Online Appendix for

The Legacy of Colonial Medicine in Central Africa

Sara Lowes

UC San Diego, NBER, and CIFAR

Eduardo Montero

University of Michigan

3 February 2021

Contents

List of Tables

List of Figures

Appendix A. Data Sources and Variable Definitions

A.1. Data and Replication Overview

Replication data and code are available in Lowes and Montero (2020). For the purposes of replication documentation, we list every data set used in the paper, the corresponding citation, access modality, and source location.

By original data, we mean the data exactly as downloaded from the source. By analysis data, we mean data modified for our analysis; this is provided in cases where we are unable to directly provide access to the original data (because of access permissions). We describe in each case the process for acquiring access to the original data; we also provide the do files for creating the analysis data from the original data in Lowes and Montero (2020).

1. **Dataset name**: Aiddata World Bank Geocoded Research Release, Version 1.4.2 (see Appendix A.6)

   - **Citation**: AidData (2017)

   - **Access modality**: The paper uses the World Bank Geocoded Research Release, Version 1.4.2 data, which is publicly available data.

   - **Source location**: The original data sets and documentation can be downloaded at https://www.aiddata.org/data/world-bank-geocoded-research-release-level-1-v1-4-2. For the replication, we provide the original data in the data repository: http://doi.org/10.3886/E120570V1.

   - **Note**: Original data accessed from https://www.aiddata.org/data/world-bank-geocoded-research-release-level-1-v1-4-2 on March 20th, 2017.

2. **Dataset name**: Afrobarometer (various, see Appendix A.5)

   - **Citation**: Afrobarometer (2013, 2015); BenYishay et al. (2017)

   - **Access modality**: The paper uses the Afrobarometer data, which is publicly available data. Users must register to access the data. Instructions to register to access the data are provided here: http://afrobarometer.org/data/data-use-policy.

   - **Source location**: The original data sets and documentation can be downloaded at http://afrobarometer.org/data/geocoded-data. We use Rounds 5 and 6 for Gabon and Cameroon. For the replication, we provide the analysis data in the data repository: http://doi.org/10.3886/E120570V1.

   - **Note**: Original data accessed from http://afrobarometer.org/data/geocoded-data on September 19, 2017.

3. **Dataset name**: Atlantic Slave Trade Data (see Appendix A.8)

   - **Citation**: Nunn and Wantchekon (2011,a)

   - **Access modality**: The paper uses the Atlantic Slave Trade data compiled by Nunn and Wantchekon (2011), which is publicly available data.

   - **Source location**: The original data sets and documentation can be downloaded at https://www.openicpsr.org/openicpsr/project/112479/version/V1/view. For the replication, we provide the original data in the data repository: http://doi.org/10.3886/E120570V1.

4

- **Note**: Original data accessed from https://www.openicpsr.org/openicpsr/project/112479/version/V1/view on October 27, 2015.

4. **Dataset name**: Atlas of the Biosphere: Suitability for Agriculture (see Appendix A.2)

    - **Citation**: Ramankutty et al. (2002)
    - **Access modality**: The paper uses the Atlas of the Biosphere: Suitability for Agriculture data, which is publicly available data.
    - **Source location**: The original data sets and documentation can be downloaded at https://nelson.wisc.edu/sage/data-and-models/atlas/maps.php?datasetid=19&includerelatedlinks=1&dataset=19. For the replication, we provide the original data in the data repository: http://doi.org/10.3886/E120570V1.
    - **Note**: Original data accessed from https://nelson.wisc.edu/sage/data-and-models/atlas/data.php?incdataset=Suitability%20for%20Agriculture on November 1, 2015.

5. **Dataset name**: Demographic and Health Surveys (DHS) and AIDS Indicator Survey (AIS) (various, see Appendix A.3)

    - **Citation**: ICF (1991-2014)
    - **Access modality**: The paper uses DHS and AIS data, which does not allow for redistribution. Users must register to access the data. Instructions to register to access the data are provided here: https://www.dhsprogram.com/data/Access-Instructions.cfm.
    - **Source location**: Once registered, all original DHS data sets and documentation can be downloaded at https://www.dhsprogram.com/Data/.
    - **Note**: Original data accessed from https://www.dhsprogram.com/Data/ in between February 2012 and October 2016 (various dates as new data sets were added to the DHS).

6. **Dataset name**: Ethnographic Atlas Pre-Colonial Data (see Appendix A.8)

    - **Citation**: Murdock (1967); Alsan (2015)
    - **Access modality**: The paper uses the Ethnographic Atlas data, which is publicly available data, merged with the Murdock map boundaries Murdock (1959) by Alsan (2015).
    - **Source location**: The original data sets and documentation can be downloaded at https://www.openicpsr.org/openicpsr/project/112921/version/V1/view. For the replication, we provide the original data in the data repository: http://doi.org/10.3886/E120570V1.
    - **Note**: Original data accessed from https://www.openicpsr.org/openicpsr/project/112921/version/V1/view on October 27, 2015.

7. **Dataset name**: GADM Administrative Boundary Data

    - **Citation**: GADM (2015a,b)
    - **Access modality**: The paper uses the GADM Administrative Boundary data, which is publicly available data.
    - **Source location**: The original data sets and documentation can be downloaded at https://gadm.org/old_versions.html. For AEF countries we use GADM Version 2.8. For Cameroon we use GADM Version 2.5

- **Note**: Original data accessed from https://gadm.org/old_versions.html in July and November 2015.

8. **Dataset name**: Global Agro-Ecological Zones (see Appendix A.2)

   - **Citation**: Fischer et al. (2000)
   - **Access modality**: The paper uses the Global AEZ data, which is publicly available data. However, you must register to gain access.
   - **Source location**: The original data sets and documentation can be downloaded using the FAO data portal at http://gaez.fao.org/Main.html. The datasets downloaded are: res03_crav6190l_sulr_csv for cassava, res03_crav6190l_sulr_fml.tif for foxtail millet and res03_crav6190l_sulr_pml.tif for pearl millet. For the replication, we provide the analysis data in the data repository: http://doi.org/10.3886/E120570V1.
   - **Note**: Original data accessed from http://gaez.fao.org/Main.html on June 10, 2014.

9. **Dataset name**: Hydrosheds Data (see Appendix A.2)

   - **Citation**: Lehner et al. (2008)
   - **Access modality**: The paper uses the Hydrosheds elevation data, which is publicly available data.
   - **Source location**: The original data sets and documentation can be downloaded at https://www.hydrosheds.org/images/inpages/HydroSHEDS_TechDoc_v1_2.pdf. We use the following data set in generic grid format at 30 second resolution for Africa: "af_dem_30s". For the replication, we provide the original data in the data repository: http://doi.org/10.3886/E120570V1.
   - **Note**: Original data accessed from www.hydrosheds.org on November 1, 2015.

10. **Dataset name**: IPUMS DHS (various, see Appendix A.4)

    - **Citation**: IPUMS and ICF (1991-2014)
    - **Access modality**: The paper uses the IPUMS DHS data, which does not allow for redistribution. Users must register to access the data. Instructions to register to access the data are provided here: https://www.idhsdata.org/idhs/index.shtml.
    - **Source location**: Once registered, all original IPUMS DHS data sets and documentation can be downloaded at https://www.idhsdata.org/idhs/index.shtml.
    - **Note**: Original data accessed from https://www.idhsdata.org/idhs/index.shtml on August 27, 2019.

