THE EFFECT OF THE TSETSE FLY ON AFRICAN DEVELOPMENT

Marcella Alsan*†

June 17, 2014

Abstract

The TseTse fly is unique to Africa and transmits a parasite harmful to humans and lethal to livestock. This paper tests the hypothesis that the TseTse reduced the ability of Africans to generate an agricultural surplus historically. Ethnic groups inhabiting TseTse-suitable areas were less likely to use domesticated animals and the plow, less likely to be politically centralized and had a lower population density. These correlations are not found in the Tropics outside of Africa, where the fly does not exist. The evidence suggests current economic performance is affected by the TseTse through the channel of precolonial political centralization.

JEL Codes: I15, N57, O57

*I am grateful to David Cutler, Paul Farmer, Claudia Goldin, Michael Kremer and Nathan Nunn for encouragement and detailed feedback. For many helpful comments and suggestions, I thank three anonymous Referees, Ran Abramitzky, Alberto Alesina, Robert Bates, Jay Bhattacharya, Hoyt Bleakley, Melissa Dell, Rebecca Diamond, Oeindrila Dube, Stanley Engerman, James Feigenbaum, James Fenske, Daniel Fetter, Erica Field, Edward Glaeser, Joshua Gottlieb, Richard Hornbeck, Lawrence Katz, Grant Miller, Philip Osafo-Kwaako, Orlando Patterson, James Robinson, Dana Rotz and participants at seminars at AEA, BU School of Public Health, Dartmouth, Georgetown, Harvard, J-PAL History of Poverty, LSE, Massachusetts General Hospital, NBER DAE, NEUDC 2012, OSU, Oxford, SCU, Stanford, WGAPE, and UCLA-Anderson. For assistance with FAO data and GIS I thank Giuliano Cecchi, Rafaelle Mattioli, William Wint and Jeff Blossom. Luca Maini and Zach Levin provided excellent research assistance. Funding provided by NIH T32 AI 007433, Harvard Global Health Institute *Fellowship for Neglected Tropical Diseases* and the Harvard Department of Economics *Lab for Economic Analysis and Policy*. All errors are my own. I declare that I have no relevant or material financial interests that relate to the research described in the paper.

†Correspondence: Stanford University 117 Encina Commons, Room186 Stanford CA, 94305, U.S.A. (e-mail: malsan@stanford.edu; website: http://healthpolicy.stanford.edu/people/marcella_alsan).

Economists have become increasingly interested in exploring the deep historical roots of comparative economic development.[1] While much of the empirical research with an African focus has concerned the legacy of the colonial powers (Nunn, 2008; Nunn and Wantchekon, 2011; Wantchekon, Novta, and Klasnja, 2013; Jedwab and Moradi, 2013) an important finding to emerge is that political centralization prior to colonization can explain a nontrivial amount of variation in African economic development today (Fenske, 2013; Gennaioli and Rainer, 2007; Michalopoulos and Papaioannou, 2013, 2014). This finding is consistent with a view widely held among African scholars that the continent's historically low population density, or relative land abundance, has played a pivotal role in shaping its development (Hopkins, 1973; Iliffe, 1995; Fenske, 2013; Herbst, 2000).[2] This view maintains abundant arable land weakened state development by hindering the ability to broadcast power over sparsely settled territories. Moreover, since labor was scarce and shadow wages high, the wage labor market was anemic with household and slave labor used instead.

But why was land in historical Africa relatively abundant? The anthropologist Jack Goody (1971) argued agricultural technologies used to improve food production in much of the rest of the Old World were slow to diffuse in Africa with adverse consequences for development.[3] However, an important and unique feature of African ecology that could have affected its ability to adopt technologies, agricultural productivity, population density and

[1]For reviews on the historical origins of comparative development, see Spolare and Wacziarg (2013) and Nunn (2009). The literature has focused on cultural, genetic, biogeographical, human capital and institutional factors that may influence the proximate determinants of wealth. Recent empirical work has demonstrated genetics may affect diffusion of technology or social cohesion (Wacziarg and Spolare, 2009; Ashraf and Galor, 2013), thus influencing economic outcomes. In addition, Galor and Michalopoulos (2012) develop a theory of inequality based on the natural selection for entrepreneurial traits. Olsson and Hibbs (2005) and Ashraf and Michalopoulos (2013) investigate how biogeographical features and climatic variation influenced the timing of the Neolithic revolution with long run consequences for growth. Seminal contributions to the role of human capital in long run development include Putterman and Weil (2010) and Glaeser et al. (2004). A large literature has emerged describing the linkages between the historical institutional environment and variation in development outcomes today (Acemoglu, Johnson, and Robinson, 2001; Banerjee and Iyer, 2005; Dell, 2010; Iyer, 2010).

[2]Reasons sometimes given for the continued relevance of precolonial African institutions include the relatively late and brief period of colonization and the strength of such institutions.

[3]In related work, Comin, Easterly, and Gong (2010) find evidence that the level of technology five hundred years ago predicts differences in income and technology in the present day. Table I of their paper describes mechanisms to explain this finding, such as spillovers to other sectors, complementarities between technologies and feedback from technology to science.

institutional development that has yet to be examined empirically is the TseTse fly.

Economists, historians and biologists have debated the role of the TseTse on African development. By circumscribing the use of domesticated animals as a source of draft power, and precluding the adoption of technologies complementary to draft power, the TseTse has been hypothesized to have hindered the ability of Africans to generate an agricultural surplus and easily transport goods overland. The entomologist T.A.M. Nash (1969, p. 31) writes, "It seems reasonable to suppose that for hundreds of years tsetse dictated that the economy of the African should be based on the hoe and the head-load, a dictatorship which he is now being freed by the petrol engine and the railway locomotive." Others have expressed skepticism that the TseTse could explain why African technology lagged behind Eurasia (Chaves, Engerman, and Robinson, 2013).

This paper is the first to investigate whether the TseTse fly affected Africa's precolonial agricultural technologies, patterns of subsistence, population density and institutions. The TseTse (*Glossina spp.*) is only found in Africa.[4] The fly feeds strictly on vertebrate blood and transmits *Trypanosomiasis*, a parasite causing sleeping sickness in humans and nagana in domesticated animals.[5] Livestock tend to be more affected than people since there are more types of trypanosomes that can infect them and the fly preferentially feeds on nonhuman animal hosts (Leak, 1999; Owaga, 1985; Vale, Flint, and Hall, 1986; Vale, 1974).[6]

[4]The TseTse is a prehistoric species that originated about 100 million years ago (Krafsur, 2009). Climate changes, continental separation and glaciations are believed to have isolated the TseTse in Africa during the late Miocene or Pliocene epoch (Lambrecht, 1964, p. 2).

[5]The TseTse is unique to Africa and TseTse-transmitted *Trypanosomiasis* is the subject of this paper. *T. cruzi* is in South and Central America and causes Chagas disease in humans. Three forms of trypanosomes causing disease in domesticated animals (*T. equiperdum, T. evansi* and *T. vivax)* have spread beyond Africa. *T. equiperdum* is a sexually transmitted infection of horses and will not be discussed further. *T. evansi* and *T. vivax* are believed to have been spread during the process of European colonization (FAO, 1998, p. 137). Animal *Trypanosomiasis* outside of Africa was not as virulent as within Africa since it lacked a specialized vector for transmission (e.g. the TseTse) and a large reservoir population of immune wild game. Further background on the biology is provided in Section III.1 and Appendix C.

[6]European explorers were convinced that the TseTse did not harm humans: "During my hunting excursions along the Teoge, I encountered the most extraordinary of insects, the Tsetse. Among the several scourges to which the traveller is subjected in the South African wilderness, one of the greatest is this insect; not, it is true, as to the wayfarer's own person, for he himself escapes almost unscathed, but as regards the horses and cattle" (Anderson and Fragmont, 1857, p. 488-489). Livingstone (1857, p. 80-81) remarked, "A most remarkable feature in the bite of the Tsetse fly is its perfect harmlessness in man and wild animals."

Crucial for identification of the impact of the TseTse is its specific, non-monotonic temperature and humidity requirements for viability. These physiological relationships have been elucidated through controlled laboratory experimentation on the fly (Bursell, 1960; Jackson, 1949; Mellanby, 1937; Rajagopal and Bursell, 1965; Terblanche et al., 2008). The exact functional forms relating TseTse birth and death rates to climate are derived from the experimental data. Using insect population growth modeling, gridded climate data and geospatial software, the potential steady state TseTse population can be calculated. The TseTse suitability index (TSI) is the standardized value (Z-score) of this steady state population. The TSI is then linked to precolonial anthropological observations on African agricultural practices, institutions and urbanization. A detailed description of the ethnographic data is provided below. The regressions compare highly TseTse suitable areas to less TseTse suitable areas within Africa controlling separately for the individual factors in the TSI and their first-order interaction (robustness tests also include higher order terms).[7]

The TseTse is estimated to have had substantial effects on precolonial Africa: a one standard deviation increase in the TSI is associated with a 23 percentage point decrease in the likelihood an African ethnic group had large domesticated animals, a nine percentage point decrease in intensive cultivation and a six percentage point reduction in plow use. A one standard deviation increase in the TSI is correlated with a 53 percent reduction in historical population density. Motivated by the land abundance literature, two institutions are explored in this paper: political centralization and indigenous slavery. A one standard deviation increase in the TSI is associated with a ten percentage point increase in the likelihood an ethnic group used slaves and an eight percentage point decrease in the probability it was centralized.

The main threat to the validity of the analysis is that the TSI may be spuriously capturing the latent negative effects of the Tropics. In addition to controlling for a rich set of geographic

[7] The reduced form relationship between the TSI and precolonial outcomes is emphasized since a reliable historical map of TseTse is unavailable. Focusing on the TSI as opposed to the observed fly distribution also potentially addresses endogeneity concerns related to more advanced, centralized ethnic groups being better able to control the fly (see discussion in footnote 16).

and climate variables, two additional steps are taken. First, the TSI is constructed for ethnic groups located in the Tropics outside of Africa. The TSI predicts significant, detrimental effects on development only within Tropical Africa. Second, the formulas for the TSI are perturbed in quantitatively slight but physiologically important ways. These "perturbed" indices do not have the same correlations with agricultural and institutional outcomes within Africa. One limitation of the study is that data on the distribution of the TseTse at a time before the precolonial period is unavailable. If such data were available, and if the observed historical TseTse measure had more explanatory power than the TSI in a regression that included both terms, this would further add to the validity of the index. Absent such data, it is difficult to completely exclude the possibility that the TSI is capturing some generic aspect of the Tropics. However, the evidence presented is consistent with the idea that the TSI within Africa is mainly capturing the effect of the TseTse.

This paper concerns institutional origins and thus the fundamental determinants of prosperity. Simulating historical African development under a lower burden of TseTse indicates that Africa would have been characterized by modest increases in intensive cultivation and political centralization. This finding is consistent with archeological evidence of more advanced civilizations supported by intensive agricultural systems in places where the fly could not survive, such as Great Zimbabwe. The TSI has a negative correlation with current economic outcomes as measured by satellite light density or the observed cattle distribution in Africa. The modern analysis is performed at the district level and is robust to including country fixed effects. The evidence suggests that the relationship between the TSI and satellite lights is driven by the TseTse's effect on shaping historical institutions, particularly political centralization. In contrast, the correlation between the TSI and the current distribution of cattle is not significantly affected by the addition of controls for historical institutions. This set of findings provides evidence in favor of the Engerman and Sokoloff (2000) point of view on how endowments, such as the disease environment, may shape institutions and thereby have persistent effects on economic development while underscoring there may be a direct influence of TseTse-transmitted *Trypanosomiasis* on animal husbandry in Africa today.

The rest of the paper is organized as follows. The next section provides an overview of the historical importance of livestock and discusses how TseTse-transmitted *Trypanosomiasis* could have influenced subsistence strategies and state development. Section II describes the data construction and empirical framework. Section III presents the main historical results, the results of the placebo test and counterfactual simulation. Section IV examines the effect of the TseTse fly on modern economic development and Section V concludes.

I Historical Background

I.1 The Role of Livestock in Development

Communicable disease has often been explored as a cause of Africa's underdevelopment (Bloom and Sachs, 1998; Gallup and Sachs, 2001; Sachs and Malaney, 2002). Although the literature has investigated the role of human pathogens on economic performance, it is largely silent on the impact of veterinary disease.[8] This is peculiar given the role livestock played in agriculture and as a form of transport throughout history. Prior to mechanization, domesticated animals were an important input into the agricultural production process. Livestock improved yields by providing manure for fertilizer, made use of leguminous fodder and served as a source of draft power. The agricultural revolution in England relied upon domesticated animals (Allen, 1999; Overton, 1996). The process was summarized by a farmer in 1795: "No dung–no turnips–no bullocks–no barley–no clover nor...wheat" [quoted in Overton, 1996].

The adoption of domesticated animals and associated technologies also affected culture.[9]

[8] Livingstone (1857) mentions the TseTse 67 times in his work, *Missionary Travels and Researches in South Africa*; by contrast, malaria is mentioned six times.

[9] Some have also viewed large domesticated animals, particularly the horse, as crucial for conquest (Maudlin, 2006) and power consolidation. For example, the southward expansion of ethnic groups using cavalry in Northern Nigeria (e.g., Nupe and Oyo) was believed to have been limited by the TseTse (Law, 1977). Law (1977, p. 198) writes, "Oyo operations against Dahomey were restricted to brief raids, as the cavalry could not operate during the rainy season (presumably because of the danger from trypanosomiasis) and were hampered by the problem of securing fodder for the horses. Consequently, although they could overrun the country and defeat any Dahomian army which stood and fought, they could not effect a complete and permanent conquest, so that in the end, Dahomey had to be left autonomous and tributary."

Alesina, Giuliano, and Nunn (2013) show that historical plow use is predictive of present-day gender norms, suggesting the comparative advantage men have in upper body strength led to a reduction in the role women played in farming. Technical progress in agricultural techniques (plow use, harnessing) led to higher returns from animal power. Pierre Bonnassie (2009, p. 40) argues these technical changes aided the decline of slavery in Western Europe: "on the one hand, water power, and on the other, an increased return accrued from animal labour (the return was quintupled in the case of the horse) took the place of human energy (represented by slave labour) in the most laborious and common of tasks."[10]

Although livestock disease has beleaguered farmers worldwide, African Animal *Trypanosomiasis* was particularly detrimental. Nagana infects all forms of ungulates, whereas most other pathogens (i.e., glanders, rinderpest, footroot) have a predilection for a particular species (Brown and Gilfoyle, 2010).[11] Rapid antigenic variation, the switching of proteins on the surface of the trypanosome so that a host cannot recognize the infectious agent, thwarts the animal's humoral immune response (Borst and Rudenko, 1994). A parasite that quickly kills its host will itself become extinct. African Animal *Trypanosomiasis* was able to survive because wild game served as a reservoir population in which they circulated. In general, for an acutely lethal parasitic disease to thrive it requires a host population that is immune (such as the big game of Africa), and an efficient vector (such as the TseTse) which has coevolved for its transmission (part of the life cycle for the parasite is completed in the fly).[12] Evidence of the hardship TseTse posed to the keeping of livestock comes from the

[10] The price of a horse imported into the Oyo empire was at least twice the price of a slave (Law, 1977, p. 185).

[11] Anthrax and Brucella are as broad as nagana in infectious scope, though Brucella is not fatal to adult animals and immunity to Anthrax has been shown to occur naturally in livestock (Turnbull et al., 1992). Vaccines exist today for both Brucella and Anthrax but not for *Trypanosomiasis*.

[12] Several different features lead to immunity in wild game including the presence of a trypanosome lethal factor (Mulla and Rickman, 1988). After 8,000 years, certain breeds of cattle ("trypanotolerant" breeds such as the N'dama) have been noted to have reduced susceptibility to *Trypanosomiasis*, though will succumb with a high enough parasite load (Murray, Trail, and D'Ieteren, 1990). As mentioned in footnote 5, there are animal trypanosomes outside of Africa, specifically *T. vivax* and *T. evansi* (also known as surra). Surra is a disease primarily affecting camels and horses.

The impact of animal *Trypanosomiasis* (*T. vivax* and *T. evansi)* outside of Africa differs from that within Africa in that there are no specialized vectors for transmission and it is therefore much less efficient. Specifically, biting flies act as needles that mechanically transmit the parasite from host to host. However, in colonial record. Commissioner H.H. Johnston (1894) described the TseTse as the "greatest curse" nature laid upon Africa and remarked the "value of the country would be centupled" in the absence of the fly. Early colonists often resorted to the less advanced technologies that characterized the region (such as human porterage). The problem was especially acute given that the TseTse's ecological niche was in fertile areas in Africa. As can be seen in Figure I, the suitability of land for agriculture (see Nunn and Qian (2011) for a description) and the TSI are positively correlated. It is remarkable that no African ethnic group historically occupied a place that was both inhospitable for the fly and yet highly suitable for agriculture.

