institutions? Lipset (1959) argued that state legitimacy is related to the capacity of a political system to convince its citizens that the prevailing institutions are not only appropriate but also the proper ones for them. Trust in state institutions is therefore a key element on which legitimacy is built. I next examine the statistical relationship between state history and trust in state institutions by estimating an analogue equation to (3) where the dependent variable is $Trust_{i,e,a,d,c}$ instead of State Legitimacy. $Trust$ is each individual's answer to different questions on several levels of trust. The question asked "How much do you trust in each of the following" and then it listed specific individuals or institutions. I recoded each original answer to a 5-point scale where 1 is "not at all" and 5 is "a lot". Following the methodology in Logan (2013), I coded the answers "don't know" at the mid-point. All the variables on the right-hand side of equation (3) are the same defined above for equation (2).

*Trust in State Institutions.* Table 16 presents IV estimates of the relationship between state history at the district level and measure of respondent's trust (OLS results along with Fuller (1) estimates are also reported in Table 16). All the specifications include the set of controls $I^{'}_{i,e,a,d,c}$, $A^{'}_{a,d,c}$, $D^{'}_{d,c}$, and the set of predetermined district-level controls using the IV validity exercise (migratory distance to Addis Adaba, ecological diversity, absolute latitude, suitability for millet and sorghum, and climate volatility measures ). In column 1 I focus on trust in institutions. This measure is based on the first principal components of each individual's answer regarding the level of trust in police and courts of law.[64]I find a strong positive statistical association between trust in institutions and state history at the district level.

*Do individuals living in districts with relative high historical exposure to statehood trust more in general?* Results in column 2 and 3 of Table 16 suggest that my previous results was not just picking up a higher level of trust. Respondents living in districts with long history of statehood do not trust more in relatives (column 2) or compatriots (column 3). All the coefficients are not statistically different from zero under usual levels of confidence. In fact, all the point estimates are negative and of the relative small size.

*State History and Trust in Politicians.* The point estimate in column 4 of Table 16 suggests that individual's trust in politicians is not strongly statistically related to state history at the district level (a first principal components of each individual's trust level in the president -or Prime Minister for some countries-, the parliament -or national assembly-, and the opposition political parties). Adding separately each of these components of this measure lead to similar results (not shown).

[Insert Table 16 here]

*The role of the Traditional Leaders.* Colonization did not eliminate several important pre-colonial obligations of the African traditional leaders. In fact, local traditional leaders still play an important role on the allocation of land and the resolution of local disputes. The way they still exercise public authority vary between and within countries (Logan, 2013). Michalopoulos and Papaioannou (2013b) document the strong influence of traditional leaders in governing the local community. It is precisely the interaction between local leaders and pre-colonial centralization what provides the foundation for Gennaioli and Rainer (2007b)'s main argument. I next analyze whether state history can explain popular support for local traditional leaders. The coefficient estimate for $\beta$ in column 1 of Table 17 shows that state history is not statistically associated to a measure of trust in local councilors. On the

[64]There is no question regarding trust on tax department in Afrobarometer 4.

contraty, results in column 2 suggests that individuals living in district with relative higher historical exposure to statehood tend to trust more in local traditional leaders. There is also a strong positive statistical association between state history and individuals' perception on the influence of traditional leaders governing the local community (column 3).[65] Moreover, the relationship between state history and preferred (instead of actual) influence of the traditional leaders is even statistically and economically stronger (column 4).[66]

[Insert Table 17 here]

## 5 Conclusion

This paper adds to a vibrant and growing literature in economics that seeks to better understand the role that historical factors play in shaping contemporary development outcomes. In particular, it contributes to the understanding of the developmental role of history of statehood by rigorously looking at the statistical relationship between state history and violent conflict at the local level. To do it, I introduce of a novel index of state antiquity at the local level. Motivated by the possibility that the construction of index may result in substantial measurement error and thus introduces a bias in OLS estimates, I follow a instrumental variable approach using the timing elapsed since the neolithic revolution as a source of exogenous variation. Both the IV and OLS estimates uncover a robust and strong empirical negative relationship between state history and contemporary conflict in a cross-section of 558 grid cells. I find similar results when using an alternative measure to account for the degreee of historical influence from political centralization based on the proximity to historical cities over the time period 1000 - 1800 CE.

I exploit then panel data variation in the prevalence of conflict, weather-induced productivity shocks, and the interaction of my state history index with those shocks to document that location with relatively high historical exposure to state capacity are remarkably less prone to experience conflict when hit by a negative agricultural productivity shock. The two

[65] The individuals answered the following question "How much influence do traditional leaders currently have in governing your local community?". The variable is coded in a 5-point scale from 1 (none) to 5 (great deal of influence). Again, I coded the answer "don't know" at the mid-point.

[66] The preferred influence variable is based on the following question "Do you think that the amount of influence traditional leaders have in governing your local community should increase, stay the same, or decrease?". The variable is coded in a 6-point scale from 1 (decrease a lot) to 5 (increase a lot). Again, I coded the answer "don't know" at the mid-point whereas and as missing value when people refused to answer.

aforementioned results are consistent with my hypothesis that locations with long history of state capacity should be better equipped of mechanisms to avoid conflict.

I then turn to specific mechanisms and examine an explanation for the uncovered statistical relationship between the history of statehood and conflict. I show that state history can be linked to people's positive attitudes towards state institutions. In particular, I show that key state institutions are regarded as more legitimate and trustworthy by people living in district with long history of state capacity.

Bearing in mind that identifying a causal effect of historical presence of statehood on contemporary conflict is a difficult task, I present empirical evidence that hard-to-account-for factors manifested in differences in long-run exposure to centralized institutions crucially matters to understand contemporary conflict.

## References


[1] Afrobarometer. 2007. "Round 4 Survey Manual." Available at: http://www.afrobarometer.org/files/documents/survey_manuals/survey_manual_r4_english.pdf

[2] Ahlerup, P. and O. Olson. 2012. "The Roots of Ethnic Diversity." Journal of Economic Growth 17(2), 71-102

[3] Ajayi, J. F. and M. Crowder. 1985. "Historical Atlas of Africa." Cambridge University Press,

[4] Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg. 2003. "Fractionalization." Journal of Economic Growth, 8:155-94.

[5] Alesina, A., P. Giuliano, and N. Nunn. 2013. "On the Origin of Gender Roles: Women and the Plough." Quarterly Journal of Economics 128, no. 2: 469-530.

[6] Alsan, M., 2013. "The Effects of the Tse-Tse Fly on African Development." mimeo.

[7] Angrist, J. D. and J.-S. Pischke. 2009. "Mostly Harmless Econometrics : An Empiricistís Companion." Princeton: Princeton University Press.

[8] Arbatli, E., Q. Ashraf, and O. Galor. 2013. "The Nature of Civil Conflict." mimeo.

[9] Ashraf, Q. and S. Michalopoulos. 2013. "Climatic Fluctuations and the Diffusion of Agriculture." NBER Working Paper 18765.

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

[11] Ashraf, Q. and O. Galor. 2013. "Genetic Diversity and the Origins of Cultural Fragmentation." American Economic Review Papers and Proceedings, 103(3): 528-533.

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

[13] Bandyopadhyay, S. and E. Green. 2012. "Pre-Colonial Political Centralization and Contemporary Development in Africa." Afrobarometer Working Paper No. 141.

[14] Bates, R. H. 1983. "Essays on the Political Economy of Rural Africa." Cambridge University Press, New York.

[15] Bates, R. H. 2013. "The Imperial Peace in Colonial Africa and Africa's Underdevelopmet". in Emmanuel Akyeampong, Robert H. Bates, Nathan Nunn, and James A. Robinson, eds., Africa's Development in Historical Perspective, Cambridge: Cambridge University Press, forthcoming.

[16] Barr, A. 2003. "Trust and expected trustworthiness: experimental evidence from Zimbabwean villages." Economic Journal, 113(489), 614-630.

[17] Barraclough. G. 1979. "The Times Atlas of World History. " Maplewood, N.J.

[18] Baum,C. F., M. E. Schaffer, and S. Stillman. 2007)."Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing," The Stata Journal, 7(4), 465–506.

[19] Ben-Amos Girshick, P., and J. Thornton. 2001. "Civil War in the Kingdom of Benin, 1689-1721: Continuity or Political Change?," Journal of African History, 42(3), 353-376.

[20] Benjaminsen, T.A., K. Alinon,, H. Buhaug, and J.T Buseth. 2012. "Does Climate Change Drive Land-use Conflicts in the Sahel," Journal of Peace Research 49(1): 97–111.

[21] Besley, T., and M. Reynal-Querol. 2012. "The Legacy of Historical Conflict. Evidence from Africa." Working paper.

[22] Blattman, C, and E. Miguel. 2010. "Civil War." Journal of Economic Literature, 48(1): 3-57.

[23] Bockstette, V., A. Chanda, and L. Putterman. 2002. "States and markets: The advantage of and early start." Journal of Economic Growth, 7, 347–369.

[24] Borcan, O., O. Olsson, and L. Putterman. 2013. "State History over Six Millenia: New Data and Results". mimeo.

[25] Brown, C., R. Meeks, K. Hunu, and W. Yu. 2011. "Hydroclimate risk to economic growth in sub-Saharan Africa." Climatic Change, 106 (4), 621-647.

[26] Buhaug, H. and J. K. Rød. 2006. "Local determinants of African civil wars, 1970-2001." Political Geography 25(3): 315 -335

[27] Burke, M., S.M. Hsiang, and E. Miguel. 2013. "Quantifying the Climatic Influence on Human Conflict, Violence and Political Instability." mimeo.

[28] Carter, D. 2011. "Sources of State Legitimacy in Contemporary South Africa: A Theory of Political Goods." Afrobarometer Working Paper No134.