11. **Dataset name**: Malaria Ecology Index (see Appendix A.2)

    - **Citation**: Kiszewski et al. (2004)
    - **Access modality**: The paper uses the Malaria Ecology Index data, which is publicly available data.
    - **Source location**: The original data sets and documentation can be downloaded at https://sites.google.com/site/gordoncmccord/datasets. We use the raster format for GIS. For the replication, we provide the original data in the data repository: http://doi.org/10.3886/E120570V1.
    - **Note**: Original data accessed from https://sites.google.com/site/gordoncmccord/datasets on June 23, 2015.

12. **Dataset name**: Missionary Data (see Appendix A.8)

    - **Citation**: Roome (1924); Nunn (2010)
    - **Access modality**: The paper uses the Roome (1924) data of missionary presence digitized by (Nunn, 2010), which is publicly available data.
    - **Source location**: The original data sets and documentation can be downloaded at https://scholar.harvard.edu/nunn/pages/data-0. We download the Roome (1924) map of the locations of Catholic and Protestant missionaries. For the replication, we provide the analysis data in the data repository: http://doi.org/10.3886/E120570V1.
    - **Note**: Original data accessed from https://scholar.harvard.edu/nunn/pages/data-0 on October 27, 2015.

13. **Dataset name**: Tsetse fly Suitability Index (see Appendix A.2)

    - **Citation**: Alsan (2015,b)
    - **Access modality**: The paper uses the TSI data, which was constructed by Alsan (2015).
    - **Source location**: The original data set was provided by Marcella Alsan. For the replication, we provide the code to construct analysis data from the TSI data in the data repository: http://doi.org/10.3886/E120570V1.

14. **Dataset name**: WorldClim 1.4 Climate Data (see Appendix A.2)

    - **Citation**: Hijmans et al. (2005)
    - **Access modality**: The paper uses the WorldClim 1.4 climate data, which is publicly available data.
    - **Source location**: The original data sets and documentation can be downloaded at https://worldclim.org/data/v1.4/worldclim14.html. We use the following data sets in generic grid format and 10 minute resolution: "tavg_10m" and "prec_10m". For the replication, we provide the original data in the data repository: http://doi.org/10.3886/E120570V1.
    - **Note**: Original data accessed from https://worldclim.org/data/v1.4/worldclim14.html on November 1, 2015.

15. **Dataset name**: World Health Organization Subnational Immunization Coverage data (see Appendix A.7)

    - **Citation**: World Health Organization (2018)
    - **Access modality**: The paper uses the WHO Subnational Immunization Coverage data, which does not allow for redistribution. Users must register to access the data. Instructions to register to access the data are provided here:].
    - **Source location**: The original data sets and documentation can be downloaded at https://www.who.int/immunization/monitoring_surveillance/data/subnational/en/. For the replication, we provide the analysis data in the data repository: http://doi.org/10.3886/E120570V1.
    - **Note**: Original data accessed from vpdata@who.int on September 11, 2019.

A.2. *Geographic Data and Variables*

- **Elevation:** The elevation data are provided by the Hydrosheds Database created by Lehner et al. (2008) and available at https://www.hydrosheds.org/downloads. This data provides elevation information in meters at the 30 arc-second resolution (approximately at the 1 km^2 level near the equator). The elevation measure is constructed using NASAs SRTM satellite images (http://www2.jpl.nasa.gov/srtm/).

- **Precipitation:** Precipitation data are provided by the Global Climate Database created by Hijmans et al. (2005) and available at http://www.worldclim.org/. This data provides monthly average rainfall in millimeters. We calculate the average rainfall for each month in each region of interest and average this over the twelve months to obtain our yearly precipitation measure in millimeters of rainfall per year.

- **Land Suitability:** Land suitability is the soil component of the land quality index created by the Atlas of the Biosphere available at http://www.sage.wisc.edu/iamdata/ used in Michalopoulos (2012) and Ramankutty et al. (2002). This data uses soil characteristics (namely soil carbon density and the acidity or alkalinity of soil) and combines them using the best functional form to match known actual cropland area and interpolates this measure to be available for most of the world at the 0.5 degree in latitude by longitude level. (The online appendix in Michalopoulos (2012) provides a detailed description of the functional forms used to create this dataset.) This measure is normalized to be between 0 and 1, where higher values indicate higher soil suitability for agriculture.

- **Crop Suitability:** Crop suitability refers to the average suitability for rain-fed, low-input crops provided by the FAO's Global Agro-Ecological Zones website: http://www.iiasa.ac.at/Research/LUC/GAEZ/index.htm (Fischer et al., 2000). FAO crop suitability model uses data on elevation, precipitation, soil and slope constraints to construct estimates of crop suitability at the 1 km^2 level for different crops. We normalized the measure to be between 0 and 100, where higher values indicate higher crop suitability.

- **Relative Suitability:** We use the crop suitability for rain-fed, low-input cassava and millet provided by the FAO's Global Ecological Zones Fischer et al. (2000). To construct the millet suitability, we use the maximum suitability of either foxtail and pearl millet. We use the the difference between the logarithm of cassava suitability and the logarithm of millet suitability.[30]

- **Malaria Suitability:** Malaria data uses the Malaria Ecology index created by Kiszewski et al. (2004). The index was created by Kiszewski et al. (2004) to approximate the prevalence of severe forms of malaria. It is created from equations relating the human-feeding tendency of the Anopheles mosquito to the malaria mortality rate using parameters from various field studies and adjusts for the mosquito type that is most prevalent in a region.

- **Tsetse Fly Suitability:** The tsetse suitability index (TSI) is from Alsan (2015). The TSI is constructed by Alsan (2015) using global climate data and parameters from laboratory experiments on the relationship between tsetse fly population birth and death rates and climate variables. The TSI is measured as the Z-score of the potential steady-state tsetse fly population.

---

[30]Specifically, we calculate the measure as: log(Cassava Suitability $+1$) $- log$(Millet Suitability $+1$) as the logarithm of zero is undefined.

A.3. DHS Survey Data and Variables

Survey data on health outcomes for individuals are from the DHS surveys for our countries of interest: Cameroon, Central African Republic, Chad, Republic of Congo, and Gabon (ICF, 1991-2014). We combine DHS survey data, GPS data on cluster locations, and data sets on HIV/other biomarkers, when available. Note that for the GPS locations of DHS clusters, coordinates are displaced by up to 5km for all urban clusters, and 99% of rural clusters and up to 10 km for 1% of rural clusters. Importantly, this displacement is random, and simply induces classical measurement error. The survey data and detailed information on the sampling procedure and variable definitions is available at: http://dhsprogram.com/data/Data-Variables-and-Definitions.cfm.

- **Cameroon:** For Cameroon, we use the 2004 and 2011 Standard DHS surveys, for which there are GPS locations of clusters and HIV/other biomarker data. The 2004 survey was administered between February and August 2004 and the 2011 survey was administered by January and August 2011. Cameroon also has DHS surveys from 1991 and 1998. However, for these years, no HIV/other biomarker data were collected.

- **Central African Republic:** For Central African Republic, we use the 1994-95 Standard DHS survey, for which there are GPS locations of clusters. There is no HIV/other biomarker data available for CAR. The 1994 survey was administered between September 1994 and March 1995. There is 2010 MICS survey for CAR which is not yet publicly available.

- **Chad:** For Chad, we use the 1996-97, 2004, and 2014-15 Standard DHS surveys. The only year for which the GPS locations of clusters is available is 2014-15. HIV/other biomarker data was collected for 1996-1997 (only child height and weight) and for 2014-15. The 1996-97 survey was administered between December 1996 and July 1997, the 2004 survey was administered between July 2004 and December 2005, and the 2014-15 survey was administered between October 2014 and April 2015.

