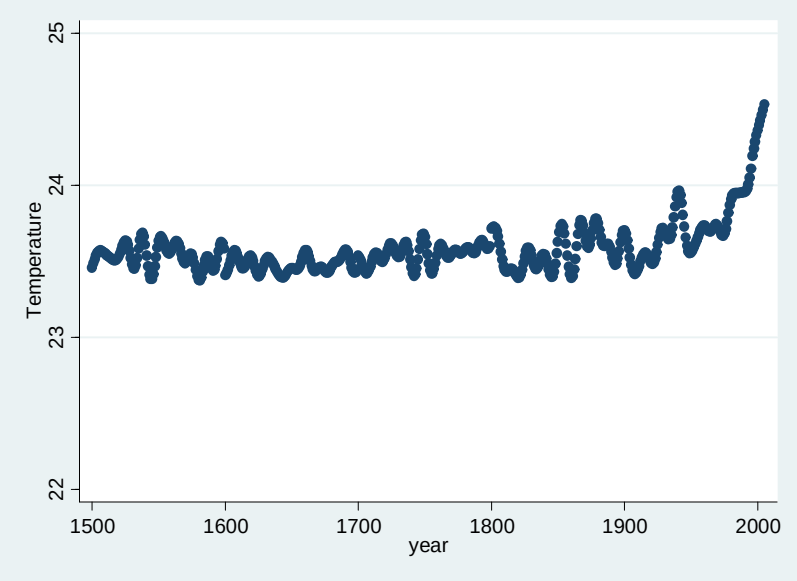
THE EFFECT OF THE TSETSE FLY ON AFRICAN DEVELOPMENT

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

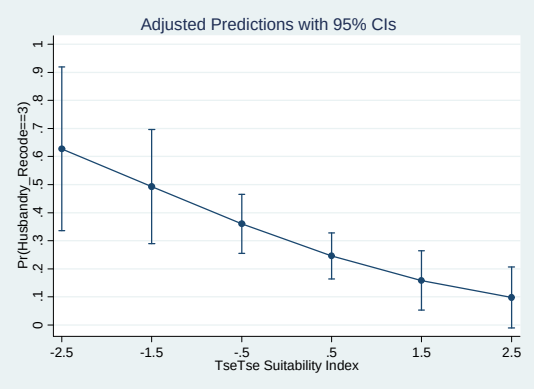
	TSI	Temp	RH	Itx	Mal 1900	Mal Ecology	Abs Lat	Longitude	River	Coast	Alt	Ag Suit	Prop Tropics
Historical TSI	1.00												
Historical Temperature	0.38	1.00											
Historical RH	0.36	–0.45	1.00										
Interaction	0.67	0.04	0.87	1.00									
Malaria 1900	0.41	0.24	0.41	0.59	1.00								
Malaria 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

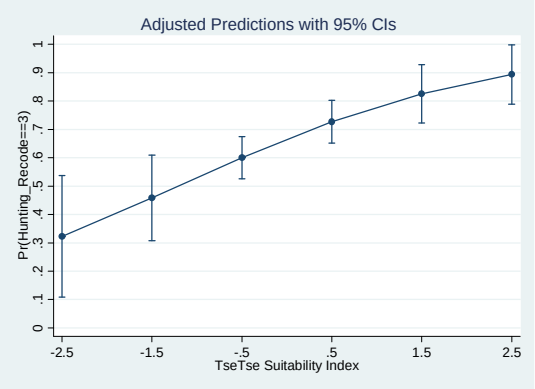
Notes: The sample of ethnic groups from the *Ethnographic Atlas* employed in the analysis is shown above. Shading is used to represent the 44 clusters.

Figure A.IV: Subsistence Patterns and TseTse Suitability

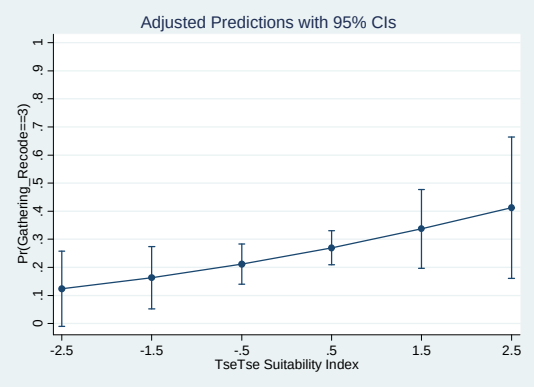
A. Husbandry



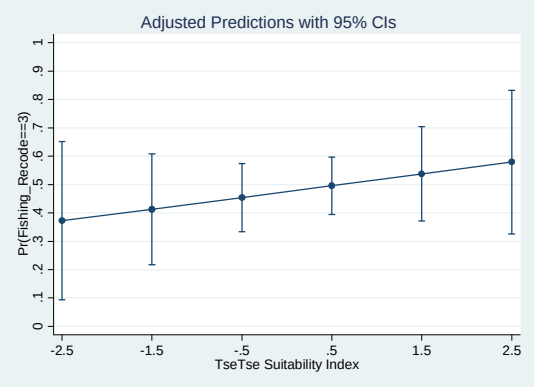
B. Hunting



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)	(2)	(3)	(4)
	ψ=0.5	ψ=1.0	ψ=1.5	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 and Sardeshmukh (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 and Sachs (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 and Snow, 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 and Chown, 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.