[29] Chandler, T. 1987. "Four Thousand Years of Urban Growth: An Historical Census." Edwin Mellon Press, Lewiston, NY.

[30] Cederman, L. E. 2008. ''Articulating the Geo-Cultural Logic of Nationalist Insurgency.'' In Order, Conflict, and Violence, edited by Stathis N. Kalyvas, Ian Shapiro, and Tarek Masoud. Cambridge: Cambridge University Press.

[31] Cervellati, M., U. Sunde., and S. Valmori, "Disease Environment and Civil Conflicts," 2011. IZA Discussion Paper No. 5614.

[32] Conley, T.G. 1999. "GMM Estimation with Cross Sectional Dependence." Journal of Econometrics, 92(1): 1-45

[33] Collier, P. 1998. "The Political Economy of Ethnicity." Annual World Bank Conference on Development Economics, ed. B. Preskovic and J. E. Stigletz, 387-399. Washington, D.C.: The World Bank.

[34] Collier, P. and A. Hoeffler. 2004. "Greed and grievance in civil wars," Oxford Economic Papers, 56, 663-695.

[35] Collier, P. , A. Hoeffler, and M. Sonderbom. 2008. "Post-conflict risk," Journal of Peace Research, 45 (4): 461-478.

[36] Dee, D.P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C. , Thépaut, J.-N. , Vitart, F.. 2011. "The ERA-Interim reanalysis: configuration and performance of the data assimilation system," Quarterly Journal of the Royal Meteorological Society, 137 (656), 553-597.

[37] Depetris-Chauvin, N., Mulangu, F., and Porto, G. 2012 "Food Production and Consumption Trends in Sub-Saharan Africa: Prospects for the Transformation of the Agricultural Sector". Background paper for the First African Human Development Report, UNDP.

[38] Diamond, J. 1997. "Guns, Germs and Steel." W.W. Norton & Co., New York.

[39] Durante, R. 2009. "Risk, Cooperation and the Economic Origins of Social Trust: an Empirical Investigation," MPRA Paper 25887, University Library of Munich, Germany.

[40] Earth System Research Laboratory, NOAA, U.S. Department of Commerce. 2009. "NOAA-CIRES Twentieth Century Global Reanalysis Version II." Research Data Archive at the National Center for Atmospheric Research, Computational and Information.

[41] Easterly, W. and R. Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." The Quarterly Journal of Economics, 112(4): 1203-1250.

[42] Eggimann, G. 2000. "La Population des villes des tiers mondes de 1500 la 1950," Universite de Geneve, Centre d'histoire economique comparee.

[43] Eck, K. 2012. "In Data We Trust? A Comparison of UCDP GED and ACLED Conflict Events Datasets," Cooperation and Conflict, 47(1) 124-141.

[44] Ehret, C. 2002. "The civilizations of Africa: a history to 1800." University Press of Virginia.

[45] Ejiogu. E. C. 2011. “State Building in the Niger Basin in the Common Era and Beyond, 1000–Mid 1800s: The Case of Yorubaland.” Journal of Asian and African Studies 2011 46: 593.

[46] Englebert, P. 2000a. “State Legitimacy and Development in Africa.” Lynne Rienner Publishers.

[47] Englebert, P. 2000b. “Pre-Colonial Institutions and Post-Colonial States and Economic Development in Tropical Africa.” Political Reseach Qarterly 53(7) : 7-36.

[48] Esteban, J., L. Mayoral, and D. Ray. 2012. “Ethnicity and Conflict: An Empirical Study.” American Economic Review, 102(4): 1310-1342

[49] Fearon, J. D., and D. D. Laitin. 2003. ‘‘Ethnicity, Insurgency, and Civil War.’’ American Political Science Review 97 (1): 75-90.

[50] Fearon, J. D. 2005. “Primary commodities exports and civil war,” Journal of Conflict Resolution, 49, 4, 483-507.

[51] Fenske, J. 2012. “Ecology, trade and states in pre-colonial Africa.” Forthcoming, Journal of the European Economic Association.

[52] Gennaioli, N. and I. Rainer. 2007a. “Precolonial Centralization and Institutional Quality in Africa.” In M. Gradstein and K. Konrad, ed., Institutions and Norms in Economic Development, MIT Press.

[53] Gennaioli, N. and I. Rainer. 2007b. “The Modern Impact of Precolonial Centralization in Africa,” Journal of Economic Growth. 12:185-234

[54] Gilley, B. 2006. “The meaning and measure of state legitimacy: Results for 72 countries”. European Journal of Political Research 45: 499-525.

[55] Green, Elliott. 2012. “On the Size and Shape of African States.” International Studies Quarterly 56(2): 229–244.

[56] Harari, M. and E. La Ferrara. 2013. “Conflict, Climate and Cells: A Disaggregated Analysis.” CEPR Discussion Paper no. 9277. London, Centre for Economic Policy Research. http://www.cepr.org/pubs/dps/DP9277.asp.

[57] Hariri, J.G. 2012. “The Autocratic Legacy of Early Statehood” American Political Science Review, 106 ( 3) 471-494.

[58] Haber, S. 2012. "Rainfall and Democracy: Climate, Technology, and the Evolution of Economic and Political Institutions." mimeo.

[59] Herbert, L. S. 1966. "The Origins of African Kingdoms." Cahiers d'études africaines. Vol.6 N°23., 402-407.

[60] Herbst, J. 2000 "States and Power in Africa: Comparative Lessons in Authority and Control." Princeton University Press.

[61] Huillery, E. 2009. "History Matters: The Long Term Impact of Colonial Public Investments in French West Africa," American Economic Journal - Applied Economics, 1(2): 176-215.

[62] Hyslop, D., and G. Imbens. 2001. "Bias from Classical and Other Forms of Measurement Error," Journal of Business and Economic Statistics.

[63] La Porta, R., F. Lopez-De Silanes, A. Schleifer, and R. Vishny. 1999. "The quality of government", Journal of Law, Economic, and Organisation, 15: 222-279.

[64] Law, R. 1977. "The Oyo empire: c. 1600-c. 1836: a west African imperialism in the era of the Atlantic slave trade." Clarendon press.

[65] Lewis, H. S. 1966. "The Origins of African Kingdoms" In: Cahiers d'études africaines. 6 (23), 402-407.

[66] Liu, H., F. Prugnolle, A. Manica, and F. Balloux. 2006. "A geographically explicit genetic model of worldwide human-settlement history." American Journal of Human Genetics, 79, 230–237.

[67] Logan , C. 2013. "The Roots of Resilience: Exploring Popular Support for African Traditional Authorities." African Affairs, 112 (448): 353-376.

[68] McEvedy, C. 1995. "Penguin Atlas of African history." New ed., rev. ed. London; New York: Penguin Books.

[69] Michalopoulos, S. 2012. "The Origins of Ethnnolinguistic Diversity." American Economic Review, 102(4): 1508-39.

[70] Michalopoulos, S. and E. Papaioannou. 2012. "The Long Run Effects of the Scramble for Africa," NBER w17620.

[71] Michalopoulos, S. and E. Papaioannou. 2013."Pre-colonial Ethnic Institutions and Contemporary African Development." Econometrica, 81(1): 113–152.

[72] Michalopoulos, S and E. Papaioannou. 2013b. "On the Ethnic Origins of African Development: Traditional Chiefs and Pre-colonial Political Centralization." mimeo.

[73] Miguel, E., and M. K. Gugerty. 2005. "Ethnic Diversity, Social Sanctions and Public Goods in Kenya," Journal of Public Economics, 89, 2325–2368.

[74] Miguel, E., S. Satyanath, and E. Sergenti. 2004. "Economic Shocks and Civil Conáict: An Instrumental Variables Approach." Journal of Political Economy 112 (4): 725-53.

[75] Moore, J. L., L. Manne, T. Brooks, N. D. Burgess, R. Davies, C. Rahbek, P. Williams and A. Balmford. 2002. "The Distribution of Cultural and Biological Diversity in Africa." Proceedings: Biological Sciences 269 (1501): 1645-1653

[76] Murdock, G. P. 1959. "Africa: Its People and their Culture History," McGraw-Hill Book Company, New York.

[77] Murdock, G. P. 1967. "Ethnographic Atlas." University of Pittsburgh Press.

[78] Nelson, A. 2004. "African Population Database Documentation, UNEP GRID Sioux Falls." Available at: http://na.unep.net/siouxfalls/datasets/datalist.php

[79] Nunn, N. 2008. "The Long Term Effects of Africa's Slave Trades." Quarterly Journal of Economics. 123(1):139-176.

[80] Nunn, N. 2013. "Historical Development." Harvard University. mimeo

[81] Nunn N, and D. Puga . 2012. "Ruggedness: The Blessing of Bad Geography in Africa." Review of Economics and Statistics;94(1):20-36.

[82] Nunn N, and L. Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." American Economic Review. 101(7):3221-3252.

[83] Olson, O, and C. Paik. 2012. "A Western Reversal Since the Neolithic? The Long-Run Impact of Early Agriculture." mimeo.

[84] Osafo-Kwaako, P. and J. Robinson. 2013. "Political centralization in pre-colonial Africa," Journal of Comparative Economics, 41 (1): 6-21.

[85] Özak, Ö. 2012. "The Voyage of Homo-œconomicus: some economic measures of distance." mimeo.

[86] Özak, Ö. 2012. "Distance to the Technological Frontier and Economic Development," Departmental Working Papers 1201, Southern Methodist University, Department of Economics.

[87] Petersen, M.B. and S.E. Skaaning. 2010. "Ultimate Causes of State Formation: The Signifficance of Biogeography, Diffusion, and Neolithic Revolutions," Historical Social Research 35(3), 200-226.