- **Republic of Congo:** For the Republic of Congo, we use the 2009 Standard AIS surveys and 2011-12 Standard DHS surveys. There are no GPS data available for the cluster locations. HIV/other biomarker data are only available for the 2009 AIDS Indicator Survey (AIS). The 2009 AIS survey was administered between March and June 2009, and the 2011-12 DHS survey was administered between September 2011 and February 2012. We do not use the 2005 Standard DHS survey, for which there is neither district data nor HIV/other biomarker data.

- **Gabon:** For Gabon, we use the 2012 Standard DHS survey, for which GPS locations of clusters are available and for which HIV/other biomarker data are available. The 2012 survey was administered between January 2012 and May 2012. We do not use the 2000 Standard DHS survey, for which there is neither GPS data nor HIV/other biomarker data.

Below we explain the variable definitions for the variables used in this paper from the DHS, and AIS surveys:

- **Educational Attainment:** Educational Attainment is a 0 to 3 categorical variable that measures the highest education level attained, where 0 is no education, 1 is primary education, 2 is secondary and 3 is higher education.

- **Wealth Factor:** Wealth Factor is an index generated by the DHS using principle component analysis on asset ownership for each individual.

- **Wealth Index:** Wealth Index is a is a 1 to 5 categorical variable where 1 is poorest quintile and 5 is richest quintile (in the entire DRC 2007 sample) from the Wealth Factor Score.

Table A1: Summary of Available DHS Data and Data in Sample

			Demographic and Health Survey Data			
Country	**Type**	**Year**	**Survey Data**	**GPS Data**	**In Blood Test Sample**	**In Vaccination Index Sample**
Cameroon	Standard DHS	2011	Yes	Yes	Yes	Yes
Cameroon	Standard DHS	2004	Yes	Yes	Yes	Yes
Cameroon	Standard DHS	1998	Yes	Not Available	No	No
Cameroon	Standard DHS	1991	Yes	Yes	No	No
CAR	MICS	2010	Not Available	Not Available	No	No
CAR	Standard DHS	1994-95	Yes	Yes	No	Yes
Chad	Standard DHS	2014	Yes	Yes	Yes	Yes
Chad	Standard DHS	2004	Yes	No	No	Yes
Chad	Standard DHS	1996-97	Yes	No	No	Yes
Congo	Standard DHS	2011-12	Yes	No	No	Yes
Congo	Standard AIS	2009	Yes	No	Yes	Yes
Congo	Standard DHS	2005	Yes	No District	No	No
Gabon	Standard DHS	2012	Yes	Yes	Yes	Yes
Gabon	Standard DHS	2000	Yes	No	No	No

Notes: "Yes" means that the data was collected and is available from DHS. "No" means the data was not collected by the DHS. "Not available" means the data was collected by the DHS but is not made available by DHS. "No District" means that the district information is not reported in addition to there being no GPS data.

- **Hemoglobin Blood Test Refused** is an indicator variable for refusing consent to taking a blood test to test hemoglobin levels.

- **HIV Blood Test Refused** is an indicator variable for refusing consent to taking a blood test to test for HIV.

- **Blood Test Refused** is an indicator variable we define as equal to 1 if an individual refuses consent to take a blood test (either for HIV or hemoglobin levels).

The DHS collects data on health behavior for a subsample of the sampled female population about their children (about a third of the entire sample). The following variables are only defined for this subsample of children:

- **Vaccination Index:** is the share of vaccines reported for children in the DHS out of the 9 possible vaccines included in the DHS survey. The vaccines recorded are the polio 0, bcg, dpt 1, polio 1, dpt 2, polio 2, dpt 3, polio 3, and measles vaccinations.

A.4. IPUMS Survey Data and Variables

For some of the supplemental analyses we also use data from IPUMS DHS, which harmonizes DHS data across countries and survey waves IPUMS and ICF (1991-2014). Data from the IPUMS-DHS Integrated sample for non-French and non-Belgian colonies with GPS coordinates are: Angola (2015), Ethiopia (2000, 2005, 2010, 2016), Ghana (1993, 1998, 2003, 2008), Kenya (2003, 2008, 2014), Lesotho (2004, 2009, 2014), Malawi (2000, 2004, 2010, 2015), Niger(1998), Namibia (2000), Tanzania (1999), and Uganda (2000).

A.5. Afrobarometer Data and Variables

We use Afrobarometer data from Cameroon and Gabon. For Cameroon we use Round 5 and Round 6 data and for Gabon we use Round 6 data (Afrobarometer, 2013, 2015). We use data available from AidData.org to georeference the observations (BenYishay et al., 2017). For each

round we use all of the questions related to trust.

Round 5 Questions - Cameroon:

- **Trust Questions:**

  – "How much do you trust each of the following, or haven't you heard enough about it to say?" The question is asked about: the President, Parliament, the electoral commission, the tax department, the local government council, the ruling party, opposition parties, the police, the army, the courts of law. The response options are Not at all, Just a little, Somewhat, A lot, Haven't heard/ Don't know. [Question 59A-J]

  – "Generally speaking, would you say that most people can be trusted or that you must be very careful in dealing with people?" The response options are most people can be trusted, must be very careful, and don't know. [Question 87]

  – "How much do you trust each of the following types of people?" The response options are Not at all, Just a little, Somewhat, A lot, Don't know. The question is asked about: Your relatives, Your neighbors, Other people you know. [Question 88]

- **Health Questions:**

  – "Are the following services present in the primary sampling unit / enumeration are or in easy walking distance: health center?" The response options are: Yes, No, Can't determine. [EA-FAC D]

  – "Over the past year, how often, if ever, have you or anyone in your family gone without medicines or medical treatment?" The response options are: Never, Just once or twice, Several times, Many times, Always, Don't know. [Question 8C].

Round 6 Questions - Cameroon and Gabon:

- **Trust Questions:**

  – "How much do you trust each of the following, or haven't you heard enough about it to say?" The question is asked about: the President, Parliament, the electoral commission, the tax department, the local government council, the ruling party, opposition parties, the police, the army, the courts of law, traditional leaders, religious leaders. The response options are Not at all, Just a little, Somewhat, A lot, Haven't heard/ Don't know. [Question 52A-L].

- **Health Questions:**

  – "Are the following services present in the primary sampling unit / enumeration are or in easy walking distance: health center?" The response options are: Yes, No, Can't determine. [Question EA-FAC D].

  – "In the last 12 months have you had contact with a public clinic or hospital". The response options are: Yes, No, Don't know. If yes, "How easy or difficult was it to obtain the medical care you needed?" The response options are No contact, Very easy, Easy, Difficult, Very difficult, Don't know. [Question 55C].

  – "Over the past year, how often, if ever, have you or anyone in your family gone without medicines or medical treatment?" The response options are: Never, Just once or twice, Several times, Many times, Always, Don't know. [Question 8C].

A.6. World Bank Project Data and Variables

The World Bank project data was accessed from AidData (2017). We use the following variables from the data:

- **IEG Outcome Rating:** The world bank project outcome rating is defined as "the extent to which the operation's major relevant objectives were achieved, or are expected to be achieved, efficiently" and uses the following six-point scale: Highly Satisfactory, Satisfactory, Moderately Satisfactory, Moderately Unsatisfactory, Unsatisfactory, and Highly Unsatisfactory.

- **Latitude and Longitude:** AidData (2017) records the geocoded location of each project as the longitude and latitude of each project that could be geolocated.

- **Sector:** the data records the (at most) 5 sectors used for each project along with the IEG evaluation rating. We define a project as part of a sector if any of the 5 sector variables include a sector in its values.

- **Total Commitments:** refers to the total amount of funds in dollars committed for each project.