I.2 Subsistence Strategies and the TseTse fly

In addition to the role of large domesticated animals in intensive agriculture and transportation, the TseTse fly has been purported to have specific effects on the patterns of subsistence within Africa. This is important because the subsistence strategy a group adopts has ramifications for its population size and social structure. In particular, the TseTse fly is believed to have inhibited the spread of animal husbandry and pastoralism in Africa. Jared Diamond (1997, p. 186) writes that "the spread southward of Fertile Crescent domestic animals through Africa was stopped or slowed by climate and disease, especially by trypanosome diseases carried by tsetse flies." Archeological evidence points to a stalled pattern of large domesticated animal diffusion in African prehistory, in contrast to the prolific spread of ceramics, and suggests that *Trypanosomiasis* may provide an explanation for the lag (see

order for transmission to be successful the interval between fly feeds has to be very short as the trypanosome dies when blood dries. Such a short interval between feeds typically only occurs when flies are interrupted while taking a blood meal. The TseTse can not only mechanically transmit but can also cyclically transmit, which means the trypanosome parasite actually multiplies in the TseTse gut and infective forms are stored in the salivary gland. The flies are infected for the rest of their lives. Whenever the TseTse takes a blood meal, the fly emits salivary anticoagulant that helps it feed and the trypanosomes are injected into the host blood stream along with the anticoagulant. This translates into a much higher transmission rate (rate at which an initial infected animal/human can give rise to new infections) and more subspecies of trypanosomes that have developed to survive within the TseTse and can harm many different species of large domesticated animals within Africa (FAO, 1998, p. 6). For example, in Africa there are not only the trypanosomes mentioned above but also those that rely on the TseTse to complete part of their replication cycle: *T. congolense, T. simiae, T. godfreyi, T. brucei brucei, T. brucei rhodesiense* and *T. brucei gambiense.* Finally in Africa, wild game are immune to the trypanosome parasites and thus serve as a reservoir of the disease. This does not seem to be the case outside of Africa (Luckins and Dwinger, 2004; FAO, 1998, p.140), meaning that the disease cannot be as acutely lethal in domesticated animals or else it will not be propagated.

review by Gifford-Gonzalez (2000)). The migratory patterns of pastoralist groups living on the edge of the Sahel were heavily influenced by the seasonal expansion and contraction of TseTse fly belts (Ingold, 1987, p. 182). Furthermore, the relative survival advantage of wildlife over domesticated animals might have encouraged the persistence of hunting and gathering as opposed to the adoption of more advanced food production strategies that relied on animal husbandry.[13]

I.3 Political Centralization and the TseTse fly

Precolonial political centralization has previously been shown to be positively correlated with modern African development (Gennaioli and Rainer, 2007; Michalopoulos and Papaioannou, 2013, 2014), though the mechanisms which gave rise to heterogeneity in African political centralization are an active area of research. There are several ways the TseTse fly could have affected centralization; and they are related to the purported role of the TseTse in shaping patterns of subsistence. If ethnic groups in TseTse-infested areas were more likely to rely on hunting and gathering (e.g., foraging), this would imply a fairly mobile society with almost no occupational choice. Foraging societies do not establish permanent settlements, rather they function as isolated bands without authority above the local level. Indeed, as populations rise within the band, the strategy of many foraging groups is to fission into smaller subgroups so as to avoid conflict over resources.

Jeffrey Herbst (2000) in his book, *States and Power in Africa: Comparative Lessons in Authority and Control,* highlights two major factors as impediments to state-building and extension of authority in Africa–low population density and high transport costs.[14] The

[13] I thank an anonymous referee for pointing this out.

[14] Herbst (2000, p. 11) writes: "The fundamental problem facing state-builders in Africa–be they pre-colonial kings, colonial governors, or presidents in the independent era–has been to project authority over inhospitable territories that contain relatively low densities of people." John Iliffe (1995, p. 70) also writes, "In the West African Savannah, underpopulation was the chief obstacle to state formation." Robert Bates (1983, p. 35) demonstrated a correlation between population density and state centralization. However, Robinson and Osafo-Kwaako (2013) have recently called into question the utility of Eurasian models of state formation for explaining state centralization in historical Africa. Their paper draws upon data from the Standard Cross Cultural Sample and finds that, within sub-Saharan Africa, there is a statistically positive but insignificant effect of population density on political centralization.

TseTse fly could have exerted influence on both factors. Extensive farming, similar to foraging, supports a limited number of individuals per hectare of land.[15] Extensive farming would be advantageous in TseTse-infested areas as burning brush would scare away wildlife that were attractants for the fly. Low population densities, in general, are favored if communicable disease (e.g. sleeping sickness) is highly prevalent since it reduces the probability of pathogen transmission. Transport costs would have been affected by the TseTse since sending messages, carrying goods or military transport over land would have been hampered by the lack of large domesticated animals.

The TseTse not only encouraged certain forms of subsistence agriculture, but possibly discouraged others. As discussed above, precolonial intensive agriculture was characterized by manure for fertilizer and plowing to aerate and loosen the soil. Intensive farming brings with it numerous social changes, such as permanent settlements due to the immobility of the food source and surplus crops which can be used to support a large, non-agricultural workforce, including a ruling class. Surplus crops and people also form the tax base to support a central authority. In summary, through its effect on subsistence patterns, which influence settlement structure, population density, occupational specialization, fiscal capacity as well as by increasing transport costs, the TseTse could have had a detrimental impact on precolonial centralization.

II Data Construction and Empirical Framework

The next subsection describes the creation of the TSI, which is used for identification throughout the paper. This is followed by a description of the ethnographic data and the main estimating equation.

[15] Shifting agriculture can produce up to 0.2 quintals of grain per hectare sustaining a maximum of 10 inhabitants per square kilometer (Mazoyer, Roudart, and Membrez, 2006, p. 116) whereas intensively farming the land using animal powered technologies such as the plow and fertilizer can support a maximum of 55 inhabitants per square kilometer (Mazoyer et al., 2006, p. 282).

II.1 Population Growth Model of the TseTse fly

There are two main reasons to develop the TSI. First, a reliable map of the precolonial TseTse distribution is not available. Using a 20th century map of the TseTse distribution may lead to skewed results since climate change may have altered the location of the fly.[16] The TSI can be constructed with historical climate data thereby mitigating this concern. Second, a measure of potential as opposed to observed TseTse purges the estimates of bias arising from states with stronger institutions being able to better control the fly. Let B represent the TseTse birth rate, which is temperature dependent, and M represent the mortality of adult flies from dessication, which depends on the saturation deficit: a combination of both humidity and temperature.[17] In the absence of intraspecies competition, the TseTse has a constant, net nonnegative growth rate , $\frac{\dot{N}}{N} = \Lambda$, defined as: $\Lambda = B(t) - M(t,h)$. The equations for B and M are found by fitting curves to the data points from laboratory experiments, as in Figure II panels (A) and (B).

Substituting in climate data from Africa, B typically exceeds M and there is no steady state. A second form of mortality attributable to intraspecies competition is introduced (density dependent mortality): $\Delta = \phi\,(N)^{\psi}$. The steady state equilibrium population is therefore: $N^* = \left(\frac{\Lambda}{\phi}\right)^{\frac{1}{\psi}}$. Intuitively, the steady state population will be bigger the larger the difference between the birth and death rate.[18] The historical analysis uses climate data

[16] Moore and Messina (2010) model how climate change has affected the TseTse distribution within Kenya. Paleoclimatic data from tree cores and ice rings (Mann et al., 2008) demonstrate changes in Africa's temperature, especially in the latter half of the 20th century (Appendix Figure A.I) which might affect fly distribution. Note that the average temperature is fairly constant over the 16th to mid-20th century.

A modern map of TseTse, produced by Ford and Katondo (1977), provides a binary measure of the fly and is heavily based on colonial surveillance. Several historians document the use of sleeping sickness surveillance and control measures by colonial powers to expropriate land and control the indigenous population. In *Lords of the Fly: Sleeping Sickness Control in British East Africa*, Hoppe (2003, p. 154) writes: "Local people did not necessarily think that the colonial agents were lying about sleeping sickness and tsetse. However, most communities were not experiencing a sleeping sickness epidemic. They experienced medical examinations, the British expropriation of land and labor, the loss of investments in homes and farms, and denied access to resources while colonial officials had free access." Motivated by such historical accounts, in Appendix Table A.IX, a subsample of the Ford and Katondo map is used to implement two-stage least-squares. The sample is limited to ethnic groups in countries that have TseTse surveillance data from sources other than the colonial powers.

[17] Further details on insect physiology can be found in Appendix C, see also Schowalter (2011).

[18] Implementing this model requires a choice of parameter values. Fortunately, May et al. (1974) have

from the 20th century reanalysis version 2 (20CRv2) (Compo, Whitaker, and Sardeshmukh, 2011). The average of daily mean temperature and relative humidity for the first year of available data (1871) are used. The TSI (the normalized N^*) is computed for each ethnic group. A 3-dimensional figure of the relationship between climate and the steady state population is shown in Figure II panel (C). The TSI joined with a continental map of African ethnic groups is shown in Figure III panel (A), and the FAO suitability for rainfed agriculture is shown in panel (B).

II.2 Estimating Equation

Motivated by the discussion above, the empirical analysis focuses on how the TseTse affected agricultural practices, urbanization, institutions and subsistence strategies. Figure IV panels (A) and (B) provide a visual representation of the reduced form relationship between many of these outcomes and the TSI by plotting their average by TSI quartile. The main estimating equation is presented below and further explores this within-Africa heterogeneity:

$$Outcome_j = \alpha + \delta TSI_j + \mathbf{X}_j'\Omega + \varepsilon_j, \tag{1}$$

where $Outcome_j$ represents one of the precolonial outcomes associated with ethnic group j. The vector X_j' denotes the set of climate variables comprising the TSI (i.e., temperature and humidity) and their first order interaction as well as the proportion of land area in the Tropics. The identification strategy can be seen graphically as the difference between Figure II panel (C), the nonlinear interaction of temperature and humidity that comprises the TSI, and panel (D), the first order interaction between temperature and humidity. It should be emphasized that similar point estimates are obtained when including higher order terms in climate (see Appendix Table A.VIII column (9)).

The *Ethnographic Atlas* is a worldwide database that includes the historical features of

studied the stability conditions of a similar model, which rest upon ψ, and have determined the steady state oscillates about its equilibrium for parameter values $2 > \psi > 0$. For the TseTse Suitability Index, $\phi = 0.025$ and $\psi = 1.25$, though a sensitivity analysis (Appendix Table A.III) shows results are robust over the entire range. The particular value of ψ was chosen based on the results of experimentation with *Aedes aegypti* ($\psi = 0.922 \pm 0.47$) (Legros et al., 2009, p.14).

1267 ethnic groups, of which 533 are in Africa and is the main dataset used in this paper (Murdock, 1967).[19] The data in the *Atlas* are cross-sectional and are meant to capture the characteristics of ethnic groups prior to European settlement. The observations are coded by Murdock and summarize field work performed by anthropologists primarily during the 19th and early 20th century. Murdock's book *Africa: Its Peoples and Their Culture History*, provides estimates of the population of many African ethnic groups and will be used as a measure of population density. The location of cities with over 20,000 inhabitants in the year 1800 is included as an alternative measure of urbanization (Chandler, 1987). The outcome data from the *Atlas* are spatially combined with *Murdock's Tribal Map of African Ethnicities* (Murdock, 1959b), which includes the location of 843 ethnic group areas. Ethnic groups from the *Ethnographic Atlas* are joined to the *Map* using the procedure described by Fenske (2013).[20] There is no map of boundaries for ethnic groups outside of Africa, a point returned to below.

X_j' includes other plausibly exogenous controls. Absolute latitude and proportion of land area in the Tropics control for different agro-ecological zones. Irrigation, trade and fishing, as well as an alternative form of transportation, would have been influenced by waterways, and access to such is also included in the analysis. Longitude captures differences in the Eastern and Western parts of the continent. Soil fertility and other environmental conditions might also affect aggregate agricultural productivity, and thus a summary measure of agricultural suitability, developed by the Food and Agricultural Organization (FAO), is included as a covariate. *Plasmodium Falciparum* has been singled out as an obstacle to growth in Africa due to its affect on human health (Bloom and Sachs, 1998; Gallup and Sachs, 2001). Others have argued that genetic and acquired immunity confer a certain

[19] Although there are 1267 ethnic groups originally, Chilcotin and Tokelau are entered twice with slightly different outcomes in the database. Since the placebo analysis compares Tropical Africa to the rest of the Tropics, only Tokelau is dropped (a total of two observations) due to this duplication.

[20] Since there is not a one-to-one match for all observations in the *Atlas* and those in the *Map*, Fenske developed an algorithm to join unmatched ethnic groups based on an alternative name, supergroup or location. 523 mainland-associated ethnic groups are matched this way–one outlier observation is dropped bringing the sample to 522. The file can be found in the Web Appendix to Fenske (2013): `http://www.jamesfenske.com/`.

tolerance to severe malaria in all but the very young and thus its historical impact may be overstated (Weil and Depetris-Chauvin, 2013). The malaria ecology index by Kiszweski et al. (2004) is used to approximate the prevalence of different forms of malaria. Altitude reflects the privileged position of the African highlands–relatively free from insect vectors, easier to defend and with ample precipitation. A correlation matrix between the TSI and these geographic and climate features is in Appendix Figure A.II.

It is unlikely that each ethnic group can be thought of as an independent observation, given that many share a common cultural ancestry. This will not be entirely captured by standard errors that control for spatial correlation, since migratory patterns (such as the Bantu expansion) may place groups far apart despite a common lineage. The most popular way to deal with spatial correlation in the data is to use Conley's covariance matrix, a weighted average of spatial autocovariances, with the weights declining linearly to zero until a pre-specified cut point is reached (Conley, 1999). This would be inadequate in a setting where spatial and genealogical correlation are both at work. Errors are therefore clustered at the level of cultural provinces, which are groupings devised by Murdock. Reconstruction of the attributes of ethnic groups is based on written and archeological records, linguistic evidence, common cultigens and the conservatism of certain features in societal organization (Murdock, 1959a, p. 42). These provinces capture both spatial and genealogical correlation. The sample of ethnic groups included in the analysis is shown in Appendix Figure A.III, with shading to represent the 44 different clusters.[21]

III Results

III.1 Historical Results

The TseTse prefer to feed on nonhuman animals and, unlike wild game, livestock are not immune. The presence of large domesticated animals at the ethnic group level is coded as a binary variable equal to one if the ethnic group used bovines, camelids, or equines.

[21] Note that 478 of 522 ethnicities are represented in the map and 44 ethnicities are joined to one of the represented 478 using the algorithm developed by Fenske (2013). Conley standard errors and multiway clustering are reported as robustness checks (see Table III).

Each cell in Table I reports the coefficient on the TSI from a separate regression using Equation (1). Table I column (1) controls only for the climate variables in the TSI and the proportion of land area in the Tropics. A one standard deviation increase in the TSI decreases the probability of an ethnic group possessing large domesticated animals by 21.6 percentage points. Moving across the columns, geographic and malaria controls are added. The point estimate remains stable, reducing concern for selection on unobservables (Altonji, Elder, and Taber, 2005). The preferred specification is reported in Table I column (4) and includes geographic, climate and malaria controls. A one standard deviation increase in the TSI is associated with a statistically significant, 23.1 percentage point decrease in the probability an ethnic group possesses large domesticated animals which is one-third of the sample mean.[22] African agricultural technology may not have advanced in many places because of the TseTse. First, without draft animals, a plow is hard to use. Second, shifting agriculture is a labor saving technique, since the number of hours necessary to burn a forest and remove the stumps and rocks is far fewer then what would be needed to continuously farm the same plot (Boserup, 1966). Third, intensive agriculture requires fertilizer, since the soil is rapidly depleted of nutrients by repeated cultivation. Without animal dung, farmers would need the long fallow of shifting cultivation to allow time for the soil to replenish. The correlation between TseTse suitability and intensive cultivation is negative: a one standard deviation increase in the TSI decreases intensive agriculture by nine percentage points, which is about one-third of the sample mean. (The complement of intensive cultivation is shifting or no agriculture). A one standard deviation increase in the TSI is also associated with a six percentage point decrease in plow use. Agricultural practices and cultural norms are often intertwined. For example, Alesina, Giuliano, and Nunn (2013) find evidence in support of the Boserupian hypothesis that historical plow use led to a gendered division of labor in agriculture. Table I row (4) reports the results of a (0,1) variable indicating whether females

[22] Livestock might not have been able to graze on the best pasturage, or otherwise used as productively in the presence of the TseTse. Milking livestock is negatively associated with TseTse suitability, suggesting an effect on the intensive margin (Appendix Table A.IV).

perform the majority of agricultural tasks. A one standard deviation increase in the TSI is associated with a statistically significant 20.6 percentage point increase in the probability that females are participating heavily in agriculture.