[88] Pinhasi, R., J. Fort, and A. J. Ammerman. 2005. "Tracing the Origin and Spread of Agriculture in Europe." PLoS Biology, 3(12): 2220-2228.

[89] Putterman, L. 2006. "Agricultural Transition Year: Country Data Set," Brown University. mimeo.

[90] Raleigh, C. and H. Hegre. 2009. "Population size, concentration, and civil war. A geographically disaggregated analysis," Political Geography, 8 (4): 224-238.

[91] Ramachandran, S., O. Deshpande, C. C. Roseman, N.A. Rosenberg, M.W. Feldman, and L. Cavalli-Sforza. 2005. "Support for the relationship of genetic and geographic distance in human populations for a serial founder effect origininating in Africa." Proceedings of the National Academy of Science USA, 102(44), 15942–15947.

[92] Reid, R. 2012. "Warfare in African History." Cambridge University Press.

[93] Schlenker, W., and D. B. Lobell. 2010. "Robust negative impacts of climate change on African agriculture." Environ. Res. Lett. 5 (1).

[94] Selway, J. and K. Templeman. 2012."The Myth of Consociationalism? Conflict Reduction in Divided Societies." Comparative Political Studies 45 (12): 1-30.

[95] Shaw, T. editor. 1993. "The Archaeology of Africa: Foods, Metals, and Towns." London: Routledge.

[96] Spoloare, E. and R. Wacziarg. 2013. "How Deep Are the Roots of Economic Development?" Journal of Economic Literature, 51(2), 1-45.

[97] Staiger, D., and J. H. Stock. 1997. "Instrumental variables regression with weak instruments." Econometrica 65(3): 557–86.

[98] Stock, J. H., J. H. Wright, and M. Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." Journal of Business and Economic Statistics, 20(4): 518-529.

[99] Sundberg, R, M. Lindgren and A. Padskocimaite. 2010. "UCDP GED Codebook version 1.0-2011," Department of Peace and Conflict Research, Uppsala University.

[100] Sundberg, R, and E. Melander. 2013. "Introducing the UCDP Georeferenced Event Dataset," Journal of Peace Research, 50(4), 523-532.

[101] Thyne, C. L. 2012. "Information, Commitment, and Intra-War Bargaining: The effect of Governmental Constriants on Civil War Duration." International Studies Quarterly. 56 (2) 307-321.

[102] Vansina, J. 1962. "A Comparison of African Kingdoms." Journal of the International African Institute, 32 (4) 324-335.

[103] Vicente-Serrano S.M., Begueria S., Lopez-Moreno J.I., Angulo M., El Kenawy A. 2010. "A global 0.5 gridded dataset (1901-2006) of a multiscalar drought index considering the joint effects of precipitation and temperature." Journal of Hydrometeorology 11(4), 1033-1043.

[104] Whatley, W., and R. Gillezeau. 2011a. "The Fundamental Impact of the Slave Trade on African Economies." In Economic Evolution and Revolution in Historical Time, ed. P. Rhode, J. Rosenbloom and D. Weiman. Stanford: Stanford University Press.

[105] Whatley, W., and R. Gillezeau. 2011b. "The Impact of the Transatlantic Slave Trade on Ethnic Stratification in Africa." American Economic Review, 101(3): 571-76.

[106] White, F. 1983. "The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa." Natural resources research, 20:1–356.

[107] Wig, T. 2013. "Peace from the Past: Pre-colonial Political Institutions and Contemporary Ethnic Civil Wars in Africa." mimeo.

[108] Wrigley, N., Holt, T., Steel, D., and Tranmer, M. 1996. "Analysing, modelling and resolving the ecological fallacy." In P. Longley, & M. Batty (Eds.), Spatial analysis: Modelling in a GIS environment (pp. 23-40). Cambridge, UK: Geo-information International.

## Appendix A: Variable Definitions (UNDER CONSTRUCTION).

### Cross section

*Conflict Prevalence*: fraction of years with at least one conflict event in the grid cell during the period 1989-2010. Own calculations based on UCDP GED, version 1.5 (November 2012).

*Conflict Onset:* fraction of years with at least one conflict onset in the grid cell. An onset is the first confrontation event within a dyad. Own calculations based on UCDP GED, version 1.5 (November 2012).

*Local State History:* See main text for definition.

*Area:* Total land area of the grid cell (in square kilometers). Own calculations.

*Distance Ocean:* Distance from the centroid of the grid cell to the nearest ocean (in hundred of kilometers). Own calculations.

*Distance Major River:* Distance from the centroid of the grid cell to the nearest major river (in hundred of kilometers). Major river is ... Own calculations.

*Capital Dummy:* Variable taking value 1 if the capital city of the country to which the grid cell was assigned is in the grid cell. Own calculations.

*Distance Capital:* Distance from the centroid of the grid cell to the capital city of the country to which the grid cell was assigned (in kilometers). Own Calculations.

*Total Road Length:* Total length of major roads intersecting the grid cell (in hundred of kilometers). Own calculations. SOURCE.

*Mean elevation:* Average elevation of the terrain (in meters above the sea level). Own calculations. The value is computed by averaging across all original pixels within a 2 by 2 degree grid cell. The data comes from National Oceanic and Atmospheric Administration (NOAA) and U.S. National Geophysical Data Center, TerrainBase, release 1.0, Boulder, Colorado. Available at http://www.sage.wisc.edu/atlas/data.php?incdataset=Topography

*Ruggedness:* Average ruggedness of the terrain based within a grid cell. The ruggedness measure comes from Nunn and Puga (2012).

30-by-30 arc-second cell

*Natural Resources Dummy:* Variable taking value 1 if at least one natural resource site (either gems, diamond, gas or oil) is located in the grid cell. Location of the natural resource sites comes from CITATION

*Number of Countries in Grid:* Total number of countries that are intersected by the grid cell. South Sudan is included in Sudan.

*Ethnic Fractionalization in 1960:* This variable is computed at the grid level $i$ with the following formula: $F_i = 1 - \sum_{g=1}^{n} \alpha_{i,g}^2$. Where $\alpha_{i,g}$ is the fraction of total population in grid cell i that live in the portion of the historical homeland of group $g$ that is intersected by the grid $i$. Population counts are from 1960 and comes from UNEP GRID Sioux Falls (Nelson 2004). The spatial distribution of ethnic groups is based on Murdock's (1959) map.

*Ln of Population Density:* Log of 1 + population density in 1960 (people per squared kilometer). Population data comes from UNEP GRID Sioux Falls (Nelson 2004).

*Pre-Colonial Variables:* The following variables are 1960 population-weighted averages of traits of ethnic groups whose historical homelands intersect a given grid cell. The weights are the aforementioned $\alpha_{i,g}$. Pre-colonial dependence variables denote subsistence income shares derived from hunting, fishing, pastoralism, and agricultural (variables v2, v3, v4, and v5 in the Ethnographic Atlas (1967) respectively). TALK ABOUT MATCHING

Pre-Colonial Hunting Dependence Pre-Colonial Fishing Dependence Pre-Colonial Pastoralism Dependence Pre-Colonial Agricultural Dependence Pre-Colonial Settlement Pattern TO BE CONTINUED....

## Appendix B: Construction of the Instrument

Recent works in economics (Olson and Paik, 2012, and Ashraf and Michalopoulos, 2012) have used Pinhasi et al (2005)'s data on location and calibrated C14-dates estimated for archaeological sites to construct the timing of the initial adoption of agriculture at fine geographical level (such as region within countries). Unfortunately, Pinhasi et al (2005)'s data only cover neolithic settlements in Europe and the Middle East. In addition, African archaeology is relative new and has not yet accumulated the density of data as in Europe and America (Shaw et al, 1993). However, we do know that, unlike other regions, plant domestication did not spread from a single point within the African continent: adoptions of indigenous crops independently occurred at least in five different regions of Africa.[67] I use a variety of archaeological sources to compile geographic location and date of earliest plant domestication for 24 representative archaeological sites in Sub-Saharan Africa.[68] I then

[67] Evidence places those regions of first domestication in the Western Sahel (pearl millet), Middle Niger delta (african wild rice), West Africa (yam), Sudan-Chad (sorghum), and Ethiopia (teff).

[68] Table A.2 in appendix lists the archaeological sites.

follow Olson and Paik (2012) and construct a continuous raster map of the time elapsed since the neolithic by interpolating the earliest date of crop adoption from my compiled data of archaeological sites. The interpolation is based on an inverse distance weighting method which relies on a underlying assumption that the closest archeological site provides the best information on the approximate earliest date of crops adoption. Formally, the date of earliest domestication of crops for the location $i$ (i.e: a pixel in the raster map) will be given by:

$$\widehat{T}(L_i) = \sum_{s}^{24} \gamma_i T(L_s)$$

Where $T(L_s)$ is the date of earliest domestication of crops in the location $L$ of archeological site $s$ and $\gamma_i = d_{i,s}^{-p} / \sum_{s}^{24} d_{i,s}^{-p}$ is a weight factor with $d_{i,s}$ being distance between location $i$ and location of site $s$. Finally, $p$ is a power parameter determined by minimizing the root-mean-square prediction error.

Figure 5 (in main text) depicts the geographical distribution of the predicted time elapsed since the neolithic revolution based on archaeological data. The diamond figures represent the location of the archaeological sites (see appendix for list of sites with their source of references), the numbers next to each figure represent the date of earliest domestication of crops.[69] Note that the earliest date of domestication (7200 years before present times -BP-) is located in Faiyum region of Egypt and does not represent a case of domestication of an indigenous crop rather a diffusion of agriculture from the Fertile Crescent, which in fact spread even southward over central Sudan around 5000 BP. Sorghum remains dating to 4000 BP in the Adrar Bous site in the Ténéré desert (Niger) represents the earliest evidence of indigenous crop domestication. Archaeological sites in Mali, Ghana, and Mauritania are the earliest evidence of domestication of pearl millet around 3500 BP. By 2600 BP agriculture already spread into northern Senegal, and by 2300 BP in southern Cameroon. The archaeological site with the earliest date of domestication, around 3200BP, in East Africa is located in Kenya. Agriculture adoption happened much later in Malawi (1900 BP), northern South Africa (1700 BP), Zambia (1400 BP), and northwestern Botswana (1100 BP).