A.7. WHO Sub-National Vaccination Data and Variables

The WHO Sub-National Vaccination data was released for 2016 and 2017 from contacting the WHO (World Health Organization, 2018). The data are describe in detail at: https://www.who.int/immunization/monitoring_surveillance/data/subnational/en/ (last accessed September 11th, 2019). The data contains the vaccination rates for a number of vaccines across different sub-national units for many countries and reported population rates for each sub-national unit. We construct the variable **Vaccination Index:** as the average vaccination coverage for a sub-national unit for all vaccinations reported in a country. We construct the geographic controls for each sub-national unit reported in the WHO data.

A.8. Other Colonial and Precolonial Data and Variables

- **Atlantic Slave Trade**: We use data from Nunn and Wantchekon (2011) on the number of slaves taken from each ethnic group – where ethnic groups are defined using maps from Murdock (1959) – during the Atlantic slave trade.

- **Missionary Locations**: Missionary post location data are from Nunn (2010) and is available at http://scholar.harvard.edu/nunn/pages/data-0 in the form of a GIS shapefile. This shapefile was created by Nathan Nunn by digitizing maps from *Ethnographic Survey of Africa: Showing the Tribes and Languages; also the Stations of Missionary Societies"* published by Roome (1924).

- **Precolonial Data**: Precolonial data are from the Ethnographic Atlas created by Murdock (1967). Not all ethnic groups have data for pre-colonial controls, so we interpolate values using the nearest neighbors within each country and include an indicator for whether or not the ethnic group had pre-colonial data reported in Murdock.

A.9. Colonial Medicine Data

The historical data were collected from the *Service Historique de la Defense* archives in Toulon, France in January 2013. Below we include a list of all of the documents collected from the archives

and their associated call numbers in the archive's catalog. Because Cameroon had special status with the UN, the French were required to submit annual reports to the UN on their activities there. Thus, we also include the information on the UN reports consulted. Many of the reports listed below are hundreds of pages long and include dozens of tables and figures. Therefore, in this appendix we do not list the page numbers of the tables and figures we digitize.

Table A2: Documents Collected From *Service Historique de la Defense*

AEF Reports			
Document Name	**Year**	**Catalog**	**Call #**
AEF 1933 Medical Report	1933	2013 ZK 005	121
AEF 1934	1934	2013 ZK 005	121
AEF 1935 Administrative	1935	2013 ZK 005	121
AEF 1935 Medical	1935	2013 ZK 005	7
AEF 1936 Administrative	1936	2013 ZK 005	121
AEF 1936 Medical	1936	2013 ZK 005	121
AEF 1939 Administrative	1939	2013 ZK 005	160
AEF 1939 Medical	1939	2013 ZK 005	160
AEF 1940 Administrative	1940	2013 ZK 005	160
AEF 1940 Medical	1940	2013 ZK 005	160
AEF 1941 Administrative	1941	2013 ZK 005	160
AEF 1941 Medical	1941	2013 ZK 005	160
AEF 1942 Administrative	1942	2013 ZK 005	160
AEF 1942 Medical	1942	2013 ZK 005	160
AEF 1943 Administrative	1943	2013 ZK 005	160
AEF 1943 Medical	1943	2013 ZK 005	160
AEF 1944 Administrative	1944	2013 ZK 005	160
AEF 1944 Medical	1944	2013 ZK 005	160
AEF 1945 Administrative	1945	2013 ZK 005	89
AEF 1945 Medical	1945	2013 ZK 005	89
AEF 1946 Administrative	1946	2013 ZK 005	89
AEF 1946 Medical	1946	2013 ZK 005	89
AEF 1947 Medical	1947	2013 ZK 005	89
AEF 1947 SGHMP (& misc. documents)	1947	2013 ZK 005	96
AEF 1948 SGHMP	1948	2013 ZK 005	96
AEF 1949 Administrative	1949	2013 ZK 005	91
AEF 1949 Medical	1949	2013 ZK 005	91
AEF 1949 SGHMP	1949	2013 ZK 005	119
AEF 1950 Administrative	1950	2013 ZK 005	91
AEF 1950 Medical	1950	2013 ZK 005	91
AEF 1950-1951 SGHMP	1950-1951	2013 ZK 005	119
AEF 1952 Economic and Social Statistics	1952	2013 ZK 005	90
AEF 1952 Medical	1952	2013 ZK 005	97
AEF 1953 SGHMP	1953	2013 ZK 005	119
AEF 1954 Annual Report Commentaries	1954	2013 ZK 005	119
AEF 1954 SGHMP	1954	2013 ZK 005	16
AEF 1955 SGHMP	1955	2013 ZK 005	100
AEF 1956 SGHMP	1956	2013 ZK 005	100

Table A3: Documents Collected From *Service Historique de la Defense*

Cameroon Reports			
Document Name	**Year**	**Catalog**	**Call #**
Cameroon 1936 Administrative	1936	2013 ZK 005	107
Cameroon 1939 Administrative	1939	2013 ZK 005	107
Cameroon 1939 Medical	1939	2013 ZK 005	107
Cameroon 1940 Annual Report	1940	2013 ZK 005	314
Cameroon 1941 Administrative	1941	2013 ZK 005	314
Cameroon 1941 Medical	1941	2013 ZK 005	314
Cameroon 1942 Administrative	1942	2013 ZK 005	314
Cameroon 1942 Medical	1942	2013 ZK 005	314
Cameroon 1943 Annex	1943	2013 ZK 005	111
Cameroon 1943 Annual Report with Mobile	1943	2013 ZK 005	111
Cameroon 1944 Annual Report with Mobile	1944	2013 ZK 005	111
Cameroon 1945 Annual Report with Mobile	1945	2013 ZK 005	111
Cameroon 1946 Annual Report with Mobile	1946	2013 ZK 005	111
Cameroon 1947 Annual Report with Mobile	1947	2013 ZK 005	110
Cameroon 1948 Annual Report with Mobile	1948	2013 ZK 005	110
Cameroon 1949 Annual Report with Mobile	1949	2013 ZK 005	110
Cameroon 1950 Annual Report with Mobile	1950	2013 ZK 005	110
Cameroon 1951 Annual Report with Mobile	1951	2013 ZK 005	110
Cameroon 1951 Supplementary	1951	2013 ZK 005	110
Cameroon 1952 Annual Report	1952	2013 ZK 005	108
Cameroon 1952 Climate	1952	2013 ZK 005	114
Cameroon 1953 Annual Report	1953	2013 ZK 005	108
Cameroon 1954 Annual Report	1954	2013 ZK 005	108
Cameroon 1954 Supplementary	1954	2013 ZK 005	108
Cameroon 1955 Annual Report	1955	2013 ZK 005	108
Cameroon 1955 Statistics	1955	2013 ZK 005	108
Cameroon 1955 Supplementary	1955	2013 ZK 005	112
Cameroon 1956 Annual Report	1956	2013 ZK 005	112
Cameroon 1956 Development Techniques	1956	2013 ZK 005	112
Cameroon 1956 Development Techniques v2	1956	2013 ZK 005	112
Cameroon 1956 Statistics	1956	2013 ZK 005	112
Cameroon 1957 Annual Report	1957	2013 ZK 005	112
Cameroon 1957 Development Techniques	1957	2013 ZK 005	112
Cameroon 1957 Statistics	1957	2013 ZK 005	112
Cameroon 1958 Annual Report	1958	2013 ZK 005	112
Cameroon 1968 Annual Report	1968	2013 ZK 005	112
Cameroon Unknown Map of Health Centers	Unknown	2013 ZK 005	