Both an agricultural surplus and transportation networks are important for urbanization and political centralization (Bairoch, 1988). A one standard deviation increase in the TSI is associated with a statistically significant 52.5 percent reduction in population density.[23] Consistent with the above discussion, there is a negative correlation between the TseTse and centralized states. Political centralization is constructed from the variable entitled "jurisdictional hierarchy beyond the local authority" in the *Ethnographic Atlas*. Jurisdictional hierarchy is coded so that it equals zero for groups lacking any form of centralized state, one for petty chiefdoms, two for large paramount chiefdoms/petty states and three or four for large states. Following the literature, an ethnic group is considered politically centralized if it has a value greater than one for the jurisdictional hierarchy variable. A one standard deviation increase in the TSI decreases the probability of an ethnic group being classified as centralized by 7.5 percentage points. Nieboer (1900) and Domar (1970) both observed that a high land-labor ratio (or low population density) was positively correlated with labor coercion in the historical record. The entomologist J.P. Glasgow (1963, p. 3) conjectured that the practice of indigenous slavery and the presence of the TseTse were related: "Nearer the equator the use of draught or pack animals was impossible, and such trade as occurred depended on transport by human carriers. This circumstance, we may suppose, encouraged the growth of slavery." Using the ethnographic data there is empirical support for a positive correlation between the TseTse and the practice of indigenous slavery at the ethnic group level. A one standard deviation increase in the TSI is associated with a statistically significant 10.1 percentage point increase in the probability an African ethnic group used slave labor.[24]

[23] Results reported in Appendix Table A.IV demonstrates a similar negative correlation between urbanization and the TSI when the outcome variable is a binary indicator for having a city in 1800 with over 20,000 inhabitants.

[24] To adjust for problems of multiple inference and draw general conclusions about the effects of the TseTse,

The first five variables in the *Ethnographic Atlas* characterize a group's subsistence strategy. Each variable is categorical (with eight to ten categories) and captures how much a group depends on a given food production system. The variables are entitled gathering (v1), hunting (v2), fishing (v3), animal husbandry (v4) , and agriculture (v5) and range from 0 to 100 percent dependence. For ease of interpretation, these data are analyzed using OLS, and the results are reported in Table II, though graphical results of marginal effects of the TSI following ordered logit estimation are provided in Appendix Figure A.IV.[25] The results in Table II support the idea the TseTse affected greatly the pattern of food production in Africa. Each column is a separate outcome, and all specifications include the full set of geographic and climate controls in Table I column (4). A one standard deviation increase in the TSI is associated with a 37.3 percent reduction in reliance on animal husbandry and a 21.6 percent increase in reliance on hunting. Gathering is positively correlated with the TSI, and this correlation might be due to the complementarities between hunting and gathering. Dependence on agriculture in the aggregate is not significantly correlated with the TSI, suggesting that the main impact of the TseTse for agriculturalists was influencing the strategy (slash and burn versus intensive farming) used in cultivation. Fishing is likely determined by access to water resources rather than the TseTse fly, thus the absence of a correlation between the TSI and this outcome is reassuring. Coefficients on the malaria ecology index are also reported. Consistent with their distinct biological mechanisms, malaria, in contrast to TseTse, has no statistical correlation with husbandry or hunting.

---

a summary index is used. Appendix Table A.II reports summary measures of the effect of the TseTse fly on agriculture, urbanization and institutions. The average effect size (AES) coefficients calculate the mean (standardized) effect of the TSI across various outcomes and is calculated following Kling, Liebman, and Katz (2007). Seemingly unrelated regression is used to estimate the sample variance of the AES estimator. The results are reported for both the full sample and the Tropics subsample. The findings reinforce those obtained when the outcomes are examined individually.

A second approach to address the multiple inference problem is to use a familywise error rate or a false discovery rate (FDR) to adjust the p-value (for an overview see Newson and the ALSPAC Study Team (2003)). Using the step-down Holland FWER, all outcomes with the exception of precolonial centralization are statistically significant at the corrected p-value cutoff of 0.025 for five percent significance. All null hypotheses of no effect of the TSI are rejected when using the Benjamini and Liu step-down FDR.

[25]The graphs in Figure A.IV plot the predicted probability (and 95 percent confidence interval) that an ethnic group falls into the top quartile of dependence on a given subsistence strategy.

III.2 Threats to Validity

The greatest threat to the validity of the analysis is that the TSI is picking up climate factors that are inhospitable to the keeping of livestock, use of the plow, or human settlement. Figure I suggests that this is not the case–agricultural suitability and TseTse suitability are positively correlated. This concern is also addressed below by showing that the TSI does not have the same predictive power in the Tropics outside of Africa in Section IV.3. A series of robustness tests are performed in this subsection to specifically tackle this concern and the results are reported in Table III.

First, if the TSI is picking up the true effect of the TseTse, slight perturbations to the formulas given in Appendix Table C.I should not produce the same results. These perturbations include manipulating the laboratory data gathered to generate the TSI, by shifting the temperature entries one standard deviation to the right and to the left of the true observations. Such perturbations generate two new curves for the birth rate of the TseTse which have the same non-monotonic shape as the original data.[26] The formula for the death rate is left intact except the threshold under which the fly enters a chill coma is raised by one standard deviation (approximately three degrees Celsius). These quantitatively slight but physiologically significant changes eliminate any significant correlation between the TSI and the main outcomes of interest, as shown in Table III columns (1) and (2).

A second concern is that the underlying physiological relationship between climate and TseTse survival is manipulated via demographic modeling to achieve a steady state. Therefore, a more straightforward TseTse suitability index is used. Instead of the TSI, the intrinsic growth rate of the fly (the birth minus the death rate), is used directly. The results are statistically significant and are reported in column (3). Note that the point estimates differ as the index in the benchmark estimates are the (normalized) steady state number of TseTse flies, and this index is simply their growth rate. Since the TSI has a negative skew, a box

[26] The formula for the perturbed curves can be compared to Table C.I: birth rate shifted left= $(-0.0058 * meantemp^2 + 0.2484 * meantemp - 1.6385)$,
and birth rate shifted right= $(-0.0058 * meantemp^2 + 0.3210 * meantemp - 3.4084)$.

plot transformation to the variable N^* is applied in column (4) and yields identical results. A third concern is that the TSI is constructed by cherry-picking parameter values. Two approaches allay these concerns. A different method to predict the fly distribution is employed. The climatic conditions for fly survival are taken from field research by Rogers and Randolph (1986) and are converted into a binary indicator of "optimal fly survival."[27] With the exception of the centralization variable, the alternative TSI produces almost identical results to the model developed above (column (5)). Second, a sensitivity analysis of parameter values is undertaken and results reported in Appendix Table A.III.

The fourth issue involves statistical inference. Conley standard errors are reported in column (6) with cutoff values of 10 degrees latitude and longitude.[28] All results remain significant at conventional levels. Standard errors are clustered by country in column (7) and by both country and cultural province in column (8). Note the standard errors are not uniformly larger using any of these alternative methods than the preferred method of clustering based on cultural relatedness. One feature of the current analysis that might be quite important but is difficult to assess is the possibility that more developed groups pushed weaker groups onto TseTse-infested areas in a time period before the Murdock *Map*. The TSI then represents not only the direct biological effects of the fly but also captures the effect of this negative selection. Observing how fixed effects for cultural relatedness, as a proxy for group ability, alter the TSI coefficient is one way to gauge the importance of this channel (Table A.VIII column (1)).[29]

[27] Rogers and Randolph (1986) define the optimum as the joint condition that the temperature lie between 22 and 27 degrees Celsius and the saturation deficit lie between 6 and 14 mm Hg. This is not the preferred method for the current analysis since they rely on field observations (which may be influenced by human activity) instead of laboratory experiments.

[28] Each degree is approximately 111 kilometers at the equator.

[29] This subsection has mentioned robustness checks for the more commonly raised concerns with the TSI identification strategy. Several other tests were performed and are mentioned here with the results gathered in Appendix Tables. Table A.V reports the results on a balanced sample. Table A.VI ignores potentially endogenous boundaries drawn by Murdock and constructs virtual countries. Table A.VII builds on the work of Kremer and Miguel (2004), who demonstrate that empirical results can be misleading if externalities associated with communicable disease are ignored. One spillover effect that could alter the interpretation of the results was if economic activity from highly TseTse suitable areas was simply displaced to ethnic groups in less TseTse suitable areas. On the other hand, diffusion of technologies might have been hindered by having neighbors that were affected by the TseTse. To investigate violations of the Stable Unit Treatment Value

III.3 The Differential Effect of the TseTse Suitability Index in Tropical Africa

To exclude the possibility that the TSI is identifying generic patterns between climate and agriculture it is necessary to check for correlations between the TSI and the outcomes of interest in the Tropics outside of Africa. The empirical analysis that follows compares areas that are equally suitable for TseTse in terms of specific climatic conditions, however, because the fly itself is restricted to Africa (see footnote 4 for a discussion), conditional on the same covariates, the marginal effect of an increase in the TSI on historical agricultural, urbanization and institutional variables should be negligible except when interacted with a dummy for the African continent. It is straightforward to extend the TseTse population growth model to other areas. The sample now includes all groups wholly located within the Tropics of Capricorn and Cancer. The estimating equation is as follows:

$$Outcome_j = \alpha + \beta\, TSI_j + \delta\, TSI_j \cdot I_j^{Africa} + \mathbf{X}_j'\Omega + I_j^{Africa}\mathbf{X}_j'\Gamma + \eta_j, \qquad (2)$$

where I_j^{Africa} is a dummy variable for ethnic groups located within Africa. This approach is similar to the main estimating equation except the TSI enters as the main effect and TseTse is identified as the specific interaction between a dummy variable for Africa and the TSI. All other geographic and climate covariates from Table I column (4) are included as main

Assumption (SUTVA), the specification above is modified to allow for specific cross-unit spatial interaction by creating spatial lags using an inverse distance weighting matrix as in Drukker et al. (2013): $y_i = \alpha_1 + \delta TSI_i + \gamma \sum_j w_{ij} TSI_j + \mathbf{X}_i'\Omega + \varepsilon_i$. Here the γ represents the coefficient on the spatial lag TSI and $\mathbf{X}$ represents the k other covariates and their spatial lags. The spatial lag TSI is often insignificant and usually has the same sign as the main effect. For example, γ for the outcome "use of large domesticated animals" is negative and significant, suggesting that part of the association between the TSI and this outcome is due to (lack of) diffusion. On the other hand, the spatial lag TSI is positive and significant for population density, suggesting that some of the effect of the association between the TSI and this outcome might be operating indirectly via migration of groups towards less TseTse suitable areas. Table A.VIII includes ten additional checks including adding various fixed effects (columns (1)-(2)), estimating with a probit (column (3)), using a different measure of malaria from Hay et al. (2004) (column (4)) and adding additional controls (columns (5)-(10)) including suitability for crops that depend on the plow (as defined in Alesina, Giuliano, and Nunn (2013)), suitability for other common crops, slave exports and land area by ethnic group. Controlling for the export slave trade is important as this was predominately an Africa-specific phenomenon that could have depopulated ethnic groups and weakened state development. Distance to railways (circa 1908) was included as a control variable and did not significantly change the results (results available on request).

effects and as interactions with the Africa dummy.[30] Thus, Africa is allowed to differ from the rest of the world in many ways, not just the TseTse.[31]

A challenge to this analysis is that only the center of ethnic groups are mapped outside of Africa, not the entire boundaries. The standard approach to overcome this problem is to draw a circular "buffer zone" around the centroid. This approach is shown in panel (C) of Figure V. If the buffer zones are chosen too large, they overlap, making it difficult to allocate territory to mutually exclusive ethnic groups. If the buffer zones are too small, they will poorly approximate the actual boundaries. The approach followed in this paper is to construct Thiessen polygons, which more nearly approximate boundaries (compare Figure V panels (A) and (B)).[32] The starting point to construct the Thiessen polygons are the centroids of the ethnic groups as reported in the *Ethnographic Atlas* or given by the *Murdock Map* for continental Africa. Using this method, boundaries for the vast majority of ethnic groups located wholly within the Tropics of Capricorn and Cancer are able to be created. It should be noted that approximately half of the non-African Tropical ethnic groups reside on islands.[33] A comparison between the δ coefficients from Equation (1) for the *Map* versus Thiessen polygons sample is shown in Table IV. The coefficients are not significantly different between the Murdock and Thiessen polygon samples.

Results from Equation (2) are shown in Table V. The second column represents the differential effect of the TSI within Africa. The coefficient on the $\left(TSI \cdot I^{Africa}\right)$ interaction

[30] Proportion of land area in the Tropics does not enter as a control since it is unity in this sample.

[31] Standard errors are clustered by language families which are broader categories than cultural provinces though still capture spatial and cultural relatedness. There are 36 clusters. This change was necessary because provinces are not available outside of Africa. Clustered and Conley (1999) standard errors are sometimes smaller than heteroskedasticity-robust standard errors. Population density from Murdock is only produced for Africa and therefore is not included in the falsification exercise.

[32] For a set of points S in Euclidean space, a Thiessen polygon (also known as a Voronoi diagram) is one such that every point in the constructed polygon is closer to one such point p than to any other point in S. Within Africa, Thiessen polygons have a higher correlation with the Murdock map boundaries than the buffer zone technique. The one drawback is that, for observations with identical centroids, the Thiessen polygons will be identical (though this would also occur with the use of buffer zones).

[33] Remote island communities (defined as ethnic groups greater than 500 kilometers from the mainland) are excluded from the analysis since they are missing key geographic controls and filling in such missing values using a nearest neighbor approach would be inaccurate. The islands that were excluded were those of the South Pacific; specifically, Micronesia, the Cook Islands, French Polynesia, Jarvis Island, Orchid Island, Palau, Pitcairn Islands, Samoa, Tonga, Tuvalu, and Wallis & Futuna.

is always significant and has the expected sign. This provides evidence that the TSI is not simply capturing a generic pattern between the Tropics and development. The first column represents the effect of the TSI in the Tropical world outside of Africa. The coefficients are close to zero for most of the outcomes. For five of the six outcomes the main TSI effect has the opposite sign of the $\left(TSI \cdot I^{Africa}\right)$ interaction and is not significant.[34] The third column is the sum of the first two columns, and represents the total effect of the TSI within Africa. The TSI does not correlate with plow use in Tropical Africa in the synthetic sample, a result which is not surprising given that it failed to reach significance for this outcome in Table IV.

III.4 Africa without the TseTse: Archeological Evidence and Simulation

The results presented so far are supportive of the hypothesis that the TseTse influenced food production, urbanization and institutional development in historical Africa. One natural question to pose is how Africa would have developed in the absence of the fly. Archeologists provide qualitative evidence that Africa without the TseTse would have been more advanced. The civilization of Great Zimbabwe was located on a plateau between the Zambezi and Limpompo rivers and has been described by archeologist Tim Connah (1987, p. 228) as a "peninsula in a sea of tsetse." Garlake (1978) noted that the boundaries of the Great Zimbabwe complex corresponded to the climatic boundaries of the TseTse described by Rogers and Randolph (Figure VI). The people of Great Zimbabwe greatly relied on cattle, as deduced from skeletal remains of livestock around the site. Their economy was complex, integrating cereal agriculture, pastoralism and trade. The elliptical building at the center of the city was by far the largest single precolonial structure in sub-Saharan Africa.[35] This evidence can be supplemented with a quantitative analysis. Using the specification presented

[34] The positive plow coefficient outside of Africa is driven by groups in three countries (India, China and Indonesia) and might be due to chance or to another geographic factor that is correlated with the TSI and particularly important for food production in those countries (such as rice suitability).