[69] To avoid an edge effect and fully cover the Sub-Saharan surface in the interpolation process I include 4 edges denoted with the letter X (See Table A.2 in appendix). The edge effect results from the fact that the inverse distance interpolation method cannot estimated a value located beyond the most extreme known value.

## Appendix C: Construction of Weather-Induced Productivity Shock

I construct a weather-induced productivity shock to the agricultural sector in two main steps which I explain in detail below. In the first step I construct five crop-specific weather shocks. In the second step I aggregate these shocks into one indicator. As in Harari and Laferrara (2013), I construct my weather shocks using the Standardized Precipitation-Evapotranspiration Index (SPEI) developed by Vicente-Serrano et al (2010).

*Drought index.* The SPEI is a multiscalar drought index, which considers the joint effects of temperature and precipitation on droughts (Vicente-Serrano et al 2010). The SPEI is based on the climatic water balance equation which depends on total precipitation and the capacity of the soil to retain water (i.e: evapotranspiration). Formally; the water balance equation for a given month $t$:

$$D_t = Prec_t - PET_t,$$

where $Prec_t$ and $PET_t$ are precipitation and potential evapotranspiration (both in mm), respectively. The PET need to be estimated using different climate inputs (such as temperature, cloud cover, and wind speeds) of which temperature is the most relevant. This water balance (deficit or superavit) can be aggregated at different scales $k$ (i.e: number of months). Then, a given $D_t^k$ is fitted to a Log-logistic distribution to obtain the $SPEI_t^k$ for a given month $t$ and scale $k$ over which water deficits/superavits accumulate. Since the SPEI is a standardized variable (with mean value of zero and standard deviation of 1), it can be compared over time and space (Vicente-Serrano et al, 2010) regardless of the election of $k$ and $t$.[70] Low and negative values of the SPEI denote relative high water balance deficits (Droughts).

As discussed in Harari and Laferrara (2013), the original SPEI series are based on CRU TS3.0 data which relies on gauge data. This poses a problem in the context of Sub-Saharan Africa where gauge data (in particular historical data) is scarce, then highly interpolated, and potentially endogenous to the existence of conflict. I therefore recalculate all the necessary SPEI series using more reliable climate data from ECMWF ERA-Interim dataset (Dee et al., 2011) and the NOAA 20th century reanalysis (Earth System Research Laboratory, NOAA, U.S. Department of Commerce, 2009), and the R package provided by the authors of the original index. In the appendix I provide the details for the calculation of all the SPEI series used to create my weather shock variable.

[70]In other words, the SPEI is measured in units of standard deviation from the historical average of the water balance (i.e: average over the period for which input climatic variables are available).

*Crop-specific weather shocks.* I focus on five staple crops: sorghum, millet, cassava, groundnuts, and maize. According to Schlenker and Lobell (2010), these crops are among the most relevant nutritional sources of calories, protein, and fat in Sub-Saharan Africa. They are also among the most relevant staple crops in terms of production (Depetris-Chauvin et al, 2012). In addition, these crops are highly dependent on rain. Although rice and wheat are also very relevant for this region, I excluded them from my analysis because they are highly irrigated (Schenkler and Lobell, 2010).[71] I then follow the main strategy in Harari and Laferrara (2013). For each grid cell and each of the five aforementioned crops I identify the planting and harvesting months.[72] Therefore, I identify the length of the growing season ($k$) and the harvest month ($t$) for each crop in each grid-cell.[73] Hence, for a given year, $SPEI_{t_{c,i}}^{k_{c,i}}$ represents a weather shock specific to the crop $c$ in grid $i$.

*Weather-Induced Agricultural Productivity Shock.* I create an aggregate weather-induced agricultural productivity shock for each grid $i$ and year $T$ by doing:

$$Negative\,Weather\,Shock_{i,T} = -\sum_{c} \theta_{c,i} \times SPEI_{t_{c,i},T}^{k_{c,i}}$$

where $k_{c,i}$ and $t_{c,i}$ are growing season length and harvest month for crop $c$ in grid $i$, respectively. $\theta_{c,i}$ are the normalized harvest shares for each crop $c$ in grid $i$.[74] There are two main departures from Harari and Laferrara (2013) regarding the methodology implemented to create the shock. First, instead of focusing in the main crop (in term of harvested area) within a set of twenty six possible crops, I focus on the five most popular rainfed crops for Sub-Saharan Africa and use their relative importance (in terms of harvested area) to weight them in the aggregation within a grid cell. Second, Harari and Laferrara (2013) define weather shock as the fraction of consecutive growing season months presenting an SPEI of 4 months of accumulation (scale 4) that is one standard deviation below the historical mean.They do mention that their results are robust to different time scales. I am less agnostic regarding the relevant scale (i.e: the number of months over which water deficits/superavits accumulate) and force it to be determined by the length of each growing season, instead.[75] My approach allows for a more parsimonious definition of shocks and makes possible the distinction between moderate and extreme drought events (namely, between an SPEI value

[71] Since spatial variation in irrigation technologies is expected to be highly correlated with weather variation, including highly irrigated-crops would underestimate the statistical relationship between crop-specific weather shocks and conflict.

[72] All the information on crop calendars comes from Mirca 2000. See appendix for details.

[73] In some regions a crop may have two growing seasons within a year; I focus only in the primary season.

[74] The shares of areas harvested for each crop are calculated based on M3-Crops. See appendix for details.

[75] I thank Santiago Bergueria -one of the authors of the SPEI- for this suggestion.

of -1 and -3).

Table 1: Summary Statistics. Grid Cell Sample

| Variable | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|
| Conflict Prevalence | 0.19 | 0.23 | 0.00 | 1.00 |
| Conflict Onset | 0.27 | 0.44 | 0.00 | 1.00 |
| Local State History 1000 - 1850 CE | 0.16 | 0.23 | 0.00 | 1.00 |
| Area (square km) | 42367 | 12239 | 122 | 49231 |
| Distance Ocean ('00 km) | 6.18 | 4.75 | 0.00 | 16.84 |
| Distance Major River ('00 km) | 4.17 | 3.25 | 0.00 | 15.58 |
| Capital Dummy | 0.07 | 0.26 | 0.00 | 1.00 |
| Distance Capital (km) | 641.7 | 435.9 | 24.7 | 1912.5 |
| Total Road Length ('00 km) | 4.36 | 4.74 | 0.00 | 38.58 |
| Mean Elevation (m) | 616.5 | 425.4 | -4.6 | 2221.9 |
| Ruggedness | 66583 | 78892 | 960 | 540434 |
| Natural Resources Dummy | 0.41 | 0.49 | 0.00 | 1.00 |
| Number of Countries in Grid | 1.62 | 0.72 | 1.00 | 4.00 |
| Cereal Suitability | 0.28 | 0.17 | 0.00 | 0.71 |
| TseTse Fly Suitability | 0.34 | 0.40 | 0.00 | 1.00 |
| Malaria Ecology early 20th Century | 5.71 | 4.90 | 0.00 | 18.52 |
| Ethnic Fractionalization in 1960 | 0.45 | 0.27 | 0.00 | 1.00 |
| ln of Population Density in 1960 | 0.10 | 0.14 | 0.00 | 0.91 |
| Pre-Colonial Hunting Dependence | 0.96 | 0.91 | 0.00 | 4.00 |
| Pre-Colonial Fishing Dependence | 0.72 | 0.72 | 0.00 | 5.21 |
| Pre-Colonial Pastoralism Dependence | 3.15 | 2.39 | 0.00 | 9.00 |
| Pre-Colonial Agricultural Dependence | 4.46 | 2.12 | 0.00 | 8.41 |
| Pre-Colonial Settlement Pattern | 4.77 | 2.23 | 1.00 | 8.00 |
| Slave Trade (log of Slave Exports/Area) | 4.31 | 4.19 | 0.00 | 14.36 |
| Distance to Historical Conflict ('00 km) | 5.74 | 3.50 | 0.09 | 16.79 |
| Ecological Diversity (Herfindhal Index) | 0.32 | 0.23 | 0.00 | 0.75 |
| Migratory Distance to Addis Adaba (weeks) | 4.70 | 2.24 | 0.03 | 9.45 |

Note: Sample Size is 558 grid cells. See full details of the variable definitions in Appendix

Table 2: OLS Estimates - Baseline Specification

| Dependent Variable: Conflict Prevalence 1989-2010 (fraction of years with at least one conflict event) | | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Local State History 1000 - 1850 CE | -0.109** | -0.191*** | -0.193*** | -0.197*** | -0.196*** |
| robust std err | (0.045) | (0.041) | (0.041) | (0.042) | (0.042) |
| spat. adj. std err (5 degrees) | (0.054) | (0.047) | (0.047) | (0.048) | (0.048) |
| spat. adj. std err (10 degrees) | (0.064) | (0.054) | (0.053) | (0.053) | (0.056) |
| std err clust country | (0.075) | (0.061) | (0.057) | (0.061) | (0.060) |
| Country Dummies | Y | Y | Y | Y | Y |
| Geo-strategic Controls | N | Y | Y | Y | Y |
| Cereal Suitability | N | N | Y | N | Y |
| Disease Environment | N | N | N | Y | Y |
| Observations | 558 | 558 | 558 | 558 | 558 |
| R-squared | 0.349 | 0.510 | 0.517 | 0.516 | 0.519 |

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ (robust case). The unit of observation is a grid cell. The geo-strategic controls are distance to ocean, distance to major river, distance to capital, capital dummy, total road length, mean elevation, ruggedness terrain, total area, dummy for natural resources sites, and number of countries intersected by the grid. Cereal suitability represents the soil suitability for cultivating cereals (FAO's GAEZ database). Disease environment control include malaria ecology in early 20th century and TseTse fly suitability (predicted distribution from FAO).