Table A4: Documents Collected From *Service Historique de la Defense*

Central African Republic Reports			
Document Name	**Year**	**Catalog**	**Call #**
CAR 1931 Annual Report	1931	2013 ZK 005	6
CAR 1932 Annual Report	1932	2013 ZK 005	6
CAR 1933 Annual Report	1933	2013 ZK 005	6
CAR 1934 Annual Report	1934	2013 ZK 005	6
CAR 1945 Annual Report	1945	2013 ZK 005	6
CAR 1946 Annual Report	1946	2013 ZK 005	6
CAR 1947 Annual Report	1947	2013 ZK 005	6
CAR 1948 Annual Report	1948	2013 ZK 005	6
CAR 1950 Annual Report	1950	2013 ZK 005	7
CAR 1951 Annual Report	1951	2013 ZK 005	7
CAR 1952 Annual Report	1952	2013 ZK 005	7
CAR 1953 Annual Report	1953	2013 ZK 005	7
CAR 1954 Annual Report	1954	2013 ZK 005	7
CAR 1954 Statistics	1954	2013 ZK 005	7
CAR 1955 Statistics	1955	2013 ZK 005	8
CAR 1956 Annual Report	1956	2013 ZK 005	8
CAR 1957 Annual Report	1957	2013 ZK 005	8
CAR 1958 Annual Report	1958	2013 ZK 005	8

Table A5: Documents Collected From *Service Historique de la Defense*

Chad Reports			
Document Name	**Year**	**Catalog**	**Call #**
Tchad 1931 Annual Report	1931	2013 ZK 005	125
Tchad 1932 Annual Report	1932	2013 ZK 005	125
Tchad 1933 Annual Report	1933	2013 ZK 005	125
Tchad 1934 Annual Report	1934	2013 ZK 005	125
Tchad 1947 Annual Report	1947	2013 ZK 005	125
Tchad 1948 Annual Report	1948	2013 ZK 005	125
Tchad 1949 Annual Report	1949	2013 ZK 005	125
Tchad 1950 Annual Report	1950	2013 ZK 005	125
Tchad 1951 Annual Report	1951	2013 ZK 005	125
Tchad 1952 Annual Report	1952	2013 ZK 005	125
Tchad 1953 Annual Report	1953	2013 ZK 005	125
Tchad 1959 Annual Report	1959	2013 ZK 005	123
Tchad 1960 Annual Report	1960	2013 ZK 005	123
Tchad 1961 Annual Report	1961	2013 ZK 005	123
Tchad 1964 Demographic Study	1964	2013 ZK 005	123

Table A6: Documents Collected From *Service Historique de la Defense*

Republic of Congo Reports			
Document Name	**Year**	**Catalog**	**Call #**
Moyen Congo 1931 Annual Report	1931	2013 ZK 005	117
Moyen Congo 1932 Annual Report	1932	2013 ZK 005	117
Moyen Congo 1933 Annual Report	1933	2013 ZK 005	117
Moyen Congo 1934 Annual Report	1934	2013 ZK 005	117
Moyen Congo 1945 Annual Report	1945	2013 ZK 005	116
Moyen Congo 1946 Administrative	1946	2013 ZK 005	116
Moyen Congo 1946 Medical	1946	2013 ZK 005	116
Moyen Congo 1947 Administrative	1947	2013 ZK 005	116
Moyen Congo 1947 Medical	1947	2013 ZK 005	116
Moyen Congo 1948 Annual Report	1948	2013 ZK 005	116
Moyen Congo 1949 Administrative	1949	2013 ZK 005	116
Moyen Congo 1949 Medical	1949	2013 ZK 005	116
Moyen Congo 1950 Medical	1950	2013 ZK 005	126
Moyen Congo 1951 Medical	1951	2013 ZK 005	126
Moyen Congo 1952 Medical	1952	2013 ZK 005	126
Moyen Congo 1953 Administrative	1953	2013 ZK 005	126
Moyen Congo 1953 Commentary	1953	2013 ZK 005	126
Moyen Congo 1954 Administrative	1954	2013 ZK 005	126
Moyen Congo 1954 Medical	1954	2013 ZK 005	126
Moyen Congo 1954 Medical Commentary	1954	2013 ZK 005	126
Moyen Congo 1955 Administrative Commentary	1955	2013 ZK 005	126
Moyen Congo 1955 Medical	1955	2013 ZK 005	126
Moyen Congo 1956 Annual Report Commentary	1956	2013 ZK 005	126
Moyen Congo 1956 Commentary	1956	2013 ZK 005	126
Moyen Congo 1956 Medical	1956	2013 ZK 005	126
Moyen Congo 1957 Annual Report	1957	2013 ZK 005	119
Moyen Congo 1957 Annual Report Commentary	1957	2013 ZK 005	119
Moyen Congo 1958 Annual Report	1958	2013 ZK 005	119

Gabon Reports			
Document Name	**Year**	**Catalog**	**Call #**
Gabon 1931 Annual Report	1931	2013 ZK 005	127
Gabon 1932 Annual Report Part 1	1932	2013 ZK 005	127
Gabon 1932 Annual Report Part 2	1932	2013 ZK 005	127
Gabon 1933 Annual Report Part 1	1933	2013 ZK 005	127
Gabon 1933 Annual Report Part 2	1933	2013 ZK 005	127
Gabon 1934 Medical	1934	2013 ZK 005	127
Gabon 1945 Annual Summary - Partial	1945	2013 ZK 005	128
Gabon 1945 Medical	1945	2013 ZK 005	128
Gabon 1948 Medical	1948	2013 ZK 005	128
Gabon 1949 Medical	1949	2013 ZK 005	128
Gabon 1950 Medical	1950	2013 ZK 005	128
Gabon 1951 Medical	1951	2013 ZK 005	128
Gabon 1952 Annual Summary - Partial	1952	2013 ZK 005	5
Gabon 1952 Statistics	1952	2013 ZK 005	5
Gabon 1953 Annual Summary - Partial	1953	2013 ZK 005	5
Gabon 1953 Statistics	1953	2013 ZK 005	5
Gabon 1954 Annual Summary	1954	2013 ZK 005	5
Gabon 1954 Statistics	1954	2013 ZK 005	5
Gabon 1955 Statistics	1955	2013 ZK 005	5
Gabon 1956 Administrative	1956	2013 ZK 005	5
Gabon 1956 Annual Summary	1956	2013 ZK 005	5
Gabon 1956 Statistics	1956	2013 ZK 005	5
Gabon 1957 Statistics	1957	2013 ZK 005	5

Table A7: Documents Collected from Other Sources

Cameroon Reports			
Document Name	**Year**	**Catalog**	**Call #**
Cameroon UN Report 1921	1921	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1922	1922	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1923	1923	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1925	1925	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1926	1926	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1927	1927	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1928	1928	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1929	1929	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1930	1930	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1931	1931	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1932	1932	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1933	1933	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1934	1934	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1935	1935	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1936	1936	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1937	1937	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1938	1938	UN	http://gallica.bnf.fr/ark:/12148/cb32848053c/date.r=.langEN (accessed on June 20th, 2013).
Cameroon UN Report 1952	1952	UN	1947, 1949-1957 available through Harvard and the center for research libs
Cameroon UN Report 1953	1953	UN	1947, 1949-1957 available through Harvard and the center for research libs
Cameroon UN Report 1954	1954	UN	1947, 1949-1957 available through Harvard and the center for research libs
Cameroon UN Report 1955	1955	UN	1947, 1949-1957 available through Harvard and the center for research libs
Cameroon UN Report 1956	1956	UN	1947, 1949-1957 available through Harvard and the center for research libs
Cameroon UN Report 1957	1957	UN	1947, 1949-1957 available through Harvard and the center for research libs

As described in Section 3.1 we collected these reports in January 2014 from the military archive. For the AEF countries we digitize data for 1927 to 1956. This data are at a sub-district level. For Cameroon, the data were reported at an ethnicity-district level for the years 1921 to 1938 and subsequently at the sub-district level from 1939 to 1950 with detailed maps of areas visited within Cameroon. When the data was not at the ethnicity-district level for Cameroon, in order to make the unit of observation consistent across reports, we aggregated the maps with sub-district information to the ethnicity-district level using area weights.[31] The report tables for Cameroon and AEF included detailed information on estimated number of people in an area, the number of people visited, the number of newly sick individuals, number of previously sick individuals, the number of lumbar punctures administered, and the number of previously sick individuals who had recovered. The reports also included narrative descriptions by the mobile team leaders of the activities undertaken by the health teams.