[35] Speculative theories for the decline of Great Zimbabwe include ecological disaster from overgrazing and climate change leading to an encroachment of fly belts onto the plateau.

in Equation (2), predicted values of the precolonial outcomes are generated for the Africa sample. The average values of these predicted outcomes are shown in Table VI, column (1). To represent a reduction in the burden of TseTse-related disease, every element in the $\left(TSI \cdot I^{Africa}\right)$ vector is reduced by one standard deviation. The average values of the predicted outcomes using the new $\left(TSI \cdot I^{Africa}\right)$ vector are shown in Table VI column (2). The outcomes for Africa are now closer to those of Eurasia. This exercise should be viewed with much caution–it does not take into account the endogenous response to an historical reduction in the burden of disease from the TseTse. Perhaps Africa would have been colonized earlier without the TseTse barrier.[36]

IV The TseTse and Current Development

Disease could affect development through its historical effect on shaping institutions and/or through contemporaneous impacts on health. There are various approaches to disentangle these two channels. One is to study eradication campaigns. Places where the fly no longer poses a threat to health would permit isolation of the historical channel. Unfortunately no large scale TseTse eradication campaigns have been successful. Another approach would be to find locations where the climate has changed sufficiently so that places historically suitable for the fly are no longer suitable or vice-versa. Since temperature changes would likely affect the Northern and Southern limits of where the fly can potentially exist, this approach would be suitable for a regression discontinuity study. However, institutional and economic variables are generally not available over the several decades one would need to capture significant climatic changes at a disaggregated level for much of Africa. The approach herein is to perform a cross-sectional analysis of the reduced form relationship between modern economic outcomes and the TSI at the district level. The modern outcomes include satellite light density collected by the US Air Force Weather Agency and processed by NOAA and

[36] For example, the late colonization of Africa is thought to be TseTse-related. Horseback-riding Muslim armies presumably had difficulty penetrating further south than the Sahel (Fukuyama, 2011, p. 91). Because it has served as a check on agricultural expansion and grazing, the TseTse could have played a role in protecting African biodiversity.

the observed distribution of cattle (FAO, 2005). The TSI is constructed with modern climate data. The modern climate data are from the East Anglia Climate Research Unit and span the years 1961-1990. The data are made available as monthly means at a 10 minute resolution. If a correlation between the contemporaneous TseTse measure and economic development is noted, the test will be to see if it remains significant after controlling for precolonial centralization, a historical institution shown to be correlated with the TSI.[37]

Inclusive economic and political institutions have been argued to be fundamental explanations for income disparities worldwide (Acemoglu and Robinson, 2012). Pluralistic political institutions that are not centralized may devolve into chaos (Acemoglu and Robinson, 2012, p. 81), whereas centralized ones can enforce rules, deliver public goods and encourage economic growth. Recent empirical studies have demonstrated a positive correlation between economic development and precolonial political centralization in Africa at the national and subnational level (Gennaioli and Rainer, 2007; Michalopoulos and Papaioannou, 2013, 2014). To investigate whether the correlation between economic outcomes and TseTse is mediated through this channel, a population weighted average of precolonial centralization in district d, of country c across ethnic groups j is constructed similar to the plow measure of Alesina, Giuliano, and Nunn (2013): $Historical\ Centralization_{d,c} = \frac{\sum_j L_{j,d,c} \cdot \mathbf{I}_j}{L_{d,c}}$, where I_j is equal to one if ethnic group j was historically centralized and zero otherwise, $L_{j,d}$ denotes the number of individuals in ethnic group j living in district d and L_d is the total population in the district. This index captures the proportion of a district's current inhabitants whose ancestors lived in a centralized society. The population data are from LandScan 2007 at a resolution of 1 kilometer. The ethnicities are mapped to their current location by using the *Ethnologue* (Lewis, 2009), a shape file that includes the current geographic distribution of

[37] An earlier version of the paper included a cross-country analysis. Motivated by comments from two anonymous referees, these findings have been replaced with a subnational analysis at the district level. The subnational findings generally support those found at the cross-country level, while allowing for the inclusion of a richer set of climate controls such as those comprising the TSI (temperature and humidity), as well as their first-order interaction and absolute latitude. At the cross-country level, the smaller sample size hinders the ability to identify the effect of the TseTse conditional on absolute latitude or a richer set of climate controls.

languages. Ethnic groups in the *Ethnographic Atlas* were matched to the *Ethnologue* thereby mitigating errors associated with assigning institutions to places where an ethnic group no longer resides.[38] Equation (1) is used to estimate the reduced form relationship between the TSI and development at the district level. In addition to the geographic and climate variables introduced above, the estimating equation will now also include country fixed effects to proxy for modern institutions and/or policies.

The results are shown in Table VII. Panel (A) reports results for the log of satellite lights and panel (B) reports results for the log number of cattle. Column (1) controls for the climate variables in the TSI as well as the proportion of land area in the Tropics. A one standard deviation increase in the TSI, conditional on these controls, is associated with a 38 percent reduction in light density and a 72 percent reduction in the number of cattle. The model in column (2) adds the malaria ecology index. This variable is not statistically significant in this specification and does not affect the TSI coefficient. Column (3) adds absolute latitude to control generally for Tropical climate, mean altitude and agricultural suitability. Lights are reflected in water causing areas close to water to register higher values (a phenomenon known as blooming); therefore, proximity to the coast and an inland body of water enter in all specifications that follow. Column (4) adds country fixed effects to control for modern institutions. Both outcome variables are impacted by this addition, with the TSI coefficient falling by 39 percent in panel (A) and by 57 percent in panel (B), though the coefficients retain significance at conventional levels. In column (5) the fraction of individuals historically belonging to ethnic groups that were centralized is added to the regression. The findings for the satellite lights outcome are similar to Michalopoulos and Papaioannou (2013, 2014)–precolonial centralization has a positive and significant effect on current economic performance.[39] For lights, but not livestock, adding precolonial centralization to the regression reduces the TSI coefficient by 35 percent and it loses its significance.

[38] I am grateful to Nathan Nunn for providing the crosswalk between the two data sources.

[39] Adding slavery does little to the TSI coefficient for either outcome, reinforcing the results of Michalopoulos and Papaioannou (2013) who show that other variables in the *Ethnographic Atlas* do not have the same strong correlation with light density as does precolonial centralization.

The estimates are imprecise but provide suggestive evidence that the direct impact of TseTse on modern economic development is negligible after controlling for historical, institutional pathways of influence.[40]

V Concluding Remarks

This study has investigated the effect of the TseTse on African development. Using insect population growth models and laboratory experiments of TseTse physiology, a suitability index for TseTse was constructed. This index was then joined with ethnographic data on precolonial African agricultural practices, institutions and urbanization. Historical TseTse suitability was correlated with less advanced agricultural practices, the use of slaves and a lower population density within but not outside of Africa.

Simulating African development with a lower burden of historical TseTse-transmitted disease demonstrated that there could have been modest increases in intensive cultivation and political centralization in the precolonial period. These results should be interpreted with caution given that they do not allow for an endogenous response to the fly's removal. However, the predictions are broadly consistent with the archaeological record which documents that relatively advanced civilizations flourished in areas of Africa inhospitable to the fly.

The findings suggest TseTse-associated disease continues to influence development mainly through its effect on precolonial centralization; thus providing support for the Engerman and Sokoloff (2000) view that endowments may shape institutions and thereby have long-run effects. On the other hand, the distribution of cattle exhibits a negative correlation with the TSI conditional on country fixed effects and local, historical institutions. This finding may reflect the continued relevance of TseTse-transmitted *Trypanosomiasis* on animal husbandry in Africa today.

[40] Population density exhibits the same pattern of correlation as light density with the TSI and is therefore not included as a control. This is not surprising since historical population density was also correlated with the TSI.

References

Acemoglu, Daron and Johnson, Simon and Robinson, James A . (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review 91*, 1369–1401.

Acemoglu, Daron and James A . Robinson. (2012). *Why Nations Fail: The Origins of Power, Prosperity and Poverty.* New York: Crown Publishers.

Alesina, Alberto and Giuliano, Paola and Nunn, Nathan. (2013). On the Origins of Gender Roles: Women and the Plough. *Quarterly Journal of Economics 128*, 469–530.

Allen, Robert C. (1999). Tracking the Agricultural Revolution in England. *The Economic History Review 52*, 209–235.

Altonji, J. and Elder, T. and Taber, C. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy 113*, 151–184.

Anderson, C.J. and Fragmont, J.C. (1857). *Lake Ngami: or, Explorations and discoveries during four years' wanderings in the wilds of southwestern Africa.* New York: Dix, Edwards & Co.

Ashraf, Quamrul and Oded Galor. (2013). The Out of Africa Hypothesis, Human Genetic Diversity, and Comparative Economic Development. *American Economic Review 103*, 1–46.

Ashraf, Quamrul and Stelios Michalopoulos. (2013). The Climatic Origins of the Neolithic Revolution: Theory and Evidence. *National Bureau of Economic Research Working Paper 18765*.

Bairoch, Paul. (1988). *Cities and Economic Development : From the Dawn of History to the Present.* Chicago: University of Chicago Press.

Banerjee, Abhijit and Iyer, Lakshmi. (2005). History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India. *American Economic Review 95*, 1190–1213.

Robert Bates. (1983). *Essays on the Political Economy of Rural Africa.* Berkeley: University of California Press.

Bloom, David E. and Sachs, Jeffrey D. (1998). Geography, Demography, and Economic Growth in Africa. *Brookings Papers on Economic Activity 1998*, 207–295.

Bonnassie, P. (2009). *From Slavery to Feudalism in South-Western Europe.* Cambridge, U.K.: Cambridge University Press.

Borst, P. and Rudenko, G. (1994). Antigenic Variation in African Trypanosomes. *Science (New York, N.Y.) 264*, 1872–3.

Boserup, Ester. (1966). *The Conditions of Agricultural Growth: The Economics of Agrarian Change Under Population Pressure.* Chicago: Aldine Pub. Co.

Brown, K. and Gilfoyle, D. (2010). *Healing the Herds: Disease, Livestock Economies, and the Globalization of Veterinary Medicine.* Athens, Ohio: Ohio University Press.

Bursell, E. (1960). Loss of water by Excretion and Defecation in the Tsetse Fly. *Journal of Experimental Biology 37*, 689–697.

Chandler, Tertius. (1987). *Four Thousand Years of Urban Growth : An Historical Census.* Lewiston, New York: St. David's University Press.

Chaves, Isaias and Engerman, Stanley L. and Robinson, James. (2013). Reinventing the Wheel: The Economic Benefits of Wheeled Transportation in Early Colonial British West Africa. *National Bureau of Economic Research Working Paper 19673*.

Comin, Diego and Easterly, William and Eric Gong. (2010). Was the Wealth of Nations Determined in 1000 BC? *American Economic Journal: Macroeconomics 2*, pp. 65–97.

Compo, GP and Whitaker, J.S. and Sardeshmukh, P.D.G. (2011). The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meterological Society 137*, 1–28.

Conley, T.G. (1999). GMM Estimation with Cross Sectional Dependence. *Journal of Econometrics 92*, 1–45.

Connah, Graham. (1987). *African Civilizations : Precolonial Cities and States in Tropical Africa : An Archaeological Perspective.* Cambridge, U.K.: Cambridge University Press.

Dell, Melissa. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica 78*, 1863–1903.

Diamond, Jared M. (1997). *Guns, Germs, and Steel : The Fates of Human Societies* (1st ed.). New York: W.W. Norton.

Domar, Evsey D. (1970). The Causes of Slavery or Serfdom: A Hypothesis. *The Journal of Economic History 30*, 18–32.

Drukker, David and Peng, Hua and Prucha, Ingmar and Raciborski, Rafal. (2013). Creating and Managing Spatial-weighting Matrices Using the spmat Command. *The Stata Journal 13*, 242–286.

Engerman, Stanley L. and Kenneth L. Sokoloff. (2000). History Lessons: Institutions, Factors Endowments, and Paths of Development in the New World. *The Journal of Economic Perspectives 14*, 217–232.

FAO. (1998). A Field Guide for the Diagnosis, Treatment and Prevention of African Animal Trypanosomiasis. Rome.

——— (2005). Observed Cattle Density Africa.

Fenske, James. (2013). Does Land Abundance Explain African Institutions? *Economic Journal*, 1363–1390.

Ford, J and Katondo, K.M. (1977). Maps of Tsetse Flies (Glossina) Distribution in Africa, 1973, according to Sub-Generic Groups on a Scale of 1: 5 000 0000. *Bulletin of Animal Health and Production in Africa 25*, 187–193.

Fukuyama, Francis. (2011). *The Origins of Political Order : From Prehuman Times to the French Revolution* (1st ed.). New York: Farrar, Straus and Giroux.

Gallup, JL and Sachs, JD. (2001). The Economic Burden of Malaria. *The American Journal of Tropical Medicine and Hygiene 64*, 85–96.

Oded Galor and Michalopoulos, Stelios. (2012). Evolution and the Natural Selection of Entrepreneurial Traits. *Journal of Economic Theory 147*, 759–780.

Garlake, P. S. (1978). Pastoralism and Zimbabwe. *The Journal of African History 19*, 479–493.

Gennaioli, Nicola and Rainer, Ilia. (2007). The Modern Impact of Precolonial Centralization in Africa. *Journal of Economic Growth 12*, 185–234.

Gifford-Gonzalez, Diane. (2000). Animal Disease Challenges to the Emergence of Pastoralism in Sub-Saharan Africa. *The African Archaeological Review 17*, 95–139.

Glaeser, Edward and Rafael La Porta and Florencio Lopez-De-Silanes and Andrei Shleifer. (2004). Do Institutions Cause Growth? *Journal of Economic Growth 9*, pp. 271–303.

Glasgow, J. P. (1963). *The Distribution and Abundance of Tsetse.* International Series of Monographs on Pure and Applied Biology. New York: Macmillan.

Jack Goody. (1971). *Technology, Tradition and the State in Africa.* London: Oxford University Press.

Hay, Simon I. and Guerra, Carlos A. and Tatem, Andrew J. and Noor, Abdisalan M. and Snow, Robert W. (2004). The Global Distribution and Population at Risk of Malaria: Past, Present, and Future. *The Lancet Infectious Diseases 4*, 327–336.

Herbst, J.I. (2000). *States and Power in Africa: Comparative Lessons in Authority and Control.* Princeton University Press.

Hopkins, A.G. (1973). *An Economic History of West Africa.* New York: Columbia University Press.

Hoppe, Kirk. (2003). *Lords of the Fly: Sleeping Sickness Control in British East Africa, 1900-1960.* Greenwood.

Iliffe, J. (1995). *Africans: The History of a Continent.* Cambridge, U.K.: Cambridge University Press.

Ingold, Tim. (1987). *The Appropriation of Nature: Essays on Human Ecology and Social Relations*. University of Iowa Press.

Iyer, Lakshmi. (2010). Direct versus indirect colonial rule in India: Long-term consequences. *Review of Economics and Statistics 92*, 693–713.

Jackson, C. H. N. (1949). The Biology of Tsetse Flies. *Biological Reviews 24*, 174–199.

Jedwab, Remi and Alexander Moradi. (2013). Colonial Investments and Long-Term Development in Africa: Evidence from Ghanian Railroads. *Working Paper*.

Johnston, H.H. (1894). *Report by Commissioner Johnston of the First Three Years' Administration of the Eastern Portion of British Central Africa*. Her Majesty's Stationery Office.

Kiszweski, A and Mellinger, A and Spielman, A and Malaney, P. and Sachs, SE and Sachs, JD. (2004). A Global Index Representing the Stability of Malaria Transmission. *The American Journal of Tropical Medicine and Hygiene 70*, 486–498.

Kling, Jefrrey and Liebman, Jeffrey and Katz, Lawrence. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica 75*, 83–119.

Krafsur, E. S. (2009). Tsetse flies: Genetics, Evolution, and Role as Vectors. *Infection Genetics and Evolution 9*, 124–141.

Michael Kremer and Edward Miguel. (2004). Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica 72*, 159–217.

Lambrecht, Frank L. (1964). Aspects of Evolution and Ecology of Tsetse Flies and Trypanosomiasis in the Prehistoric African Environment. *The Journal of African History 5*, 1–24.

Robin Law. (1977). *The Oyo Empire c.1600-c.1836: A West African Imperialism in the Era of the Atlantic Slave Trade*. Oxford: Clarendon Press.

Leak, S.G.A. (1999). *Tsetse Biology and Ecology: Their Role in the Epidemiology and Control of Trypanosomosis*. New York City: CABI Publishing in association with International Livestock Research Institute.

Legros, M and Lloyd, A and Huang, Y and Gould, F. (2009). Density-Dependent Intraspecific Competition in the Larval Stage of Aedes aegypti (Diptera: Culicidae): Revisiting the Current Paradigm. *Journal of Medical Entomology 46*, 409–419.

Lewis, M. P. (Ed.) (2009). *Ethnologue: Languages of the World, Sixteenth Edition*. Dallas, TX: SIL International.

Livingstone, D. (1857). *Missionary Travels and Researches in South Africa*. London: Harper and Brothers.

Anthony G. Luckins and Ron H. Dwinger. (2004). *The Trypanosomiases*, Chapter Nontsetse Transmitted Trypanosomiasis. New York City: CABI Publishing.

Mann, Michael E. and Zhang, Zhihua and Hughes, Malcolm K. and Bradley, Raymond S. and Miller, Sonya K. and Rutherford, Scott and Ni, Fenbiao. (2008). Proxy-based Reconstructions of Hemispheric and Global Surface Temperature Variations over the Past Two Millennia. *Proceedings of the National Academy of Sciences 105*, 13252–13257.

Maudlin, I. (2006). African Trypanosomiasis. *Annals of Tropical Medicine and Parasitology*, 679–708.

May, R.M. and Conway, G.R. and Hassell, M.P. and Southwood, T.R.E. (1974). Time-Delays, Density-Dependence and Single-Species Oscillations. *Journal of Animal Ecology 43*, 747–770.