Table 3: OLS Estimates - Accounting for Genetic and Ecological Diversity

| Dependent Variable: Conflict Prevalence 1989-2010 (fraction of years with at least one conflict event) | | | |
|---|---|---|---|
| Local State History 1000 - 1850 CE | -0.185*** | -0.209*** | -0.198*** |
| | (0.064) | (0.060) | (0.063) |
| Migratory Distance to Addis Adaba | -0.059*** | | -0.056*** |
| | (0.019) | | (0.018) |
| Ecological Diversity | | 0.117** | 0.111** |
| | | (0.052) | (0.049) |
| Country Dummies | Y | Y | Y |
| Geo-strategic Controls | Y | Y | Y |
| Cereal Suitability | Y | Y | Y |
| Disease Environment | Y | Y | Y |
| Observations | 558 | 558 | 558 |
| R-squared | 0.534 | 0.529 | 0.543 |

Robust standard errors clustered at the country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The unit of observation is a grid cell. The basic set of controls is described in Table 2. Migratory Distance to Addis Adaba proxies for genetic diversity. The longer the distance to Addis Adaba, the lower the genetic diversity. Ecological diversity is a Herfindhal index based on Vegetation types from White (1983).

Table 4: OLS Estimates - Additional Controls

| Dependent Variable: Conflict Prevalence 1989-2010 (fraction of years with at least one conflict event) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Local State History 1000 - 1850 CE | -0.199*** | -0.223*** | -0.200*** | -0.199*** | -0.194*** | -0.212*** | -0.213*** | -0.233*** |
| | (0.055) | (0.060) | (0.062) | (0.064) | (0.064) | (0.062) | (0.064) | (0.064) |
| Additional Control | Precolonial Prosperity | Ln of Pop Density | Ethnic Fraction | Slave Trade | Hist Trade Routes | Distance Hist Conflict | Light Density | All |
| Coefficient Add. Control | P-value Joint Sig. | 0.367** | 0.051 | 0.001 | -0.015 | -0.006 | 0.026* | P-value Joint Sig. |
| | [0.11] | (0.146) | (0.037) | (0.005) | (0.033) | (0.007) | (0.014) | [0.0296] |
| Country Dummies | Y | Y | Y | Y | Y | Y | Y | Y |
| Geo-strategic Controls | Y | Y | Y | Y | Y | Y | Y | Y |
| Cereal Suitability | Y | Y | Y | Y | Y | Y | Y | Y |
| Disease Environment | Y | Y | Y | Y | Y | Y | Y | Y |
| Migratory Distance to Addis Adaba | Y | Y | Y | Y | Y | Y | Y | Y |
| Ecological Diversity | Y | Y | Y | Y | Y | Y | Y | Y |
| Observations | 558 | 558 | 558 | 558 | 558 | 558 | 558 | 558 |
| R-squared | 0.549 | 0.558 | 0.545 | 0.543 | 0.543 | 0.545 | 0.550 | 0.565 |

Robust standard errors clustered at the country level in parentheses.*** $p<0.01$, ** $p<0.05$, * $p<0.1$. The unit of observation is a grid cell. The basic set of controls is described in Tables 2 and 3. See appendix for definition of the additional controls.

Table 5: OLS Estimates. Different Conflict Measures

| | Dependent Variable | | | | | |
|---|---|---|---|---|---|---|
| | All Conflicts | Battles | Violence | Riots | Log of Casualties | Conflict Onset |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Local State History 1000 - 1850 CE | -0.093 | -0.157** | -0.124* | 0.011 | -1.408*** | -0.277*** |
| | (0.061) | (0.060) | (0.063) | (0.030) | (0.469) | (0.089) |
| Dataset | ACLED | ACLED | ACLED | ACLED | UCDP-GED | UCDP-GED |
| Country Dummies | Y | Y | Y | Y | Y | Y |
| Geo-strategic Controls | Y | Y | Y | Y | Y | Y |
| Cereal Suitability | Y | Y | Y | Y | Y | Y |
| Disease Environment | Y | Y | Y | Y | Y | Y |
| Migratory Distance to Addis Adaba | Y | Y | Y | Y | Y | Y |
| Ecological Diversity | Y | Y | Y | Y | Y | Y |

Robust standard errrors clustered at the country level in parentheses.*** p<0.01, ** p<0.05, * p<0.1.
The unit of observation is a grid cell. Conflict measures in columns 1,2,3,4, and 6 represent the fraction of years with at leat one conflict event in grid. ACLED data comprises the period 1997-2010. Conflict onset is defined as the first event within a dyad. The set of controls is described in Tables 2 and 3.

Table 6: OLS Estimates. Heterogeneity Across Regions

| Dependent Variable: Conflict Prevalence 1989-2010 (fraction of years with at least one conflict event) | | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Local State History 1000 - 1850 CE | -0.133* | -0.223* | -0.171* |
| | (0.068) | (0.105) | (0.088) |
| Regions Included in Sample | West-Central | East-South | All but West |
| Country Dummies | Y | Y | Y |
| Geo-strategic Controls | Y | Y | Y |
| Cereal Suitability | Y | Y | Y |
| Disease Environment | Y | Y | Y |
| Migratory Distance to Addis Adaba | Y | Y | Y |
| Ecological Diversity | Y | Y | Y |
| Observations | 274 | 284 | 425 |
| R-squared | 0.539 | 0.632 | 0.547 |

Robust standard errors clustered at the country level in parentheses.*** p<0.01, ** p<0.05, * p<0.1. The unit of observation is a grid cell. The basic set of controls is described in Tables 2 and 3.

Table 7: OLS Estimates. Discount Factors and Importance of Medieval Period

| Dependent Variable: Conflict Prevalence 1989-2010 (fraction of years with at least one conflict event) | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 5% Discount | 10% Discount | 25% Discount | 50% Discount | 0% Discount | 0% Discount | 0% Discount |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Discounted Local State History 1000 - 1850 CE | -0.176*** | -0.157** | -0.114* | -0.107 | | | |
| | (0.061) | (0.060) | (0.057) | (0.067) | | | |
| Local Statehist 1500-1850 CE | | | | | -0.099* | | -0.039 |
| | | | | | (0.051) | | (0.042) |
| Local Statehist 1000-1500 CE | | | | | | -0.195*** | -0.175*** |
| | | | | | | (0.063) | (0.063) |
| Beta coefficient | -0.185 | -0.170 | -0.130 | -0.099 | -0.120 | -0.201 | na |
| Country Dummies | Y | Y | Y | Y | Y | Y | Y |
| Geo-strategic Controls | Y | Y | Y | Y | Y | Y | Y |
| Cereal Suitability | Y | Y | Y | Y | Y | Y | Y |
| Disease Environment | Y | Y | Y | Y | Y | Y | Y |
| Migratory Distance to Addis Adaba | Y | Y | Y | Y | Y | Y | Y |
| Ecological Diversity | Y | Y | Y | Y | Y | Y | Y |
| Observations | 558 | 558 | 558 | 558 | 558 | 558 | 558 |
| R-squared | 0.540 | 0.537 | 0.531 | 0.527 | 0.530 | 0.544 | 0.545 |

Robust standard errors clustered at the country level in parentheses.*** p<0.01, ** p<0.05, * p<0.1.
The unit of observation is a grid cell. The basic set of controls is described in Tables 2 and 3.

Table 8: OLS Estimates - Intensive vs Extensive Margin of Political Centralization

| Dependent Variable: Conflict Prevalence 1989-2010 (fraction of years with at least one conflict event) | | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| State History Score 1800 CE | -0.048 | | | |
| | (0.038) | | | |
| Ethnic Centralization (v33 Eth. Atlas) | | -0.019 | | |
| | | (0.015) | | |
| Local State History 1000 - 1850 CE | | | -0.207*** | -0.217** |
| | | | (0.057) | (0.101) |
| Country Dummies | Y | Y | Y | Y |
| Geo-strategic Controls | Y | Y | Y | Y |
| Cereal Suitability | Y | Y | Y | Y |
| Disease Environment | Y | Y | Y | Y |
| Migratory Distance to Addis Adaba | Y | Y | Y | Y |
| Ecological Diversity | Y | Y | Y | Y |
| Sample | Full | Full | Local State Hisory > 0 | >100 years of State |
| Observations | 558 | 558 | 335 | 229 |
| R-squared | 0.524 | 0.524 | 0.559 | 0.600 |

Robust standard errors clustered at the country level in parentheses.*** p<0.01, ** p<0.05, * p<0.1. The unit of observation is a grid cell. State History Score 1800 CE represents the fraction of grid which was under a centralized state during the period 1800 - 1850 CE.Ethnic Centralization is 1960-population weighted average of Ethnographic Atlas's variable v33 (JurisdictionalHierarchy Beyond Local Community) ranging from 1 to 5 (Large States).The basic set of controls is described in Tables 2 and 3.