---

[31]Specifically, consider an area of a sub-district that was visited by the campaigns according the the report tables and map; for instance, Figure B2a is an example of a map documenting areas visited in 1941 in Cameroon. We first digitized the maps of visits and the sub-district campaign data. Second, we overlay the map of ethnicity-districts constructed using previous reports, Third, we aggregate the sub-district information to the ethnicity-district level using the area share of each ethnicity-district visited to aggregate the campaign table data. We considered an ethnic-district as visited for this aggregation if at least 5% of the ethnic groups area was visited.

Appendix B. Additional Tables and Figures

B.1. Historical Data Examples

Figure B1: Examples of Reports

(a) Example of Archival Data from Gabon (1954)

(b) Example of Archival Data from Cameroon (1934)

Figure B2: Examples of Maps

(a) Regions Visited in Cameroon (1941)
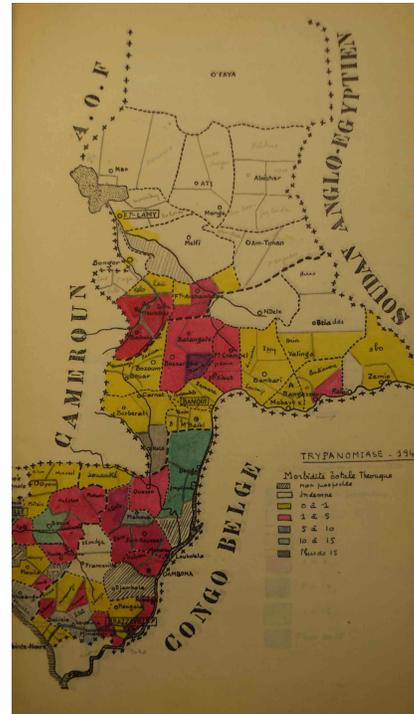
(b) Sleeping Sickness Prevalence AEF (1944)

B.2. Maps

Figure B3: Sleeping Sickness Rates at First Measurement
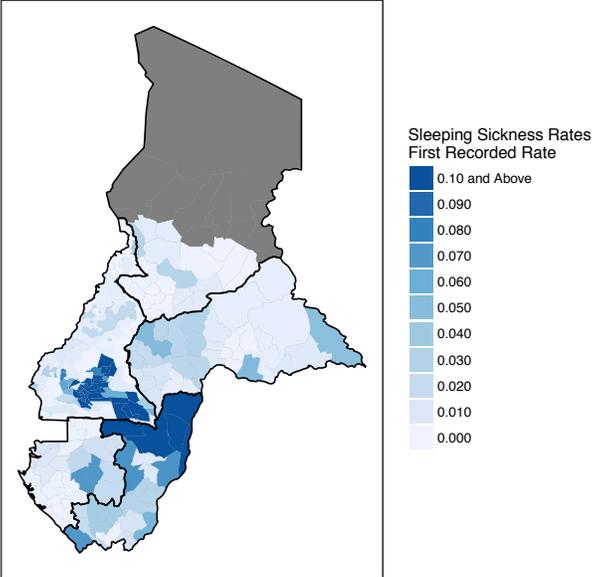
Figure B4: Average Reported Sleeping Sickness During First Three Visits
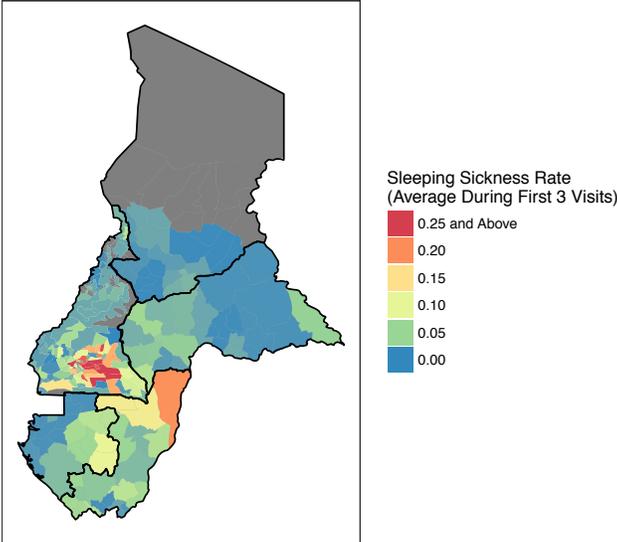
Figure B5: Sleeping Sickness Year of First Measurement
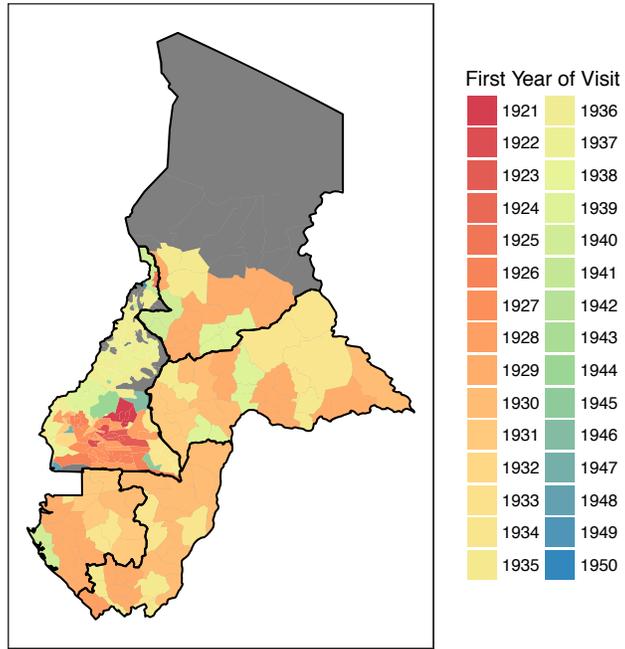
Figure B6: Sleeping Sickness - Maximum Prevalence Recorded
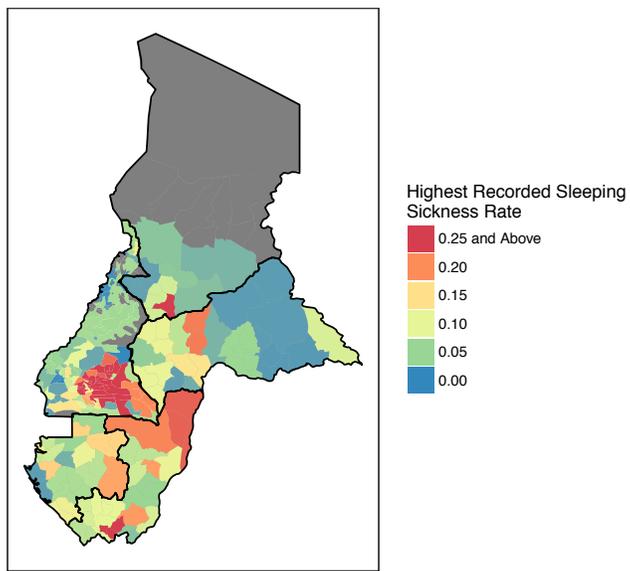
Figure B7: DHS Clusters or Observations Per District