Mazoyer, M. and Roudart, L. and Membrez, J.H. (2006). *A History of World Agriculture: From the Neolithic Age to the Current Crisis.* New York: Earthscan.

Mellanby, K. (1937). Experimental work with the Tsetse-Fly, Glossina Palpalis in Uganda. *Bulletin of Entomological Research 28*, 611–632.

Michalopoulos, S. and Papaioannou, E. (2013). Pre-colonial Ethnic Institutions and Contemporary African Development. *Econometrica 81*, 113–152.

Michalopoulos, Stelios and Papaioannou, Elias. (2014). National Institutions and Subnational Development in Africa. *The Quarterly Journal of Economics*, 151–213.

Moore, Nathan and Joseph Messina. (2010). A Landscape and Climate Data Logistic Model of Tsetse Distribution in Kenya. *PLOS One 5*, 1–10.

Mulla, A. F. and Rickman, L. R. (1988). How do African Game Animals Control Trypanosome Infections? *Parasitology Today 4*, 352–354.

Murdock, George Peter. (1959a). *Africa, its Peoples and their Culture History.* New York: McGraw-Hill.

——— (1959b). Tribal Map of Africa.

——— (1967). *Ethnographic Atlas.* University of Pittsburgh Press.

Murray, M. and Trail, J. C. and D'Ieteren, G. D. (1990). Trypanotolerance in Cattle and Prospects for the Control of Trypanosomiasis by Selective Breeding. *Revue Scientifique et Technique*, 369–386.

Nash, T.A.M. (1969). *Africa's Bane: The Tsetse Fly.* London: Collins.

Newson, Roger and the ALSPAC Study Team. (2003). Multiple-test procedures and smile plots. *The Stata Journal 3*, 109–132.

Nieboer, H.J. (1900). *Slavery as an Industrial System: Ethnological Researches.* The Hague: Martinus Nijhoff.

Nunn, Nathan. (2008). The Long-term Effects of Africa's Slave Trades. *The Quarterly Journal of Economics 123*, 139–176.

——— (2009). The Importance of History for Development. *Annual Review of Economics 1*, 65–92.

Nunn, Nathan and Qian, Nancy. (2011). The Potato's Contribution to Population and Urbanization: Evidence from A Historical Experiment. *The Quarterly Journal of Economics 126*, 593–650.

Nunn, Nathan and Leonard Wantchekon. (2011). The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review 101*, 3221–3252.

Olsson, Ola and Douglas Hibbs. (2005). Biogeography and Long-run Economic Development. *European Economic Review 49*, pp. 909–938.

Overton, M. (1996). *Agricultural Revolution in England: The Transformation of the Agrarian Economy, 1500-1850.* Cambridge, U.K.: Cambridge University Press.

Owaga, Mary L. A. (1985). Observations on the Efficacy of Buffalo Urine as a Potent Olfactory Attractant for Glossina Pallidipes Austen. *International Journal of Tropical Insect Science 6*, 561–566.

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

Rajagopal, P. K. and Bursell, E. (1965). The Effect of Temperature on the Oxygen Consumption of Tsetse Pupae. *Bulletin of Entomological Research 56*, 219–225.

Robinson, J. and Osafo-Kwaako, P. (2013). Political Centralization in Pre-Colonial Africa. *Journal of Comparative Economics 41*, 6–21.

Rogers, David J. and Randolph, Sarah E. (1986). The Distribution and Abundance of Tsetse Flies. *Journal of Animal Ecology 55*, 1007–1025.

Sachs, J. and Malaney, P. (2002). The Economic and Social Burden of Malaria. *Nature 415*, 680–685.

Schowalter, T.D. (2011). *Insect Ecology: An Ecosystem Approach.* Elsevier Science.

Spolare, Enrico and Romain Wacziarg. (2013). How Deep Are the Roots of Economic Development? *Journal of Economic Lit 51*, 325–369.

Terblanche, John S. and Clusella-Trullas, Susana and Deere, Jacques A. and Chown, Steven L. (2008). Thermal tolerance in a South-East African Population of the Tsetse Fly Glossina Pallidipes (Diptera, Glossinidae): Implications for Forecasting Climate Change Impacts. *Journal of Insect Physiology 54*, 114–127.

Turnbull, P. C. and Doganay, M. and Lindeque, P. M. and Aygen, B. and McLaughlin, J. (1992). Serology and Anthrax in Humans, Livestock and Etosha National Park Wildlife. *Epidemiology and Infection 108*, 299–313.

Vale, G. A. (1974). The Responses of Tsetse Flies (Diptera, Glossinidae) to Mobile and Stationary Baits. *Bulletin of Entomological Research 64*, 545–588.

Vale, G. A. and Flint, S. and Hall, D. R. (1986). The Field Responses of Tsetse Flies, Glossina spp. (Diptera: Glossinidae), to Odours of Host Residues. *Bulletin of Entomological Research 76*, 685–693.

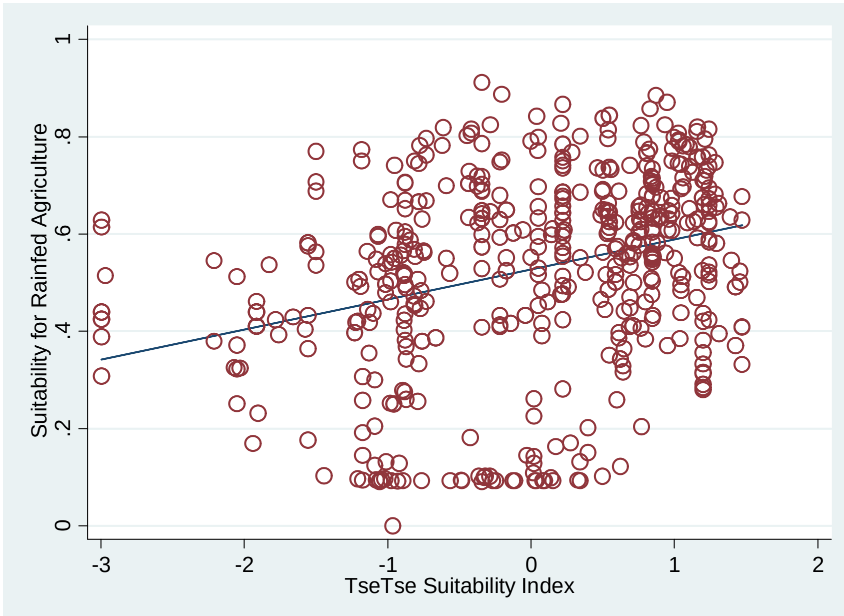
Wacziarg, Romain and Enrico Spolare. (2009). The Diffusion of Development. *Quarterly Journal of Economics 124*, pp. 469–529.

Wantchekon, L., and Novta, Natalija and Marko Klasnja. (2013). Education and Human Capital Externalities: Evidence from Colonial Benin. *Princeton Working Paper*.

Weil, David N., and Depetris-Chauvin, Emilio. (2013). The Impact of Malaria on African Development over the Longue Duree. *National Bureau of Economic Research Working Paper 19603*.

VI Figures and Tables

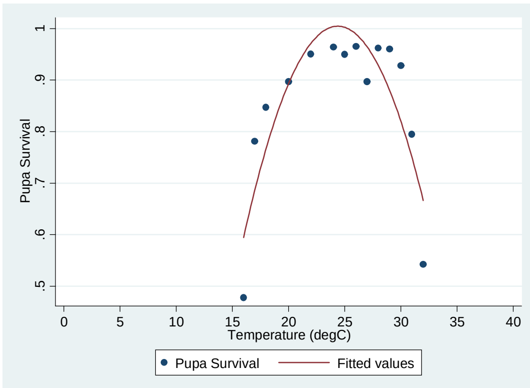
Figure I: Agricultural Suitability versus TseTse Suitability
(African Ethnic Groups)



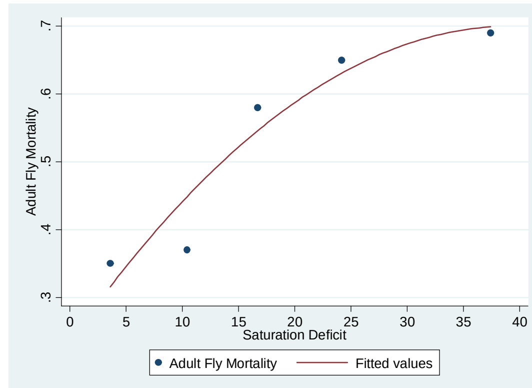
Notes: This figure demonstrates the correlation between agricultural suitability and TseTse suitability. Data on agricultural suitability for rainfed crops are from the FAO Global Agro-Ecological Zones (2002). Details on the FAO methodology can be found in Appendix B. The TseTse suitability index (TSI) is based on the author's calculations using climate data from NOAA 20th Century Reanalysis (1871). The equations for the TSI can be found in Appendix C Table C.I.

Figure II: Physiology of the TseTse Fly

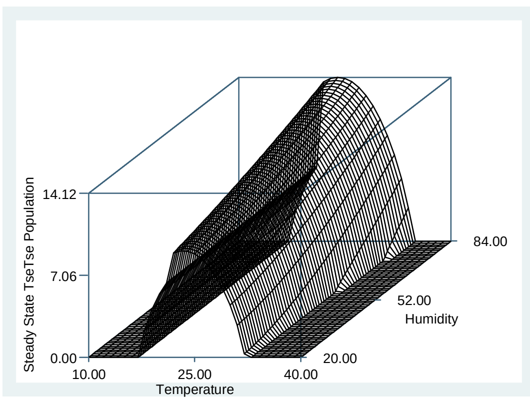
A. Pupa Survival and Temperature



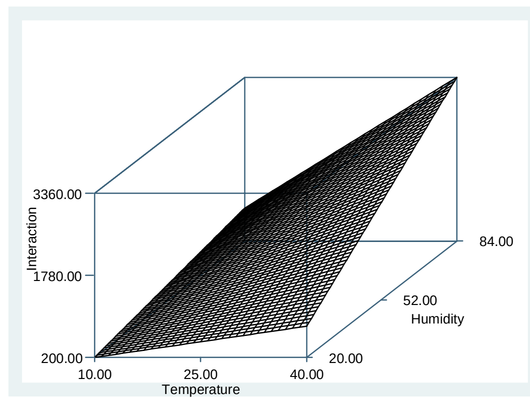
B. Adult Fly Mortality and Saturation Deficit



C. Steady State TseTse Population



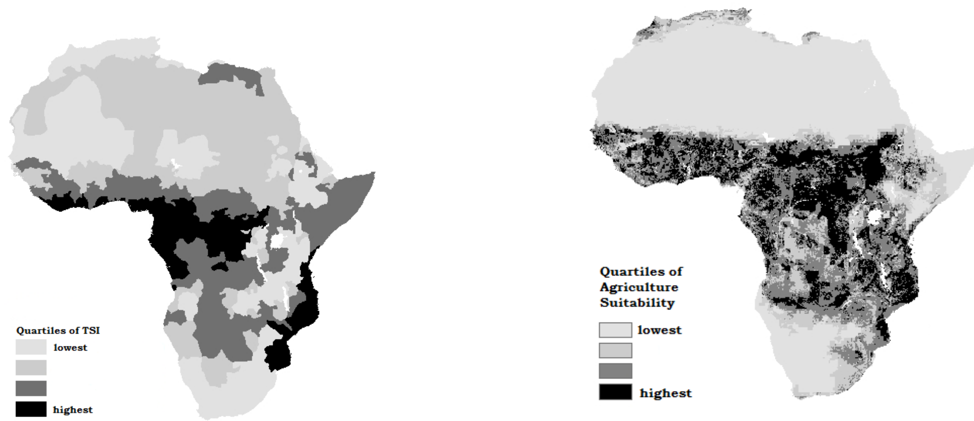
D. Linear Interaction of Climate Variables



Notes: These graphs show the relationship between TseTse physiology and climate. Panel (A) data points are from Bursell (1960) and Rajagopal and Bursell (1965). Panel (B) data points are from K. Mellanby (1937). Panel (C) depicts the steady state population of TseTse as a function of climate. Panel (D) plots the first-order interaction of humidity and temperature.

Figure III: TseTse Suitability Index and the Observed TseTse Distribution

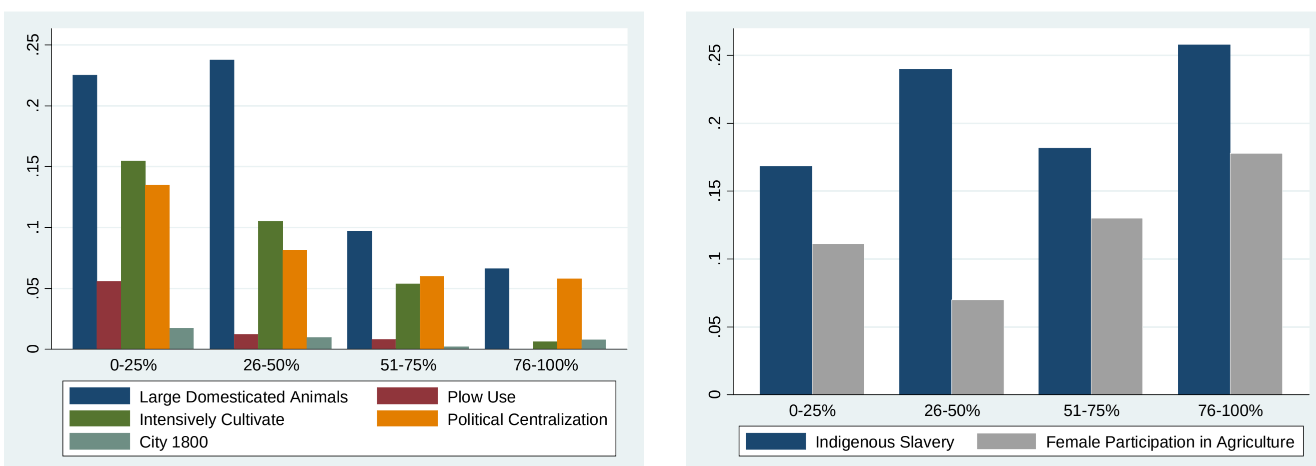
A. TseTse Suitability Index (1871) B. Suitability for Rainfed Agriculture (2002)



Notes: Panel (A) shows the historical TseTse suitability index created using climate data from NOAA's 20th century reanalysis for the year 1871. Panel (B) shows the suitability for rainfed agriculture (FAO, 2002).

Figure IV: Average Precolonial Outcomes by Quartile of TSI

A. B.



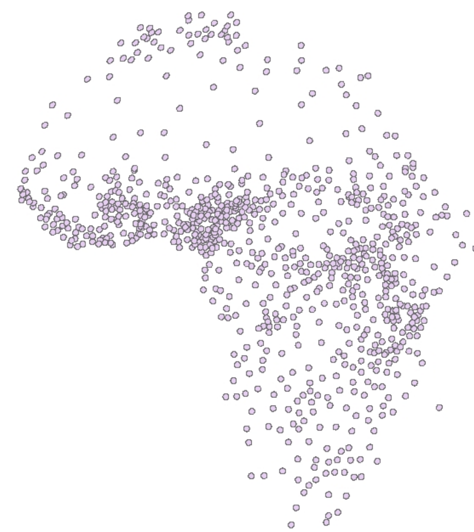
Notes: These graphs show the average of binary precolonial African outcomes by TSI quartile.

Figure V: Thiessen Polygons

A. Map

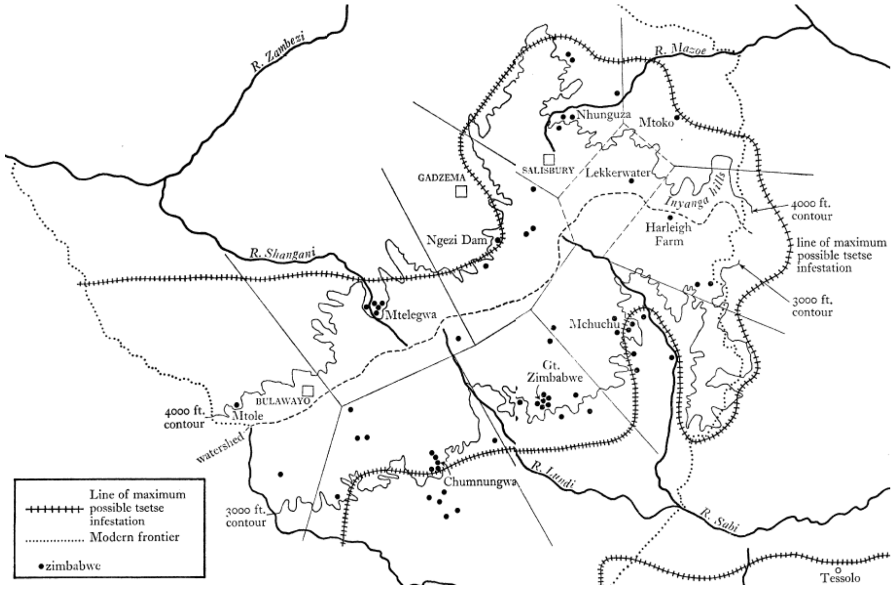
B. Thiessen Polygons

C. Buffer Zones



Notes: This figure compares the *Murdock Map* of ethnic groups in panel (A) with the constructed Thiessen polygons in panel (B) and with the buffer zone approach, panel (C).