Table 9: First-Stage. Neolithic Instrument

| Dependent Variable: Local State History 1000 - 1850 CE | | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Time Elapsed Since Neolithic | 0.185** | 0.054 | 0.250*** | |
| spat. adj. std err (5 degrees) | (0.049) | (0.078) | (0.051) | |
| spat. adj. std err (10 degrees) | (0.055) | (0.087) | (0.060) | |
| std err clust country | (0.085) | (0.147) | (0.079) | |
| Square of Time Elapsed Since Neolithic | | | | 0.034*** |
| spat. adj. std err (5 degrees) | | | | (0.007) |
| spat. adj. std err (10 degrees) | | | | (0.008) |
| std err clust country | | | | (0.012) |
| Country Dummies | Y | Y | Y | Y |
| Geo-strategic Controls | Y | Y | Y | Y |
| Cereal Suitability | Y | Y | Y | Y |
| Disease Environment | Y | Y | Y | Y |
| Migratory Distance to Addis Adaba | Y | Y | Y | Y |
| Ecological Diversity | Y | Y | Y | Y |
| Sample | Full | Central Africa | Restricted | Full |
| Observations | 558 | 141 | 465 | 558 |
| R-squared | 0.456 | 0.533 | 0.471 | 0.467 |

*** p<0.01, ** p<0.05, * p<0.1 (clustered at country level case). The unit of observation is a grid cell.
Specification in Column 3 excludes observations from Gabon, Congo, Congo DR, and Angola.
The basic set of controls is described in Tables 2 and 3.

Table 10: IV Estimates

| Dependent Variable: Conflict Prevalence 1989-2010 (fraction of years with at least one conflict event) | | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Local State History 1000 - 1850 CE | -0.712*** | -0.618*** | -0.718*** | -0.527** | -0.599*** | -0.410** | -0.386** |
| | (0.209) | (0.202) | (0.204) | (0.221) | (0.225) | (0.186) | (0.186) |
| Absolute Latitude | | -0.010** | | | | -0.003 | 0.002 |
| | | (0.004) | | | | (0.006) | (0.006) |
| Suitability for Sorghum and Millet Cultivation | | | -0.008 | | | 0.014 | -0.026 |
| | | | (0.030) | | | (0.027) | (0.029) |
| Historical Mean Temperature | | | | 0.001 | | 0.002 | 0.009 |
| | | | | (0.008) | | (0.009) | (0.005) |
| Temperature Volatility | | | | -0.072 | | -0.094* | -0.112* |
| | | | | (0.050) | | (0.051) | (0.066) |
| Square of Temperature Volatility | | | | 0.005 | | 0.007 | 0.009 |
| | | | | (0.005) | | (0.005) | (0.006) |
| Pre-Colonial Ethnic Controls (p-value joint sign.) | | | | | [0.1863] | [0.3612] | [0.0123] |
| Country Dummies | Y | Y | Y | Y | Y | Y | Y |
| Geo-strategic Controls | Y | Y | Y | Y | Y | Y | Y |
| Cereal Suitability | Y | Y | Y | Y | Y | Y | Y |
| Disease Environment | Y | Y | Y | Y | Y | Y | Y |
| Migratory Distance to Addis Adaba | Y | Y | Y | Y | Y | Y | Y |
| Ecological Diversity | Y | Y | Y | Y | Y | Y | Y |
| Specification for Neolithic Instrument | Square | Square | Square | Square | Square | Square | Linear |
| F (cluster-robust statistics, country level) | 7.964 | 7.022 | 9.018 | 9.828 | 13.234 | 16.798 | 22.534 |
| Observations | 558 | 558 | 558 | 558 | 558 | 558 | 465 |
| R-squared | 0.084 | 0.170 | 0.079 | 0.233 | 0.196 | 0.306 | 0.369 |

Robust standard errrors clustered at the country level in parentheses.*** p<0.01, ** p<0.05, * p<0.1. The unit of observation is a grid cell. The basic set of controls is described in Tables 2 and 3. Absolutelatitude corresponds to the grid's centroid. Suitability for sorghum and millet is the principal component of soil suitability for growingthese two crops (FAO's GAEZ dataset). Temperature data is from the period 1978-2010 and proxy for historical figures (See Ashraf andMichalopoulos, 2013). Temperature volatility represents the intertemporal standard deviation of monthly data. Pre-colonial ethnic controlsare 5 ethnograhic variables representing within- grid 1960-population weighted averages of settelement patterns complexity and subistenceincome shares from hunting, fishing, animal husbandry, and agriculture.
Specification in Column 7 excludes observations from Gabon, Congo, Congo DR, and Angola.

Table 11: Alternative Measure. Historical Proximity to Cities

Dependent Variable: Conflict Prevalence 1989-2010 (fraction of years with at least one conflict event)

| | (1) OLS | (2) IV | (3) IV |
|---|---|---|---|
| Historical Proximity Cities | -0.542** | | |
| | (0.257) | | |
| Local State History 1000 - 1850 CE | | -0.601*** | -0.488*** |
| | | (0.156) | (0.163) |
| Instrument | | Proximity Cities | Proximity Cities and Neolithic Sq. |
| F (cluster-robust statistics, country level) | | 8.674 | 12.825 |
| Hansen J Statistic Over. Test | | | 0.2075 |
| Country Dummies | Y | Y | Y |
| Geo-strategic Controls | Y | Y | Y |
| Cereal Suitability | Y | Y | Y |
| Disease Environment | Y | Y | Y |
| Migratory Distance to Addis Adaba | Y | Y | Y |
| Ecological Diversity | Y | Y | Y |
| Absolute Latitude | Y | Y | Y |
| Suitability for Sorghum and Millet | Y | Y | Y |
| Intertemporal Temperature Volatility | Y | Y | Y |
| Observations | 558 | 558 | 558 |
| R-squared | 0.573 | 0.211 | 0.274 |

Robust standard errors clustered at the country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The unit of observation is a grid cell. All the controls are described in Tables 2, 3, and 10.

Table 12: Conflict, State History, and Weather Shocks -Panel Data Evidence (1989-2010)-

| Dependent Variable: Conflict Prevalence (1 if at least one conflict event in grid-year) | | | | | |
|---|---|---|---|---|---|
| | (1) OLS | (2) IV | (3) IV-FE | (4) IV-FE | (5) IV-FE |
| Shock*Local State History 1000 - 1850 CE | -0.035* | -0.245*** | -0.229*** | -0.223*** | -0.170*** |
| | (0.020) | (0.071) | (0.070) | (0.069) | (0.056) |
| Negative Weather Shock | 0.017*** | 0.051*** | 0.052*** | 0.050*** | 0.038*** |
| | (0.005) | (0.011) | (0.011) | (0.011) | (0.009) |
| Local State History 1000 - 1850 CE | -0.197*** | -0.574*** | | | |
| | (0.041) | (0.160) | | | |
| Lagged Conflict | | | | | 0.275*** |
| | | | | | (0.017) |
| Country Dummies | Y | Y | N | N | N |
| Grid FE | N | N | Y | Y | Y |
| Year FE | Y | Y | Y | Y | Y |
| Geo-strategic Controls | Y | Y | N | N | N |
| Cereal Suitability | Y | Y | N | N | N |
| Disease Environment | Y | Y | N | N | N |
| Migratory Distance to Addis Adaba | Y | Y | N | N | N |
| Ecological Diversity | Y | Y | N | N | N |
| Yearly Precipitation and Temperature Variation | N | N | N | Y | Y |
| F (cluster-robust statistics, country level) | | 24.9 | 40.1 | 40.1 | 39.9 |
| Observations (grids) [years] | 12,276 (558) [22] | 12,276 (558) [22] | 12,276 (558) [22] | 12,276 (558) [22] | 11,718 (558) [21] |

Robust standard errors clustered at the grid level in parentheses.*** $p<0.01$, ** $p<0.05$, * $p<0.1$. The unit of observation is a grid-year.
See appendix for details regarding the weather shock. Yearly precipation variation is the standard deviation of monthly precipation whithin a year.
Yearly temperature variation is the average of monthly deviation from historical mean monthly temperature. The set of controls is described in Tables 2, 3, and 10.

Table 13: Conflict, State History, and Weather Shocks -Panel Data Evidence (1989-2010)-

| Dependent Variable: Conflict Prevalence (1 if at least one conflict event in year) | | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Shock*Local State History 1000 - 1850 CE | -0.177*** | -0.128*** | -0.116** | -0.162*** | -0.151*** | -0.145*** |
| | (0.056) | (0.044) | (0.047) | (0.051) | (0.057) | (0.052) |
| Negative Weather Shock | 0.039*** | 0.061*** | 0.011 | -0.006 | 0.045*** | 0.019 |
| | (0.009) | (0.015) | (0.008) | (0.007) | (0.010) | (0.017) |
| Additional Interacted Control | | Light Density | Cereal Suitability | Agric. Pre-Colonial Dep. | Temp. Volatility | ALL (p-value joint sig.) |
| Shock*Additional Interacted Control | | 0.010*** | 0.065*** | 0.010*** | -0.003* | [0.0000] |
| | | (0.003) | (0.018) | (0.002) | (0.002) | |
| Lagged Conflict | 0.276*** | 0.275*** | 0.274*** | 0.276*** | 0.274*** | 0.276*** |
| | (0.017) | (0.017) | (0.017) | (0.017) | (0.017) | (0.017) |
| Grid FE | Y | Y | Y | Y | Y | Y |
| Year FE | Y | Y | Y | Y | Y | Y |
| F (cluster-robust statistics, country level) | | 24.897 | 40.099 | 40.133 | 39.93 | 38.568 |
| Observations (grids) [years] | 11,718 (558) [21] | 11,718 (558) [21] | 11,718 (558) [21] | 11,718 (558) [21] | 11,718 (558) [21] | 11,718 (558) [21] |

Robust standard errors clustered at the grid level in parentheses.*** p<0.01, ** p<0.05, * p<0.1. The unit of observation is a grid-year. See appendix for details regarding the weather shock and additional interacted controls. Yearly precipation variation is the standard deviation of monthly precipation whithin a year.