DHS Clusters
• DHS Cluster

DHS w/o Coordinates:
Obs. per District
2000
4000
6000
8000
10000
12000
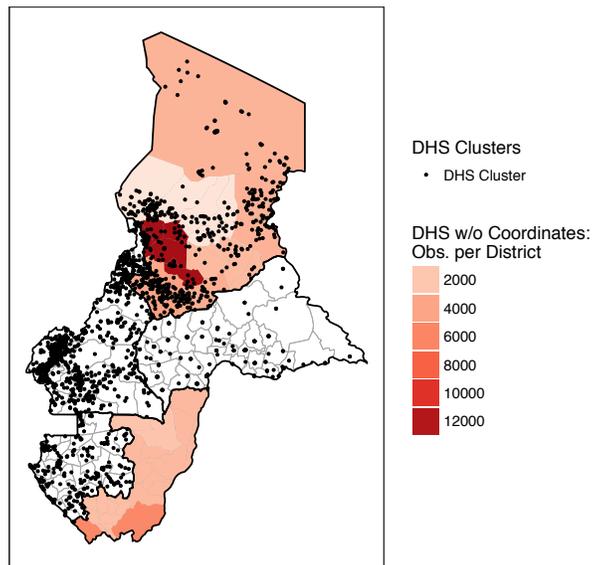
Figure B8: Maps of Variables for Instrument and Colonial Medical Campaign Visits in Cameroon and Former French Equatorial Africa
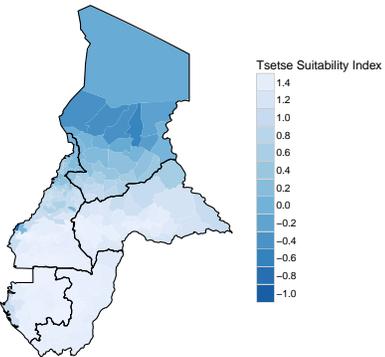
(a) Tse Tse fly Suitability
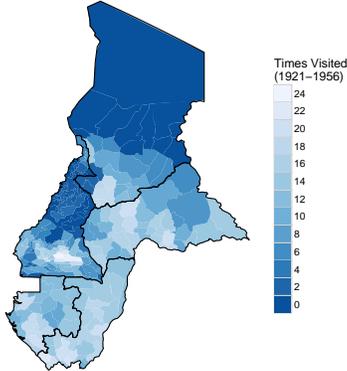
(b) Times Visited
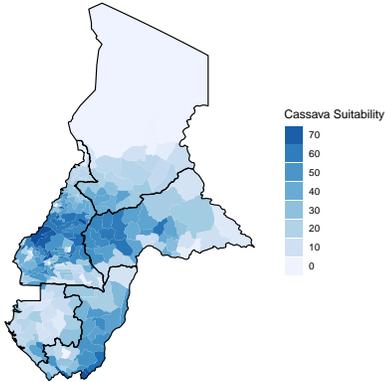
(c) Suitability for Cassava
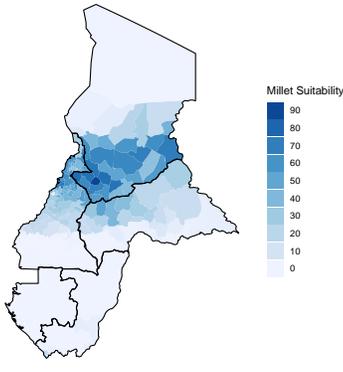
(d) Suitability for Millet
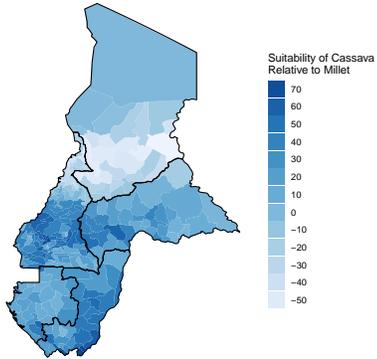
(e) Suitability for Cassava Relative to Millet
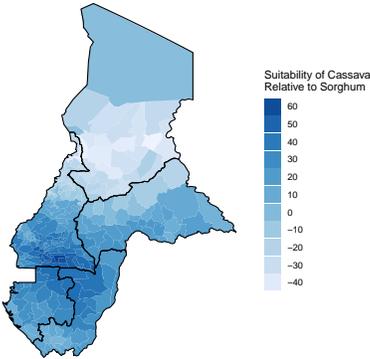
(f) Suitability for Cassava Relative to Sorghum

B.3. IV Measure: Relative Suitability Distribution Plots

Figure B9: Relative Suitability Measures (Cassava Suitability minus Millet Suitability)

(a) Relative Log Suitability

(b) Relative Suitability
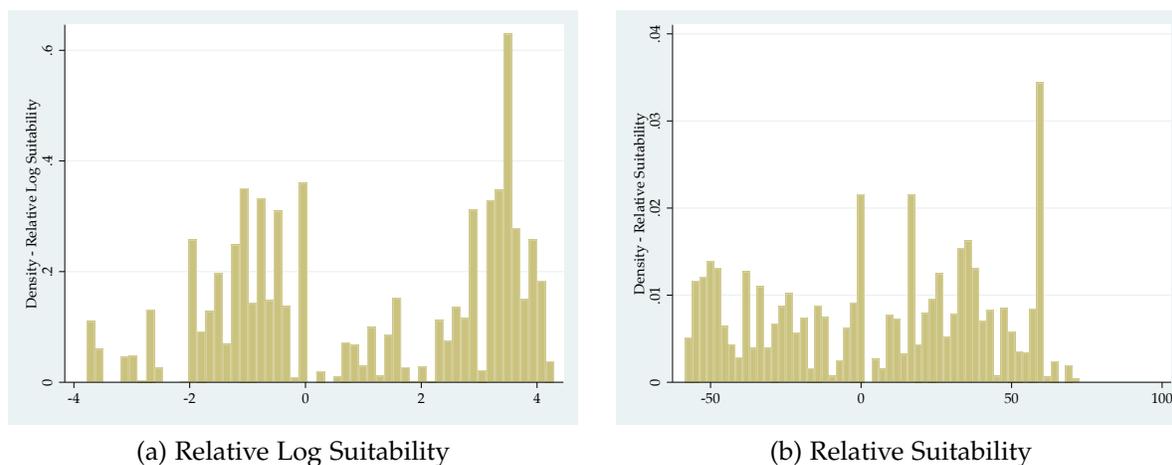
B.4. IV Falsification: Placebo Suitability Measures

This section presents the results from an additional falsification exercise for the instrumental variable used in the paper and the exercise comparing former French vs former British Cameroon. We argue that the suitability for cassava relative to millet is a predictor of sleeping sickness visits due to particular features of cassava that increase exposure to tsetse fly habitat. These features are that (1) cassava must be soaked to remove cyanide and areas for soaking often have tsetse fly and (2) cassava is more productive per hectare than millet, so less land must be cleared to produce a fixed amount of calories, leaving more tsetse fly harboring bush. We find that the relative suitability measure predicts blood test refusals and lower vaccination rates in former French Cameroon but not in former British Cameroon (where no campaigns occurred), providing evidence consistent with the exclusion restriction.

This argument also implies that we should not see similar effects if we use the suitability for crops other than cassva, since the features we highlight of cassava are specific to cassava. Thus, we can construct relative suitability measures all other crops for which the FAO GAEZ data has suitability measures relative to millet suitability, estimate the reduced form for these other crops, and examine how these placebo estimates compare to the cassava estimates.[32] The other relative suitability crop measures (relative to millet suitability) we examine are: banana, barley, buckwheat, cabbage, cacao, carrot, chickpea, citrus, coconut, coffee, cotton, cowpea, drypea, flax, greengram, groundnut, dryland rice, maize, oat, olive, onion, palm oil, phaseolus bean, pidgeon pea, rye, sorghum, soy bean, sugar beet, sugar cane, sun flower, sweet potato, tea, tobacco, tomato, wetland rice, wheat, white rice, white potato, and yams.