Figure VI: Map of Great Zimbabwe



Notes: This figure demonstrates the overlap between the boundaries of Great Zimbabwe and the line of maximum possible TseTse infestation. The figure is used with permission from Garlake (1978). The hatched line indicates the line of maximum possible TseTse infestation as described by Oxford zoologists David Rogers and Sarah Randolph. The black dots represent the *zimbabwe*–stone enclosures indicative of human settlement.

Table I

REDUCED FORM ESTIMATES OF THE RELATIONSHIP BETWEEN HISTORICAL AFRICAN DEVELOPMENT AND TSETSE SUITABILITY

Dependent Variable	(1)	(2)	(3)	(4)	No. Obs	No. Clusters	Sample mean
Agriculture							
Large Domesticated Animals	–0.216***	–0.221***	–0.210***	–0.231***	484	44	0.626
	(0.044)	(0.046)	(0.045)	(0.042)			
Intensive Agriculture	–0.092***	–0.079**	–0.083***	–0.090***	485	44	0.320
	(0.028)	(0.031)	(0.029)	(0.028)			
Plow Use	–0.052**	–0.055**	–0.052**	–0.057**	484	44	0.076
	(0.021)	(0.023)	(0.022)	(0.023)			
Female Participation in Agriculture	0.254***	0.226***	0.214***	0.206***	315	43	0.489
	(0.053)	(0.053)	(0.054)	(0.060)			
Urbanization							
Log Population Density (Murdock)	–0.880***	–0.828***	–0.802***	–0.745***	398	43	1.700
	(0.272)	(0.256)	(0.244)	(0.229)			
Institutions							
Indigenous Slavery	0.095**	0.109**	0.105**	0.101**	446	44	0.848
	(0.040)	(0.043)	(0.043)	(0.040)			
Centralization	–0.058	–0.076**	–0.075**	–0.075**	467	44	0.334
	(0.035)	(0.035)	(0.035)	(0.035)			
Climate Controls	Y	Y	Y	Y			
Malaria Controls	N	Y	Y	Y			
Waterway Controls	N	N	Y	Y			
Geography Controls	N	N	N	Y			

Notes: OLS estimates of Equation (1). Each cell in columns (1)–(4) represents a separate regression and the coefficient on the TSI is reported. The dependent variable is listed in the leftmost column of the upper panel. The mean of the dependent variables is located in the rightmost column. With the exception of the urbanization category, the data for the dependent variables are from the *Ethnographic Atlas,* and are based on anthropological observations from the late 19th and early 20th century. The data for population density are from Murdock's book *Africa Its Peoples and Their Culture History*. The climate variables are from the 20th century reanalysis for the year 1871. Climate controls refer to temperature, relative humidity and the first-order interaction between temperature and humidity as well as the proportion of land that is in the Tropics. Malaria refers to the malaria ecology index developed by Kiszewski et al., (2004). Waterway controls include whether a river was located within the ethnic group boundaries and whether the boundaries included a coast. Geography controls include mean altitude, the FAO's agricultural suitability index, longitude and absolute latitude. Robust standard errors clustered at the level of cultural province in parentheses. * ** *** Significant at 10, 5 and 1 percent levels.

Table II

REDUCED FORM ESTIMATES OF THE RELATIONSHIP BETWEEN TSETSE SUITABILITY AND SUBSISTENCE PATTERNS WITHIN HISTORICAL AFRICA

	(1)	(2)	(3)	(4)	(5)
	Husbandry	Hunting	Gathering	Agriculture	Fishing
TSI	–0.373**	0.216***	0.170**	0.046	–0.060
	(0.164)	(0.073)	(0.084)	(0.191)	(0.082)
Malaria Ecology Index	–0.005	–0.005	0.013	0.002	–0.005
	(0.019)	(0.006)	(0.009)	(0.020)	(0.014)
No. Obs	522	522	522	522	522
No. Clusters	44	44	44	44	44
R-squared	0.436	0.121	0.101	0.269	0.163

Notes: OLS estimates of Equation (1). Each column is a separate regression. The dependent variable is listed as a column heading and represents the percent dependence on a given food production strategy. Each cell reports the coefficient on the TSI or the malaria ecology index. The malaria ecology index was developed by Kiszewski et al., (2004). Other covariates (not reported) include temperature, relative humidity the first-order interaction between temperature and humidity, the proportion of land that is in the Tropics, whether a river was located within the ethnic group boundaries and whether the boundaries included a coast. In addition, the regressions include mean altitude, the FAO's agricultural suitability index, longitude and absolute latitude. Robust standard errors clustered at the level of cultural province in parentheses. * ** *** Significant at 10, 5 and 1 percent levels.

Table III

ROBUSTNESS TESTS OF THE REDUCED FORM ESTIMATES OF TSETSE SUITABILITY AND HISTORICAL AFRICAN DEVELOPMENT

Dependent Variable	Alternative TseTse Indices					Alternative Clustering		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Perturb TSI Shift Left	Perturb TSI Shift Right	Intrinsic Rate of Growth	Box-Plot	Optimal TseTse Conditions	Conley S.E.	S.E. Clustered by Country	Multi-way Clustering
Agriculture								
Large Domesticated Animals	0.016	0.008	–1.392***	–0.009***	–0.367***	–0.231***	–0.231***	–0.231***
	(0.066)	(0.083)	(0.258)	(0.002)	(0.077)	(0.040)	(0.039)	(0.037)
Intensive Agriculture	0.024	0.022	–0.567***	–0.004***	–0.127**	–0.090***	–0.090***	–0.090***
	(0.038)	(0.048)	(0.174)	(0.001)	(0.059)	(0.030)	(0.030)	(0.023)
Plow Use	–0.021	0.013	–0.325**	–0.002***	–0.044**	–0.057*	–0.057*	–0.057*
	(0.013)	(0.035)	(0.142)	(0.0009)	(0.022)	(0.031)	(0.032)	(0.031)
Female Participation in Agriculture	–0.079	–0.055	1.241***	0.008***	0.265***	0.206***	0.206***	0.206***
	(0.055)	(0.070)	(0.363)	(0.002)	(0.060)	(0.041)	(0.047)	(0.058)
Urbanization								
Log Population Density (Murdock)	–0.293	–0.229	–4.479***	–0.029***	–0.663**	–0.745***	–0.745***	–0.745***
	(0.225)	(0.291)	(1.415)	(0.009)	(0.284)	(0.197)	(0.199)	(0.234)
Institutions								
Indigenous Slavery	0.020	0.029	0.619**	0.004**	0.100*	0.101**	0.101**	0.101**
	(0.040)	(0.055)	(0.242)	(0.002)	(0.054)	(0.039)	(0.040)	(0.040)
Centralization	0.029	–0.004	–0.464**	–0.003**	–0.004	–0.075*	–0.075**	–0.075**
	(0.060)	(0.084)	(0.214)	(0.001)	(0.058)	(0.043)	(0.036)	(0.035)

Notes: The dependent variable is noted in the leftmost column. Each cell represents a separate regression and the coefficient on a measure of TseTse is reported. With the exception of the urbanization category, the data for the dependent variable are from the *Ethnographic Atlas*, and are based on anthropological observations from the late 19th and early 20th century. The data for population density are from Murdock's book *Africa Its Peoples and Their Culture History*. The climate variables are from the 20th century reanalysis for the year 1871. All specifications include temperature, relative humidity the first-order interaction between temperature and humidity and the proportion of land area in the Tropics as well as mean altitude, the FAO's agricultural suitability index, access to waterways, the malaria ecology index developed by Kiszewski et al., (2004), absolute latitude and longitude. Columns (1) and (2) replace the TSI with a TSI generated by perturbations to the formula in Table C.I as described in Section IV.2. Column (3) replaces the TSI with λ, the intrinsic rate of growth. Column (4) is a box-plot transformation of the steady state number of flies-seven observations (where N=0) are dropped in the process of constructing this variable. Column (5) reports the results from an alternative measure for TseTse suitability developed by Rogers and Randolph (1986). Column (6) reports results using Conley (1999) standard errors to account for spatial correlation with cutoffs of 10 degrees latitude and 10 degrees longitude. Column (7) clusters standard errors by country. Column (8) uses a multiway clustering method developed by Cameron, Gelbach and Miller (2011) and clusters standard errors along the country and cultural province dimensions. Robust standard errors clustered at the level of cultural province in parentheses unless otherwise indicated. * ** *** Significant at 10, 5 and 1 percent levels.

Table IV

COMPARISION OF MURDOCK MAP AND THIESSEN POLYGON COEFFICIENTS

Tropics Only

Dependent Variable	(1) Map	(2) Polygon	(3) Difference
Agriculture			
Large Domesticated Animals	–0.269***	–0.177***	–1.260
	(0.048)	(0.055)	
Intensive Agriculture	–0.114***	–0.090**	–0.451
	(0.034)	(0.041)	
Plow Use	–0.034	–0.0007	–1.020
	(0.021)	(0.025)	
Female Participation in Agriculture	0.230***	0.208***	0.230
	(0.064)	(0.071)	
Urbanization			
Log Population Density (Murdock)	–0.519**	–0.429**	–0.321
	(0.212)	(0.183)	
Institutions			
Indigenous Slavery	0.096**	0.102**	–0.094
	(0.040)	(0.050)	
Centralization	–0.091**	–0.106**	0.244
	(0.043)	(0.044)	

Notes: OLS estimates of Equation (1) using the Murdock *Map* polygons sample in column (1) and Thiessen Polygons sample in column (2). The sample is limited to groups located between the Tropics of Capricorn and Cancer. Each cell in column (1) and (2) represents a separate regression and the coefficient on the TSI is reported. The t-statistic on the difference between these two coefficients is reported in column (3). The dependent variable is noted in the leftmost column. With the exception of the urbanization category, the data for the dependent variable are from the *Ethnographic Atlas,* and are based on anthropological observations from the late 19th and early 20th century. The data for population density are from Murdock's book *Africa Its Peoples and Their Culture History.* The climate variables are from the 20th century reanalysis for the year 1871. All specifications include temperature, relative humidity and the first-order interaction between temperature and humidity as well as mean altitude, the FAO's agricultural suitability index, access to waterways, the malaria ecology index developed by Kiszewski et al., (2004), absolute latitude and longitude. Standard errors clustered by cultural province. * ** *** Significant at 10, 5 and 1 percent levels.

Table V

PLACEBO TEST			
	(1)	(2)	(3)
Dependent Variable	Main Effect TSI (β)	Africa Interaction TSI (δ)	Africa Total TSI (β+δ)
Agriculture			
Large Domesticated Animals	0.036	–0.214***	–0.177***
	(0.030)	(0.039)	(0.029)
Intensive Agriculture	–0.015	–0.075*	–0.090***
	(0.041)	(0.043)	(0.022)
Plow Use	0.069**	–0.070*	–0.0007
	(0.030)	(0.035)	(0.019)
Female Participation in Agriculture	–0.039	0.247***	0.208***
	(0.065)	(0.088)	(0.063)
Institutions			
Indigenous Slavery	–0.003	0.105**	0.102***
	(0.042)	(0.049)	(0.020)
Centralization	0.010	–0.116**	–0.106**
	(0.027)	(0.051)	(0.049)

Notes: OLS estimates of Equation (2). The dependent variable is noted in the leftmost column. The sample is limited to ethnic groups located between the Tropics of Capricorn and Cancer. Ethnic group boundaries are constructed using Thiessen polygons. Column (1) reports the coefficient on TSI. Column (2) reports the coefficient on the TSI*Africa interaction. Column (3) reports the coefficient of column (1) + column (2). Robust standard errors clustered at the ethnic language family in parentheses. * ** *** Significant at 10, 5 and 1 percent levels.

Table VI

SIMULATION			
	(1)	(2)	(3)
	Africa	Africa	
Dependent Variable	Baseline TseTse	Reduced TseTse	Eurasia
Agriculture			
Large Domesticated Animals	0.60	0.84	0.83
Intensive Agriculture	0.28	0.36	0.72
Plow Use	0.03	0.11	0.72
Female Participation in Agriculture	0.47	0.19	0.13
Institutions			
Indigenous Slavery	0.86	0.74	0.23
Centralization	0.30	0.43	0.63

Notes: This table is created by using OLS to predict the outcomes for the Tropical African sample using Equation (2). The mean values of these predicted outcomes are shown above in column (1). Then each element in the TSI*Africa vector is reduced by one standard deviation. The mean values of the predicted outcomes with the new TSI*Africa vector are shown in column (2). For comparison, the weighted average for Eurasia (e.g., with equal weights for the average from each continent) is shown in column (3).

Table VII

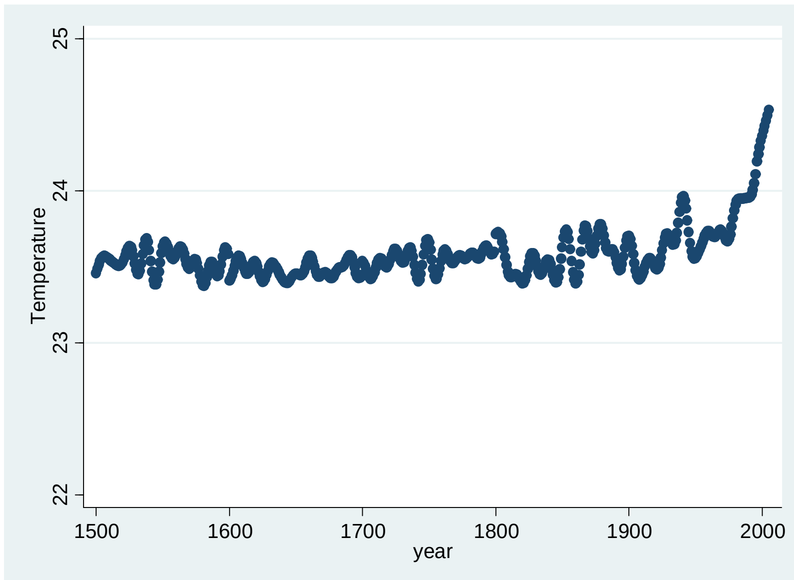
REDUCED FORM ESTIMATES OF THE RELATIONSHIP BETWEEN MODERN ECONOMIC DEVELOPMENT AND TSETSE SUITABILITY

	(1)	(2)	(3)	(4)	(5)
Panel A: Dependent Variable is the Log Mean Luminosity					
TSI	–0.480**	–0.441*	–0.744***	–0.452*	–0.296
	(0.236)	(0.234)	(0.228)	(0.252)	(0.246)
Historical Centralization					1.083***
					(0.247)
Panel B: Dependent Variable is the Log Number of Cattle					
TSI	–1.270**	–1.172**	–1.491***	–0.639*	–0.648*
	(0.473)	(0.447)	(0.390)	(0.320)	(0.323)
Historical Centralization					–0.060
					(0.319)
Climate Controls	Y	Y	Y	Y	Y
Malaria Index	N	Y	Y	Y	Y
Other Geographic Controls	N	N	Y	Y	Y
Country Fixed Effects	N	N	N	Y	Y
No. Obs	665	665	665	665	665
No. Clusters	48	48	48	48	48

Notes : This table reports OLS estimates of the TSI on log (mean luminosity + 0.01) for the year 2008 in panel (A) and log (number of cattle + 1) for the year 2005 in panel (B). Satellite light data are from NOAA and cattle data are from the FAO. Each column of each panel represents a separate regression and the cell reports the coefficient on the TSI or the coefficient on a measure of district-level precolonial historical centralization. Climate controls include temperature, relative humidity the proportion of land area in the Tropics as well as the first-order interaction of temperature and humidity. The malaria ecology index was developed by Kiszewski et al., (2004). Other geography controls include mean altitude, the FAO's agricultural suitability index, absolute latitude, longitude, a binary indicator for whether a province is located adjacent to an inland body of water exceeding 500 square kilometers in area and for proximity to the coast. Historical centralization captures the proportion of a district's current inhabitants whose ancestors lived in a centralized society and is based on data from the *Ethnographic Atlas* . Robust standard errors clustered by country are in parentheses. * ** *** Significant at 10, 5 and 1 percent levels.

THE EFFECT OF THE TSETSE FLY ON AFRICAN DEVELOPMENT
Marcella Alsan
ONLINE APPENDIX
APPENDIX A: ADDITIONAL FIGURES & TABLES

Figure A.I: Africa's Temperature Over the Long Run



Notes: Figure constructed using paleoclimatic data on temperature from Mann et al., (2008).