Table 14: State Legitimacy and State History

| Dependent Variable: Views on State Institutions Legitimacy | | | | |
|---|---|---|---|---|
| | (1) OLS | (2) IV | (3) IV | (4) IV |
| Local State History 1000 - 1850 CE (District) | 0.0224 | 4.614** | 4.051** | 4.538** |
| | (0.100) | (1.964) | (1.755) | (1.955) |
| Migratory Distance to Addis Adaba | | | -0.0296* | -0.0284* |
| | | | (0.015) | (0.015) |
| Absolute Latitude District | | | -0.0266** | -0.0266** |
| | | | (0.012) | (0.013) |
| Ecological Diversity | | | 0.0780 | 0.0653 |
| | | | (0.138) | (0.143) |
| Suitability for Sorghum and Millet Cultivation | | | 0.126** | 0.136** |
| | | | (0.053) | (0.057) |
| Historical Mean Temperature | | | 0.0117 | 0.0107 |
| | | | (0.008) | (0.009) |
| Temperature Volatility | | | -0.0978 | -0.0999 |
| | | | (0.082) | (0.087) |
| Square of Temperature Volatility | | | 0.0202 | 0.0198 |
| | | | (0.015) | (0.016) |
| F (cluster-robust statistics, district level) | | 8.477 | 9.547 | 8.985 |
| Fuller 1 Point Estimate for Local State History | | 4.581** | 4.025** | 4.506** |
| | | (1.942) | (1.738) | (1.934) |
| Individuals Controls | Y | Y | Y | Y |
| Village Controls | Y | Y | Y | Y |
| District Controls | Y | Y | Y | Y |
| Country FE | Y | Y | Y | Y |
| Ethnic FE | Y | Y | Y | Y |
| Old Conflict within 100km radius | N | N | N | Y |
| Fractionalization District | N | N | N | Y |
| Slave Trade Prevalence District | N | N | N | Y |
| Mean Slave Trade People in District | N | N | N | Y |
| Observations | 22,527 | 22,527 | 22,527 | 22,527 |

Robust standard errors clustered at the district level in parentheses. *** p<0.01, **p<0.05, * p<0.1.
The dependent variable is the first principal component of the responses to 3 questions about the legitimacy of the court decisions, police enforcement, and the tax department (see main text for details). The state history variable is calculated for a buffer of 100km radius from district's centroid. The individual controls are for age, age squared, male indicator, unemployment indicator, education ordered measure (from 0 -no formal education- to 9 -graduate education-), and living condition ordered measure (from 1 - very bad- to 5 -very good-). Village controls are 6 indicators for public good provisions: police station, school, electricity, piped water, sewage, and health clinic. District controls are distance to the capital of the country, infant mortality, per capita light density at nights, and urban indicator. See main text for definition of the additional district controls.

Table 15: State Legitimacy and State History. Internal vs External Norms

| Dependent Variable: Views on State Institutions Legitimacy | | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Local State History 1000 - 1850 CE (District) | 4.614** | | |
| | (1.964) | | |
| Local State History 1000 - 1850 CE (Ethnicity) | | 0.414 | 1.561 |
| | | (1.110) | (1.340) |
| F (cluster-robust statistics, district level) | 8.477 | 8.719 | 8.859 |
| Individuals Controls | Y | Y | Y |
| Village Controls | Y | Y | Y |
| District Controls | Y | Y | N |
| Country FE | Y | Y | N |
| Ethnic FE | Y | N | N |
| District FE | N | N | Y |
| Observations | 22,527 | 22,527 | 22,527 |

Main results are IV estimates. Robust standard errors clustered at the district level in parentheses. *** p<0.01, **p<0.05, * p<0.1 The dependent variable is the first principal componentof the responses to 3 questions about the legitimacy of the court decisions, police enforcement, and the tax department (see main text for details). The state history variable is calculated for a buffer of 100km radius from district's centroid. The set of controls is described in Table 14.

Table 16: Trust and State History

| | Dependent Variable: Trust in | | | |
|---|---|---|---|---|
| | (1) Institutions | (2) Relatives | (3) Compatriots | (4) Politicians |
| Local State History 1000 - 1850 CE (District) | 2.606** | -1.355 | -0.930 | 1.550 |
| | (1.263) | (0.859) | (1.011) | (1.169) |
| F (cluster-robust statistics, district level) | 9.597 | 9.668 | 9.63 | 12.205 |
| Fuller 1 Point Estimate for Local State History | 2.590** | -1.348 | -0.924 | 1.542 |
| | (1.251) | (0.853) | (1.004) | (1.161) |
| OLS Point Estimate for Local State History | 0.202** | -0.255*** | -0.113 | -0.0107 |
| | (0.087) | (0.064) | (0.079) | (0.107) |
| Individuals Controls | Y | Y | Y | Y |
| Village Controls | Y | Y | Y | Y |
| District Controls | Y | Y | Y | Y |
| Migratory Distance to Addis Adaba | Y | Y | Y | Y |
| Absolute Latitude District | Y | Y | Y | Y |
| Ecological Diversity | Y | Y | Y | Y |
| Suitability for Sorghum and Millet Cultivation | Y | Y | Y | Y |
| Intertemporal Temperature Volatility | Y | Y | Y | Y |
| Country FE | Y | Y | Y | Y |
| Observations | 22,525 | 22,453 | 22,155 | 22,383 |

Main results are IV estimates. Robust standard errors clustered at the district level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$ OLS and Fuller 1 point estimates are also reported .The dependent variable is the response to the question How much do you trust in (subject)? I recoded each original answer to a 5 point scale where 1 is “not at all” and 5 is “a lot”. See main text for details.
For the case of trust in institutions the dependent variable is the first principal component of the responses to the trust questions related to police and court of laws. The state history variable is calculated for a buffer of 100km radius from district’s centroid.
The set of controls is described in Table 14

Table 17: Trust in Local Policy Makers, Traditional Leaders, and State History

| | Trust | | Traditional Leader Influence | |
|---|---|---|---|---|
| | Local Councilors | Traditional Leader | Perceived | Prefered |
| | (1) | (2) | (3) | (4) |
| Local State History 1000 - 1850 CE (District) | -0.0278 | 1.787** | 3.307*** | 3.645*** |
| | (0.775) | (0.852) | (1.141) | (1.006) |
| F (cluster-robust statistics, district level) | 23.784 | 23.743 | 22.219 | 22.272 |
| Fuller 1 Point Estimate for Local State History | -0.0274 | 1.783** | 3.299*** | 3.635*** |
| | (0.773) | (0.850) | (1.137) | (1.002) |
| OLS Point Estimate for Local State History | 0.172 | 0.162 | 0.589*** | 0.168* |
| | (0.109) | (0.100) | (0.111) | (0.089) |
| Individuals Controls | Y | Y | Y | Y |
| Village Controls | Y | Y | Y | Y |
| District Controls | Y | Y | Y | Y |
| Country FE | Y | Y | Y | Y |
| Ethnic FE | Y | Y | Y | Y |
| Migratory Distance | Y | Y | Y | Y |
| Absolute Latitude | Y | Y | Y | Y |
| Suitability for Sorghum and Millet | Y | Y | Y | Y |
| Historical Climate Volatility | Y | Y | Y | Y |

Main results are IV estimates. Robust standard errors clustered at the district level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$ OLS and Fuller 1 point estimates are also reported. For the first 2 columns the dependent variable is the response to the questions How much do you trust in local councilors / traditional leaders? The perceived influence variable is based on the question How much influence do traditional leaders currently have in governing your local community? The variable is coded in a 5-point scale from 1 (none) to 5 (great deal of influence). The prefered influence variable is based on the question “Do you think that the amount of influence the traditional leaders have in governing your local community should increase, stay the same, or decrease? THe variable is coded in a 5 point scale from 1 (decrease a lot) to 5 (increase a lot). The state history variable is calculated for a buffer of 100km radius from district.

Table A.1. List of Historical States

| | Date of | |
|---|---|---|
| | Establishment (1) | Unestablishment (2) |
| Dongola (Makuria) | b 1000 | 1314 |
| Alwa | b 1000 | 1500 |
| Kanem Empire | b 1000 | 1387 |
| Kingdom of Ghana | b 1000 | 1235 |
| Pre-imperial Mali | b 1000 | 1230 |
| Pre-imperial Songhai (Gao) | b 1000 | 1340 |
| Siwahili city-states3 | b 1000 | 1500 |
| Mossi States | 1100 | a 1850 |
| Ethiopia (Abyssinia) | 1137 | a 1850 |
| Akan (Bonoman) | 1200 | 1700 |
| Imperial Mali | 1200 | 1600 |
| Buganda | 1300 | a 1850 |
| Songhai Empire | 1340 | 1590 |
| Wollof Empire | 1350 | 1549 |
| Bornu-Kanem | 1387 | a 1850 |
| Kingdom of Congo | 1390 | a 1850 |
| Kingdom of Bamum | 1398 | a 1850 |
| Yoruba (Oyo) | 1400 | a 1850 |
| Nupe Kingdom | 1400 | a 1850 |
| Darfur (Daju-Tunjur until c1600, then Sultanate of Darfur) | 1400 | a 1850 |
| Hausa States | 1400 | 1800 |
| Adal Sultanate | 1415 | 1577 |
| Mwenemutapa (Kingdom of Mutapa) | 1430 | 1760 |
| Benin Empire | 1440 | a 1850 |
| Kingdom of Butua (Butwa) | 1450 | 1683 |
| Kingdom of Rwanda | 1500 | a 1850 |
| Bunyoro-Kitara | 1500 | a 1850 |
| Kingdom of Merina | 1500 | a 1850 |
| Maravi Kingdom | 1500 | 1700 |
| Kingdom of Idah (Igala) | 1500 | a 1850 |
| Kwararafa | 1500 | 1700 |