Figure A.II: Correlation Matrix of the Historical TSI with Other Geographic and Climate Covariates

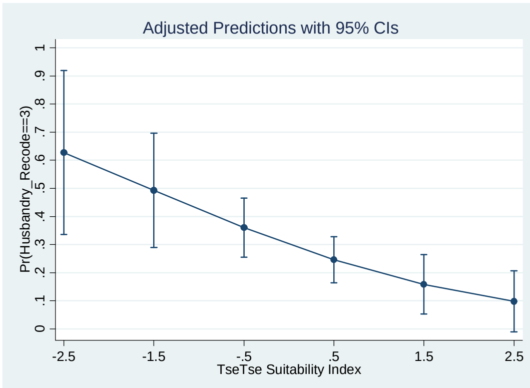
	TSI	Temp	RH	Itx	Mal 1900	Mal Ecology	Abs Lat	Longitude	River	Coast	Alt	Ag Suit	Prop Tropics
Historical TSI	1.00												
Historical Temperature	0.38	1.00											
Historical RH	0.36	–0.45	1.00										
Interaction	0.67	0.04	0.87	1.00									
Malaria 1900	0.41	0.24	0.41	0.59	1.00								
Malaria Ecology Index	0.39	0.70	0.00	0.37	0.45	1.00							
Absolute Latitude	–0.40	–0.08	–0.52	–0.61	–0.52	–0.25	1.00						
Longitude	–0.21	–0.46	0.14	–0.10	–0.20	–0.38	–0.22	1.00					
River	0.10	–0.05	0.19	0.18	0.18	0.12	–0.11	0.03	1.00				
Coast	0.04	0.00	0.06	0.07	–0.12	–0.16	0.30	–0.21	–0.03	1.00			
Mean Altitude	–0.34	–0.53	0.00	–0.29	–0.26	–0.50	0.04	0.43	0.12	–0.19	1.00		
Agricultural SI	0.31	0.15	0.38	0.51	0.56	0.49	–0.43	–0.07	0.18	–0.15	–0.23	1.00	
Proportion Tropics	0.33	0.36	0.18	0.36	0.39	0.37	–0.77	0.09	0.08	–0.32	–0.10	0.36	1.00

Figure A.III: Sample and Clusters

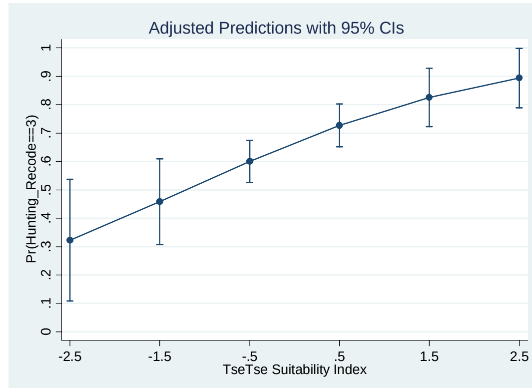
Notes: The sample of ethnic groups from the *Ethnographic Atlas* employed in the analysis is shown above. Shading is used to represent the 44 clusters.

Figure A.IV: Subsistence Patterns and TseTse Suitability

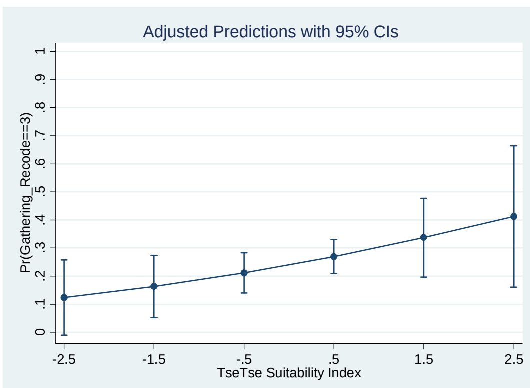
A. Husbandry



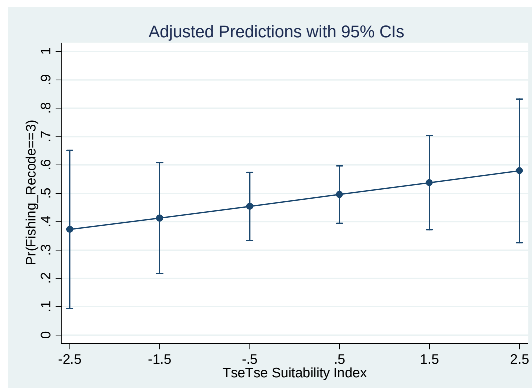
B. Hunting



C. Gathering



D. Fishing



Notes: These graphs show the relationship between TseTse suitability and subsistence patterns. The predicted probability that an ethnic group falls into the top quartile of dependence on a given subsistence strategy is plotted along with the 95 percent confidence interval after estimation using an ordered logit.

Table A.I

SUMMARY STATISTICS FOR HISTORICAL ANALYSIS			
Variable	*Mean*	*Std. Dev.*	*N*
Large Domesticated Animals	0.63	0.48	484
Female Participation in Agriculture	0.49	0.50	315
Intensive Agriculture	0.32	0.47	485
Plow Use	0.08	0.27	484
Indigenous Slavery	0.85	0.36	446
Centralization	0.33	0.47	467
Log Population Density (Murdock) inhab/km^2	1.70	1.59	398
Mean Historical Temperature (Celsius)	24.44	3.11	522
Mean Historical Relative Humidity (Percent)	57.50	14.60	522
Proportion of Land Area in the Tropics	0.93	0.24	522
Longitude	17.81	15.76	522
Absolute Latitude	9.71	7.44	522
Mean Altitude (kilometers)	0.34	0.33	522
Agricultural Suitability Index	0.53	0.20	522
Coast	0.15	0.36	522
River	0.57	0.49	522
Malaria Ecology Index	13.63	9.62	522
Historical TSI (1871)	0.01	1.00	522

SUMMARY STATISTICS FOR MODERN ANALYSIS			
Variable	*Mean*	*Std. Dev.*	*N*
Log (Luminosity + 0.01) (2008)	–2.12	2.41	665
Log (Cattle + 1) (2005)	9.70	3.22	665
Historical Centralization	0.57	0.44	665
Mean Temperature (Celsius)	23.30	4.10	665
Mean Humidity (Percent)	62.15	12.79	665
Proportion of Land Area in the Tropics	0.79	0.40	665
Longitude	18.26	17.20	665
Absolute Latitude	13.88	10.96	665
Mean Altitude (kilometers)	0.65	0.53	665
Agricultural Suitability Index	0.44	0.22	665
Coast	0.23	0.42	665
Inland Water Body	0.14	0.35	665
Malaria Ecology Index	10.74	9.39	665
Modern TSI (1961-1990)	–0.08	1.10	665

Table A.II

AES coefficients		
Grouping	Entire Sample	Tropics Only
Agriculture	–0.350***	–0.392***
	(0.044)	(0.060)
Urbanization	–0.272***	–0.232***
	(0.081)	(0.077)
Institutions	–0.193***	–0.209***
	(0.054)	(0.054)

Notes: Columns give AES estimates for the full and Tropics only subsample. The AES averages the normalized treatment effects obtained from a seemingly unrelated regression. *Agriculture* includes the following variables: use of large domesticated animals, plow use, female participation in agriculture and intensive agriculture. *Urbanization* includes the log population density from Murdock and the presence of a city in 1800 and *Institutions* include indicator variables for political centralization and the practice of indigenous slavery. All specifications include the full set of climate and geography controls in Table I column (4). Standard errors are clustered at the level of cultural province. The control group is defined as all ethnic groups with a TSI<0. The signs of female participation and indigenous slavery are reversed in order to compute the index. * ** *** Significant at 10, 5 and 1 percent levels.

Table A.III

SENSITIVITY ANALYSIS

Dependent Variable	(1) ψ=0.5	(2) ψ=1.0	(3) ψ=1.5	(4) ψ=2.0
Agricuture				
Large Domesticated Animals	–0.202***	–0.221***	–0.235***	–0.237***
	(0.040)	(0.041)	(0.043)	(0.043)
Intensive Agriculture	–0.084***	–0.090***	–0.088***	–0.081***
	(0.030)	(0.028)	(0.027)	(0.029)
Plow Use	–0.041*	–0.052**	–0.062***	–0.071***
	(0.021)	(0.022)	(0.022)	(0.024)
Female Participation in Agriculture	0.182***	0.197***	0.208***	0.202***
	(0.054)	(0.058)	(0.062)	(0.064)
Urbanization				
Log Population Density (Murdock)	–0.674***	–0.711***	–0.765***	–0.804***
	(0.217)	(0.225)	(0.228)	(0.225)
Institutions				
Indigenous Slavery	0.093**	0.098**	0.102**	0.100**
	(0.037)	(0.038)	(0.042)	(0.046)
Centralization	–0.068**	–0.074**	–0.073**	–0.068*
	(0.033)	(0.034)	(0.036)	(0.038)

Notes: OLS estimates of Equation (1). The dependent variable is noted in the leftmost column. Each cell is the coefficient on the Historical TSI as the parameter ψ in the equation for the steady state TseTse fly population is varied over the feasible range as identified by May et. al., (1974). All specifications include temperature, relative humidity the first-order interaction between temperature and humidity and the proportion of land area in the Tropics, as well as mean altitude, the FAO's agricultural suitability index, access to waterways, the malaria ecology index developed by Kiszewski et al., (2004), absolute latitude and longitude. Robust standard errors clustered at the level of cultural province in parentheses. * ** *** Significant at 10, 5 and 1 percent levels.

Table A.IV

REDUCED FORM ESTIMATES OF THE RELATIONSHIP BETWEEN CEREAL CULTIVATION, MILKING, CITIES AND TSETSE SUITABILITY

Dependent Variable	(1)	(2)	(3)	(4)
Cereal	–0.103***	–0.082**	–0.082**	–0.070**
	(0.035)	(0.033)	(0.033)	(0.030)
Milking	–0.121***	–0.143***	–0.141***	–0.177***
	(0.040)	(0.040)	(0.039)	(0.031)
Presence of an Urban	–0.035*	–0.031*	–0.028*	–0.025*
Center (1800)	(0.018)	(0.017)	(0.016)	(0.014)
Climate controls	Y	Y	Y	Y
Malaria control	N	Y	Y	Y
Waterway controls	N	N	Y	Y
Geography controls	N	N	N	Y

Notes: OLS estimates of Equation (1) are reported. The dependent variables are listed in the leftmost column of the upper panel and include whether the ethnic group practiced cereal cultivation, whether the ethnic group milked their livestock frequently and whether the ethnic group had an urban center in 1800. Each cell represents a separate regression and the coefficient on the TSI is reported. The climate variables are from the 20th century reanalysis for the year 1871. Climate controls refer to temperature, relative humidity, the first-order interaction between temperature and humidity and the proportion of land area in the Tropics. Malaria refers to the malaria ecology index developed by Kiszewski et al., (2004). Waterway controls include whether a river was located within the ethnic group boundaries and whether the boundaries included a coast. Geography controls include mean altitude, the FAO's agricultural suitability index, longitude and absolute latitude. Robust standard errors clustered at the level of cultural province in parentheses. * ** *** Significant at 10, 5 and 1 percent levels.

Table A.V

BALANCED SAMPLE	
Dependent Variable	(1)
Agriculture	
Large Domesticated Animals	–0.233***
	(0.052)
Intensive Agriculture	–0.069
	(0.041)
Plow Use	–0.048
	(0.029)
Urbanization	
Log Population Density (Murdock)	–0.821***
	(0.245)
Institutions	
Indigenous Slavery	0.123**
	(0.049)
Centralization	–0.112**
	(0.051)
No. Obs.	318
No. Clusters	42

Notes: The dependent variable is noted in the leftmost column. Each cell represents a separate regression and the coefficient on the TSI is reported. The sample is balanced and includes all outcomes except female participation in agriculture. (Including female participation in agriculture reduces the sample size to 216 though the results are qualitatively similar). Column (1) reports OLS estimates of Equation (1). All specifications include temperature, relative humidity the first-order interaction between temperature and humidity the proportion of land area in the Tropics, as well as mean altitude, the FAO's agricultural suitability index, access to waterways, the malaria ecology index developed by Kiszewski et al., (2004), absolute latitude and longitude. Standard errors clustered at the level of cultural province in parentheses. * ** *** Significant at 10, 5 and 1 percent levels.

Table A.VI

VIRTUAL COUNTRIES	
Dependent Variable	(1)
Agriculture	
Large Domesticated Animals	–0.155***
	(0.043)
Intensive Agriculture	–0.121**
	(0.041)
Plow Use	–0.144**
	(0.057)
Females in Agriculture	0.329***
	(0.065)
Urbanization	
Log Population Density (Murdock)	–1.028***
	(0.272)
Institutions	
Indigenous Slavery	0.051
	(0.065)
Centralization	–0.080*
	(0.042)

Notes: OLS estimates of Equation (1). Each cell represents a separate regression and the coefficient on the TSI is reported. The sample is comprised of virtual countries. Each country is the shape of a square approximately 160,000 $kilometers^2$ in size. Ethnic group outcomes are averaged within the virtual country boundaries. Unweighted results are reported, but similar results are obtained when weighting each average outcome by the number of observations comprising the average. All regressions include temperature, relative humidity the first-order interaction between temperature and humidity, the proportion of land area in the Tropics, as well as mean altitude, the FAO's agricultural suitability index, access to waterways, the malaria ecology index developed by Kiszewski et al., (2004), absolute latitude and longitude. Conley standard errors in parentheses with cutoffs of 40 degrees latitude and longitude. * ** *** Significant at 10, 5 and 1 percent levels.

Table A.VII

INCLUDING SPATIAL LAGS OF TSETSE SUITABILITY

Dependent Variable	(1)	(2)
Agriculture	Own TSI	Spatial Lag TSI
Large Domesticated Animals	–0.178***	–1.394**
	(0.038)	(0.673)
Intensive Agriculture	–0.051	–0.154
	(0.036)	(1.046)
Plow Use	–0.042*	0.290
	(0.023)	(0.436)
Females in Agriculture	0.130**	0.236
	(0.064)	(0.922)
Urbanization		
Log Population Density (Murdock)	–0.876***	4.173*
	(0.220)	(2.448)
Institutions		
Indigenous Slavery	0.097**	0.758
	(0.038)	(0.503)
Centralization	–0.093**	0.101
	(0.044)	(0.719)

Notes: OLS estimates of a modified version of Equation (1) with the spatial lag TSI included. The dependent variable is listed in the leftmost column. Each *row* represents a separate regression and the coefficient on the ethnic group's own TSI is reported in column (1), while the coefficient on the spatial lag TSI is reported in column (2). See footnote (29) for details. All specifications include climate controls (temperature, relative humidity, the first-order interaction between temperature and humidity and the proportion of land area in the Tropics), a control for malaria and controls for mean altitude, FAO's agricultural suitability index, longitude, absolute latitude and access to waterways and their spatial lags. Coefficients are reported with robust standard errors clustered at the level of cultural province in parentheses. * ** *** Significant at 10, 5 and 1 percent levels.

Table A.VIII

ADDITIONAL ROBUSTNESS TESTS OF THE RELATIONSHIP BETWEEN THE REDUCED FORM ESTIMATES OF TSETSE SUITABILITY AND HISTORICAL AFRICAN DEVELOPMENT

Dependent Variable	(1) Cultural Province F.E.	(2) Region F.E.	(3) Probit	(4) Alternative Malaria Index	(5) TSI Squared	(6) Plow Positive Crops	(7) Crop Suitabilities	(8) Ethnic Group Land Area	(9) Higher Order Terms	(10) Export Slave Trade
Agriculture										
Large Domesticated Animals	–0.124**	–0.202***	–0.191***	–0.259***	–0.232***	–0.228***	–0.178***	–0.218***	–0.242***	–0.231***
	(0.046)	(0.041)	(0.040)	(0.047)	(0.044)	(0.051)	(0.044)	(0.049)	(0.071)	(0.042)
Intensive Agriculture	–0.057*	–0.078**	–0.090***	–0.104***	–0.095***	–0.113***	–0.083**	–0.073**	–0.104*	–0.090***
	(0.029)	(0.038)	(0.026)	(0.026)	(0.030)	(0.033)	(0.035)	(0.032)	(0.057)	(0.028)
Plow Use	–0.037*	–0.037*	–0.025	–0.063**	–0.053**	–0.083***	–0.059***	–0.049**	0.012	–0.057**
	(0.021)	(0.020)	(0.015)	(0.023)	(0.022)	(0.029)	(0.021)	(0.022)	(0.033)	(0.023)
Female Participation in	0.114*	0.161***	0.179***	0.249***	0.206***	0.209***	0.210***	0.180**	0.138	0.206***
Agriculture	(0.067)	(0.054)	(0.056)	(0.062)	(0.060)	(0.056)	(0.067)	(0.067)	(0.087)	(0.061)
Urbanization										
Log Population Density	–0.629**	–0.734***	N.A.	–0.681***	–0.749***	–0.418*	–0.564***	–0.572***	–0.409	–0.741***
	(0.293)	(0.210)		(0.238)	(0.240)	(0.208)	(0.201)	(0.192)	(0.281)	(0.224)
Institutions										
Indigenous Slavery	0.084*	0.131***	0.068**	0.097**	0.104**	0.070	0.087*	0.097**	0.070	0.100**
	(0.049)	(0.036)	(0.035)	(0.038)	(0.040)	(0.045)	(0.043)	(0.040)	(0.076)	(0.042)
Centralization	–0.092*	–0.052	–0.074**	–0.058	–0.077**	–0.071*	–0.083*	–0.108***	–0.138**	–0.075**
	(0.051)	(0.042)	(0.034)	(0.036)	(0.036)	(0.041)	(0.043)	(0.039)	(0.064)	(0.036)

Notes: The dependent variable is noted in the leftmost column. Each cell represents a separate regression and the coefficient on a measure of TseTse is reported. With the exception of the urbanization category, the data for the dependent variable are from the *Ethnographic Atlas*, and are based on anthropological observations from the late 19th and early 20th century. The data for population density are from Murdock's book *Africa Its Peoples and Their Culture History*. The climate variables are from the 20th century reanalysis for the year 1871. All specifications include temperature, relative humidity, the first-order interaction between temperature and humidity, the proportion of land area in the Tropics, as well as mean altitude, the FAO's agricultural suitability index, access to waterways, the malaria ecology index developed by Kiszewski et al., (2004), absolute latitude and longitude. Column (1) includes cultural province fixed effects. Column (2) includes region fixed effects for the following regions: North, South, Central, East and West. Column (3) estimates Equation (1) using a probit and reports marginal effects. Column (4) uses a measure of malaria from 1900 (Hay et al., 2004) instead of the malaria ecology index. Column (5) includes TSI squared, the coefficient on the squared term (not reported), is never significant. Column (6) includes the number of crops that can potentially be grown in the ethnic group's area that usually involves a plow (see Alesina, Giuliano and Nunn (2013) for a description). Column (7) includes suitability for other crops (maize, pearl millet, rice, cereal and sorghum). Column (8) controls for the size of the ethnic homeland. Column (9) includes higher order terms in all the climate variables (temperature squared, humidity squared, the interaction of temperature and humidity squared and the proportion of land area in the Tropics squared). Column (10) includes the log {(slave exports+1)/(land area)} as a control (see Nunn and Wantchekon (2011) for a description). Robust standard errors clustered at the level of cultural province in parentheses for columns (1)-(2) and columns (4)-(10). Delta method used to compute standard errors for marginal effects. * ** *** Significant at 10, 5 and 1 percent levels.

Table A.IX

OLS AND IV ESTIMATES OF THE RELATIONSHIP BETWEEN OBSERVED TSETSE AND HISTORICAL AFRICAN DEVELOPMENT

Dependent Variable	Large Domesticated Animals	Intensive Agriculture	Plow Use	Female Participation in Agriculture	Indigenous Slavery	Centralization	Log Population Density (Murdock)
Panel A: Two-Stage Least-Squares with TSI as the Instrument							
Observed TseTse	–1.000***	–0.607***	–0.271**	1.216***	0.876***	–0.362*	–3.921**
	(0.249)	(0.201)	(0.133)	(0.392)	(0.308)	(0.208)	(1.540)
Panel B: First Stage							
TseTse Suitability Index (1871)	0.164***	0.165***	0.164***	0.175***	0.155***	0.163***	0.155***
	(0.024)	(0.024)	(0.024)	(0.033)	(0.022)	(0.025)	(0.027)
F-statistic	37.17	36.90	37.17	39.56	26.22	36.84	19.21
Panel C: Ordinary Least Squares							
Observed TseTse (1973)	–0.449***	–0.186	–0.051	0.231**	0.152	–0.064	–0.802***
	(0.121)	(0.164)	(0.034)	(0.092)	(0.103)	(0.107)	(0.269)
No. Obs	407	408	407	264	369	393	336
No. Clusters	42	42	42	41	42	42	40

Notes: Panel (A) reports the two-stage least-squares estimates. The dependent variable is noted in row (1). Panel (B) reports the corresponding first stage. Panel (C) reports the OLS coefficient from a regression of the outcome in row (1) on the fraction of tribal land observed to be TseTse-infested in 1973 (Ford and Katondo (1977) and the International Livestock Research Institute). Control variables (not reported to save space) include climate controls (temperature, relative humidity the first-order interaction of temperature and humidity and the proportion of land area in the Tropics), the malaria ecology index developed by Kiszewski et al., (2004) and other geography controls (mean altitude, the FAO's agricultural suitability index, longitude, absolute latitude and access to waterways). The sample is limited to countries where data on the presence or absence of the TseTse are available from sources other than colonial surveillance. Coefficients are reported with robust standard errors clustered at the level of cultural province in parentheses. * ** *** Significant at 10, 5 and 1 percent levels.

APPENDIX B: DATA SOURCES

Climate Data and the TseTse Suitability Index. The TSI is constructed using global, gridded daily climate variables from the 20th Century Reanalysis version 2.0: http://www.esrl.noaa.gov/psd/data/20thC_Rean. 20CRv2 is a retrospective analysis produced by the National Oceanic and Atmospheric Administration's (NOAA) Earth Science Research Laboratory Physical Sciences Division in collaboration with the University of Colorado CIRES Climate Diagnostics Center. The 20CRv2 is the earliest climate data set available (at a 2° spatial resolution) that covers Africa and includes the indicators necessary for constructing the TSI. The 20CRv2 uses advanced assimilation methods (the Kalman Ensemble technique) to develop a more accurate representation of late 19th century weather. Earlier reanalyses had difficulty recreating weather for historical periods since upper-air observations typically increased over time and data assimilation methods could not adequately adjust for variation in observation networks. The single level daily data file for air temperature and relative humidity in the earliest year of available data, 1871, was used in this study. The interested reader is referred to Compo et al. (2011) for more information on the reanalysis technique. Modern climate data used in the analysis are from the East Anglia Climate Research Unit and span the period 1961-1990: http://www.cru.uea.ac.uk/cru/data/hrg/tmc.

Population and Urbanization Data. City location geospatial data are from Chandler (1987) for the year 1800. Cities are defined by Chandler as locations with over 20,000 inhabitants. Population data are estimated by Murdock (1959a) for African ethnic groups. Population density is defined as logarithm (inhabitants per square kilometer). The gridded population data for the year 2007 are taken from LandScan™ http://www.ornl.gov/sci/landscan. The LandScan algorithm uses remote sensing spatial data and imagery analysis technologies to disaggregate census counts within an administrative boundary and estimate an ambient population (e.g., an average population over 24 hours).

Geographic Variables: GTOPO 30 is a digital elevation model of the world, developed by the United States Geological Survey (USGS). http://eros.usgs.gov. Elevation is calculated in kilometers. Data on the location of rivers were obtained from the Harvard GIS Center: http://www.gis.harvard.edu/icb/icb.do. Coast and distance to the coast were calculated using the near tool in ArcMap version 10.1.The malaria ecology index by Kiszweski et al. (2004) is used to approximate the prevalence of severe forms of malaria. This index is derived from an equation relating the human-biting tendency of the mosquito to the daily mortality rate. The parameters for the index are taken from field studies. Biting force is proxied for by the number of mosquitoes in a given area that have evidence of a human bloodmeal and mortality is based on the climatic limits for *Anopheles* survival. There is also an adjustment for the type of mosquito most prevalent in a particular region. A coarser malaria map from the year 1900 is used in a robustness check (Hay et al., 2004). Agricultural suitability refers to suitability of land for rainfed crops and is accessible from the FAO's Global Agro-Ecological Zones (2000) website: http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm. The FAO methodology characterizes the climate, soil and terrain conditions relevant to agricultural production and compares these requirements with observed conditions to develop a global data set of maximum potential crop yields under varying input levels. This index has been described in detail by Nunn and Qian (2011). Similar to their analysis, suitability for rainfed crops is used in this empirical exercise to approximate historical agricultural conditions. The suitability index is normalized to range from 0 to 1 with higher values indicating greater suitability.

Slave Exports data are from Nunn and Wantchekon (2011) and measure the number of slaves taken from each ethnic group between years t-1 and t. Estimates begin in 1400. The logarithm of 1+ the total slaves exported per ethnic group, normalized by ethnic group land area, is used as a measure of slave export intensity. This data set can be found at: http://www.economics.harvard.edu/faculty/nunn/data_nunn.

TseTse Presence data are taken from Ford and Katondo (1977)and the International Livestock Research Unit website: http://www.ilri.org/GIS.

Light density at night values are from the Defense Meterological Satellite Program's Operational Linescan System (DMSP-OLS). The files are cloud-free composites made using all available archived DMSP-OLS smooth resolution data for the calendar year of 2008. The data used in this paper contain lights from cities, towns, and other sites with persistent lighting, including gas flares. The Log (0.01 + average luminosity) is used as a dependent variable in the modern analysis. The data can be found online at: http://ngdc.noaa.gov/eog/

Cattle data in Africa (approximate year 2005) were provided by FAO, Agal Division courtesy of Elisa Palamara and Tim Robinson. The Log (1+ number of cattle) is used as a dependent variable in the modern analysis.

Precolonial Ethnographic Outcomes. Ethnographic outcomes are from the *Ethnographic Atlas.* The variables used in this analysis and their definitions are provided in the table below:

Table B.II

ETHNOGRAPHIC ATLAS VARIABLE NAMES AND DEFINITIONS

Varible Name (No.)	*Definition*
Gathering (v1)	Variable ranging from 0 (0-5% dependence) to 9 (86-100% dependence) on gathering
Hunting (v2)	Variable ranging from 0 (0-5% dependence) to 9 (86-100% dependence) on hunting
Fishing (v3)	Variable ranging from 0 (0-5% dependence) to 9 (86-100% dependence) on fishing
Husbandry (v4)	Variable ranging from 0 (0-5% dependence) to 9 (86-100% dependence) on husbandry
Agricultural (v5)	Variable ranging from 0 (0-5% dependence) to 9 (86-100% dependence) on agriculture
Centralization (v33)	Indicator variable equal to one if >=2 levels of hierarchy above the local authority
Cereal Cultivation (v29)	Indicator variable equal to one if major crop type is a cereal grain
Cultural province (chapter heading)	Grouping based on common cultural/genealogical attributes and spatial proximity
Female participation in agriculture (v54)	Indicator variable equal to one if females perform the majority of agricultural tasks
Indigenous slavery (v70)	Indicator variable equal to one for incipient/reported/hereditary slavery
Language family (v98)	Linguistic affiliation
Milking (v41)	Indicator variable equal to one for milking more often than sporadically
Plow use (v39)	Indicator variable equal to one for aboroginal or adopted plow use
Presence of Large Domesticated Animals (v40)	Indicator variable equal to one for presence of bovines, deer, camelids or equines
Intensive agriculture (v28)	Indicator variable equal to one for intensive or intensive, irrigated agriculture

APPENDIX C: INSECT PHYSIOLOGY AND THE TSETSE SUITABILITY INDEX

This appendix provides further background information on the development of the TseTse Suitability Index (TSI). Other researchers have developed suitability indices for the TseTse (Rogers and Randolph, 1986); however, they were not appropriate for this paper since they used inputs that could be considered endogenous, such as the distribution of cattle. Important for the approach herein is that results from controlled laboratory experiments are used to identify abiotic climate factors that affect insect physiology and thus determine the reproductive success of the fly. In general, insects are particularly sensitive to changes in the conditions of their environment due to their large surface area to volume ratio, which affects thermoregulation and water balance (Schowalter, 2011). The TseTse is distinguished from most other insects by its method of reproduction, known as adenotrophic viviparity, which results in a low number of offspring production per adult female. After ovulation, the egg develops in the uterus of the female until reaching the third instar larva stage (Jackson, 1949). Larva are deposited onto the soil where they quickly burrow and encapsulate to form pupa. The pupa have fat stores that last for the approximately month-long period needed to metamorphosize into adult flies (Glasgow, 1963).

Pupa metabolism is highly temperature dependent. Extremes of temperature lead the pupa to metabolize lipid reserves too quickly or slowly, thereby exhausting energy stores before metamorphosis is complete. Figure II panel (A), adapted from Bursell (1960) and Rajagopal and Bursell (1965), demonstrates pupa survival as a function of temperature. The dots represent data from lab experiments on hundreds of flies and a line representing the best quadratic fit is shown. At temperatures below 22 degrees Celsius, adult flies lapse into a chill coma (Terblanche et al., 2008). TseTse flies, especially at the young, teneral stage, are susceptible to dessication at low humidities (Mellanby, 1937). The saturation deficit is a combination of humidity and temperature–it captures the difficulty organisms have in transpiration in hot and humid weather. Figure II panel (B) shows that at lower humidities (higher saturation deficits) fly mortality increases. Teneral comes from *tener,* the Latin word for tender, and describes the softness of the fly body and immaturity of its thoracic musculature prior to the first blood meal. Insects are particularly vulnerable to desiccation during molts (Schowalter, 2011). These physiological relationships are combined in a closed population growth model of the fly. The equations used in the model of TseTse fly population growth are provided in the table below:

Table C.I

FORMULAE FOR THE POPULATION GROWTH MODEL OF THE TSETSE FLY

Variable	*Formula*
Birth Rate (B)	$B(t) = (-0.0058 \cdot t^2 + .2847 \cdot t - 2.467)$
Adult Fly Mortality (M)	$M(t, h) = (-.0003 \cdot satdef^{£2} + 0.0236 \cdot satdef + .235)$
Rate of Increase (λ)	$max((B - M), 0)$
Steady State Population	$N^* = \left(\frac{\Lambda}{\phi}\right)^{\frac{1}{\psi}}$
TseTse Suitability Index	Z-score of N*

Notes: Temperature is in degrees Celsius and denoted with a *t*. Relative Humidity is a percent and denoted *h* Saturation deficit is denoted *satdef* and is a nonlinear combination of temperature and relative humidity. In the baseline model phi=0.025 and psi=1.25. [£]Mortality was replaced by its maximum value if the mean temperature was less than 22 degrees Celsius to reflect the "chill coma" phenomenon.

REFERENCES FOR THE APPENDIX

Bursell, E. (1960). Loss of Water by Excretion and Defecation in the TseTse Fly. *Journal of Experimental Biology,* 37:689–697.

Chandler, Tertius. (1987). *Four Thousand Years of Urban Growth: An Historical Census* St. David's University Press, Lewiston, New York.

Compo, GP, Whitaker, J.S., and P.D.G. Sardeshmukh. (2011). The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meterological Society*, 137:1–28.

FAO. (2005). Observed Cattle Density in Africa, Rome.

Ford, J., and K.M. Katondo (1977). Maps of Tsetse Flies (Glossina) Distribution in Africa, 1973, according to Sub-Generic Groups on a Scale of 1:5 000 0000. *Bulletin of Animal Health and Production in Africa,* 25:187–193.

Glasgow, J.P. (1963). *The Distribution and Abundance of Tsetse.*International Series of Monographs on Pure and Applied Biology. Macmillan, New York.

C.H.N. Jackson. (1949). The Biology of Tsetse Flies. *Biological Reviews*, 24:174–179.

Kiszweski, Anthony, Mellinger, Andrew, Spielman, Andrew, Malaney, Pia, Sachs, Sonia, and Jeffrey D. Sachs. (2004). A Global Index Representing the Stability of Malaria Transmission. *The American Journal of Tropical Medicine and Hygiene,* 70:486–498.

Mellanby, K. (1937). Experimental work with the Tsetse-fly, *Glossina Palpalis* in Uganda. *Bulletin of Entomological Research,* 28:611–632.

Murdock, George P.(1959a). *Africa its Peoples and Their Culture History.* McGraw-Hill, New York.

Murdock, George P. (1959b). *Ethnographic Atlas.* University of Pittsburgh Press.

Nunn, Nathan, and Nancy Qian. (2011). The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment. *The Quarterly Journal of Economics*, 126:593–650.

Nunn, Nathan, and Leonard Wantchekon. (2011). The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review*, 101:3221–3252.

Rajagopal, P.K., and E. Bursell. (1965). The Effect of Temperature on the Oxygen Consumption of Tsetse Pupae. *Bulletin of Entomological Research,* 56:219–225.

Terblanche, John S, Clusella-Trullas, Susana, Deere, Jaques, and Steven Chowen. (2008). Thermal tolerance in a South-East African Population of the TseTse Fly *Glossina Pallidipes* (Diptera, Glossinidae): Implications for Forecasting Climate Change Impacts. *Journal of Insect Physiology,* 514:114–127.