(continuation) Table A.1. List of Historical States

| | Date of | |
|---|---|---|
| | Establishment (1) | Unestablishment (2) |
| Nkore Kingdom (Ankole) | 1500 | a 1850 |
| Kotoko Kingdom | 1500 | a 1850 |
| Mandara Kingdom (Wandala) | 1500 | a 1850 |
| Funj Sultanate | 1504 | 1821 |
| Kingdom of Bagirmi (Baguirmi Sultanate) | 1522 | a 1850 |
| Kingdom of Ndongo (Angola) | 1530 | 1670 |
| Kingdom of Jolof (Senegal) | 1550 | a 1850 |
| Kingdom of Menabe | 1550 | a 1850 |
| Awsa (Aussa Sultanate since c1730) | 1577 | a 1850 |
| Luba Empire | 1585 | a 1850 |
| Air Sultanate | 1591 | a 1850 |
| Dendi Kingdom | 1591 | a 1850 |
| Teke (Anziku Kigdom) | 1600 | a 1850 |
| Kingdom of Dahomey | 1600 | a 1850 |
| Kuba Kingdom (Bushongo) | 1625 | a 1850 |
| Wadai (Ouaddai Empire) | 1635 | a 1850 |
| Lunda Empire | 1665 | a 1850 |
| Kingdom of Burundi | 1680 | a 1850 |
| Rozwi Empire | 1684 | 1834 |
| Aro trading confederacy | 1690 | a 1850 |
| Kindom of Boina | 1690 | 1808 |
| Ashanti Empire | 1700 | a 1850 |
| Kingdom of Orungu (Gabon) | 1700 | a 1850 |
| Kong Empire | 1710 | a 1850 |
| Bamana Empire (Segu) | 1712 | a 1850 |
| Imamate of Futa Jallon | 1725 | a 1850 |
| Lozi Kingdom | 1750 | a 1850 |
| Mbailundu | 1750 | a 1850 |
| Calabar (Akwa Akpa) | 1750 | a 1850 |
| Kaarta (Baambara in Nioro) | 1753 | a 1850 |

(continuation) Table A.1. List of Historical States

| | Date of | |
|---|---|---|
| | Establishment (1) | Unestablishment (2) |
| Imamate of Futa Toro | 1776 | a 1850 |
| Gibe States | 1780 | a 1850 |
| Xhosa | 1780 | a 1850 |
| Azande Kingdom | 1800 | a 1850 |
| Swaziland (House of Dlamini) | 1800 | a 1850 |
| Ovimbundu (4) | 1800 | a 1850 |
| Yaka (4) | 1800 | a 1850 |
| Borgu States | 1800 | a 1850 |
| Sokoto Caliphate | 1804 | a1850 |
| Zulu Kingdom | 1816 | a1850 |

Note: (1) b stands for before. (2 ) a stands for after. (3) Mogadishu, Mombasa, Gedi, Pate, Lamu, Malindi, Zanzibar, Kilwa, and Sofala. (4) approximate date

Table A.2. Archaelogical Data Used to Construct Time Elapsed Since Neolithic Revolution

| **Site** | **Date BP** | **Latitude** | **Longitude** | **Reference** |
|---|---|---|---|---|
| Nqoma Site, Botswana | 1100 | -18.8 | 21.8 | Manning et al (2011) -1- |
| Shongweni Site, South Africa | 1100 | -29.9 | 30.7 | Manning et al (2011) -2- |
| Chundu Site, Zambia | 1400 | -17.6 | 25.7 | Harlan (1976) |
| Silver Leaves , South Africa | 1700 | -24.0 | 31.0 | Manning et al (2011) -3- |
| Knope Site, Malawi | 1900 | -15.8 | 35.0 | Harlan (1976) |
| Bwambé-Sommet Site, Cameroon | 2300 | 2.9 | 9.9 | Manning et al (2011) |
| Abang Minko'o Site, Cameroon | 2300 | 2.4 | 11.3 | Manning et al (2011) |
| Walaldé, Senegal | 2600 | 16.5 | -14.2 | Manning et al (2011) |
| Gajiganna, Nigeria -Chad Basin- | 2930 | 11.8 | 13.2 | Harlan (1976) |
| Jenné Jenno, Mali | 3000 | 13.9 | -4.6 | Harlan (1976) -4- |
| Ti-n-Akof, Burkina Fasso | 3000 | 14.5 | -0.2 | Manning et al (2011) |
| Deloraine Farm, Kenya | 3200 | -0.4 | 36.1 | Marshall and Hildebrand (2002) |
| Ntereso, Ghana | 3250 | 7.5 | -2.9 | Cowan and Watson (1992) |
| Tichitt site, Mauritania | 3500 | 18.4 | -9.5 | Harlan (1976) |
| Oualata site, Mauritania | 3500 | 17.3 | -7.0 | Harlan (1976) |
| Birimi, Ghana | 3500 | 10.5 | -0.4 | Manning et al (2011) |
| Lower Tilemsi Site, Mali | 3500 | 16.9 | 0.2 | Manning et al (2011) -5- |
| Adrar Bous in the Ténéré Desert in Niger | 4000 | 20.4 | 9.0 | Smith (1995) |
| El Zakiab, Sudan | 5000 | 15.8 | 32.6 | Marshall and Hildebrand (2002) |
| Faiyum, Egypt | 7200 | 29.5 | 30.6 | Shaw (1993) |
| | | | Auxiliary and Highly Speculative Data | |
| South Extreme Edge | 1100 | -34.8 | 20.8 | Based on Shongweni Site |
| Lakaton'I Anja Site, Madagascar | 1600 | -12.3 | 49.4 | Burney et al (2004) -6- |
| Taolambiby Site, Madagascar | 2300 | -23.7 | 44.6 | Burney et al (2004) -7- |
| West Extreme Edge | 3000 | 15.1 | -18.1 | Somalia (Putterman 2006) |
| East Extreme Edge | 3500 | 11.4 | 51.5 | Senegal (Putterman 2006) |

Notes: 1. Coordinates are for Tsodilo Hills. 2. Coordinates are approximate. 3. Northern Province, South Africa. Coordinates are approximate 4. Harlan (1976) mentioned a late domestication (around 2000 BP) but Putterman (2006) put first date for agriculture at 3000 BP based on Harlan's argument. 5. Coordinates are for Karkarinchinkat Site. 6. Evidence of settlement but probably not long-term occupation. Very speculative. 7. This is the earliest evidence of human presence. Very speculative.

Table A.3. Historical Cities Used in Alternative Measure

| **City** | **Source** | **Date** | **City** | **Source** | **Date** |
|---|---|---|---|---|---|
| Saint-Denis | Eggimann | 1800 | Agades | Chandler | 1600-1700 |
| Zimbabwe | Chandler | 1300-1400 | Zagha | Chandler | 1200 |
| Port Louis | Chandler | 1800 | Dongola | Chandler | 1000-1500 |
| Kilwa | Eggimann | 1200-1700 | | | |
| Loanda | Chandler, Eggimann | 1600-1800 | **Non-Sub-Saharan Cities** | | |
| Sao Salvador | Chandler, Eggimann | 1500 | Qus | Chandler | 1000-1400 |
| Loango | Chandler, Eggimann | 1700-1800 | Asyut | Chandler, Eggimann | 1200-1800 |
| Calabar | Eggimann | 1800 | Giza | Eggimann | 1800 |
| Gbara | Chandler | 1600-1800 | Bulaq | Chandler | 1000-1800 |
| Benin | Chandler, Eggimann | 1600-1800 | Tanta | Chandler | 1800 |
| Whydah | Chandler, Eggimann | 1800 | Mahalla el Kubra | Eggimann | 1800 |
| Lagos | Eggimann | 1800 | Damanhour | Eggimann | 1800 |
| Allada | Chandler | 1600 | Alexandria | Chandler, Eggimann | 1000-1800 |
| Kumasi | Chandler, Eggimann | 1700-1800 | Damietta | Chandler, Eggimann | 1200-1800 |
| Abomey | Chandler, Eggimann | 1700-1800 | Marrakech | Chandler, Eggimann | 1200-1800 |
| Bonga | Chandler | 1700-1800 | Tripoli | Chandler | 1500-1800 |
| Ife | Eggimann | 1800 | Azammur | Chandler | 1500 |
| Oyo | Chandler, Eggimann | 1400-1800 | Meknes | Chandler, Eggimann | 1300-1800 |
| Freetown | Eggimann | 1800 | Rabat-Sale | Chandler | 1000-1800 |
| Zaria | Chandler, Eggimann | 1600-1800 | Taza | Chandler | 1500 |
| Massenya | Eggimann | 1600-1800 | Tlemcen | Eggimann | 1300-1800 |
| Kebbi | Chandler, Eggimann | 1800 | Kairwan | Chandler | 1000-1800 |
| Kano | Chandler, Eggimann | 1200-1800 | Oran | Chandler | 1500-1800 |
| Gondar | Chandler, Eggimann | 1700-1800 | Tanger | Eggimann | 1800 |
| Katsina | Chandler | 1600-1800 | Ceuta | Chandler | 1200-1400 |
| Segou | Eggimann | 1700-1800 | Tagaste | Chandler | 1500-1600 |
| Sennar | Chandler, Eggimann | 1600-1800 | Constantine | Eggimann | 1400-1800 |
| Jenne | Chandler | 1300-1600 | Algiers | Eggimann | 1500-1800 |
| Axum | Chandler, Eggimann | 1000-1800 | Bejaia | Eggimann | 1200-1800 |
| Soba | Chandler | 1000-1300 | Tunis | Chandler, Eggimann | 1300-1800 |
| Gao | Chandler | 1000-1500 | Annaba | Eggimann | 1800 |
| Timbuktu | Chandler, Eggimann | 1000-1800 | | | |