Figure B10 plots the t-statistics for this exercise. For all crops for which we can construct relative suitability measures, we estimate the difference in the reduced form coefficient between former French and former British Cameroon, and plot the distribution of the t-statistics for these

---

[32]In essence, this is similar to a randomization inference exercise.

placebo suitability measures (using different bandwidths near the border as in Table 3).[33] We then plot in red the t-statistic for our instrument, the relative suitability of cassava vs millet. We find that, across both bandwidths, the estimated effect for the relative suitability of cassava vs millet across former French vs former British Cameroon is much larger than the vast majority of all other possible relative suitability measures. This suggests that it is indeed something specific about (i) cassava, and (ii) french colonial medicine history that is associated with mistrust in medicine today and provides additional support for the exclusion restriction assumption of the instrument.

Figure B10: Randomization Inference Exercise - Reduced Form on Consent to Blood Test - Alternative Relative Suitability Measures (relative to Millet Suitability)

(a) Within 100 km of Border   (b) Within 50 km of Border
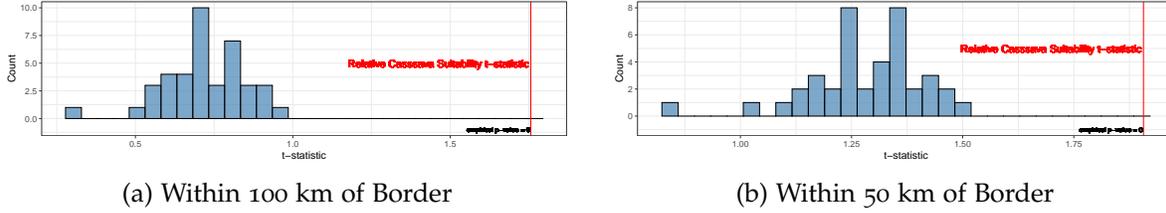
B.5. *Reduced Form for Instrument for Cameroon: Other Outcomes*

Figure B11 presents estimated differences in the reduced form of the instrument for alternative outcomes related to economic development across former French vs former British Cameroon (under alternative bandwidths from the former border). The logic is that the instrument should not be affecting other economic outcomes outside of the health outcomes examined in Table 3. Specifically, we estimate the following reduced form equation:

$$y_{irt} = \gamma RelSuit_r + \alpha_y \text{FC}_{irt} + \beta FC_{irt} \times RelSuit_r + X'_{irt}B + FC_{irt}X'_{irt}\Gamma + \delta_t + FC_{irt}\delta_t + \varepsilon_{irt} \qquad \text{(A1)}$$

where i indexes individuals, r DHS clusters, t survey year, y_{irt} an outcome of interest related to economic development, $RelSuit_r$ the relative suitability measure for cassava vs millet at the DHS cluster, FC_{irt} is an indicator equal to one if the cluster is on the former French Cameroon side of the border, X_{irt} is the vector of controls (i.e. baseline controls, climate and disease controls, colonial controls, and contemporaneous controls) defined as before, and δ_t are survey year fixed effects. Our coefficient of interest is β, which examines whether there is a differential effect across former British and French Cameroon of the relative suitability measure on economic outcomes.

We use DHS data and include the same set of geographic and historical controls as in Table 3. The results in the figure suggest that the two regions are comparable in terms of economic development in terms of the reduced form effects of the instrument on economic outcomes in these bandwidths. (The one exception is urban-rural status; however, we control for this variable

---

[33]We plot the t-statistic for the β from the following specification equation: $y_{irt} = \gamma CSuit_r + \alpha_y \text{FC}_{irt} + \beta FC_{irt} \times CSuit_r + X'_{irt}B + FC_{irt}X'_{irt}\Gamma + \delta_t + FC_{irt}\delta_t + \varepsilon_{irt}$, where i indexes individuals, r dhs clusters, t survey year, y_{irt} an indicator for whether a blood test was refused, $CSuit_r$ the suitability for a given crop at the DHS cluster, FC_{irt} an indicator equal to one if the cluster is on the former French Cameroon side of the border, X_{irt} is the vector of controls (i.e. baseline controls, climate and disease controls, colonial controls, and contemporaneous controls) defined as before, and δ_t are survey year fixed effects. Thus, β estimates if there is a differential effect across former British and French Cameroon of a given crop's relative suitability measure on blood test refusals.

throughout in all our regressions.) This suggests that differences in economic development at the border are unlikely to explain the results presented in Table 3.[34]

Figure B11: Estimated Differences in Reduced Form for Instrument between Former British and French Cameroon - Alternative Outcomes

(a) Within 100 km of Border

(b) Within 50 km of Border
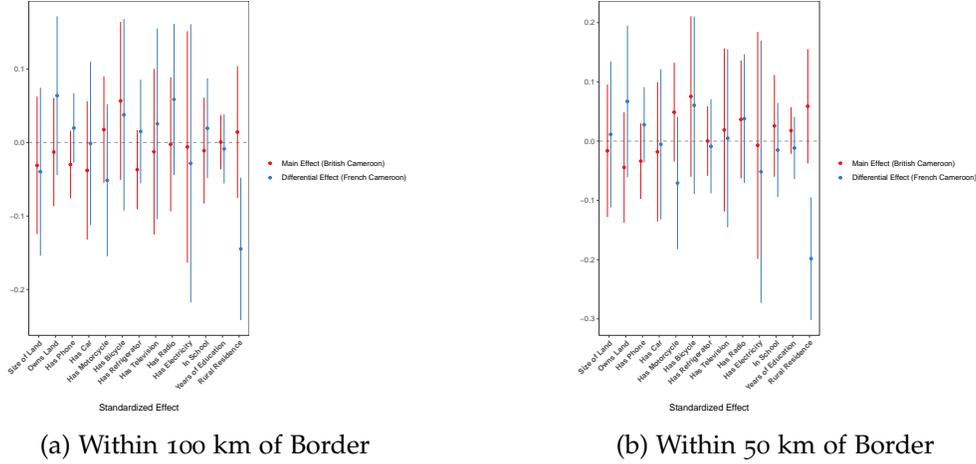
B.6. DHS Convergence

Given that the exposure to the medical campaigns happened between the 1920s and 1950s and our sample population was, for the most part, not alive during this period, it is possible that the effect on mistrust may diminish overtime. We can examine whether this is the case. To do this, we compare cohorts with varying levels of historical exposure to the medical campaigns born within five years of each other by estimating a regression that includes fixed effects for each 5-year cohort along with the interactions between the *Share of Years Visited$_r$* (SYV_r) and cohort fixed effects. Formally, we estimate the following specification:

$$y_{irct} = \gamma SYV_r + \alpha_y C_{irct} + \gamma_y C_{irct} \times SYV_r + \mathbf{X}'_{irct}\mathbf{B} + \mathbf{X}'_r\mathbf{\Gamma} + \delta_{ct} + \varepsilon_{irct} \qquad \text{(A2)}$$

where C_{irct} are 5-year cohort fixed effects and the other variables are defined as in equation (1).

Figure B12 plots the estimated cohort coefficients for refusal to consent to blood test, first with survey fixed effects and second with the full set of controls. We see no evidence of convergence across cohorts: the estimated coefficients for each cohort are similar, particularly once controls are added, and the effects do not seem to be driven by older cohorts.

---

[34]Additionally, the main effect of our instrument – γ in our empirical equation above – on these outcomes is consistently very small and rarely statistically significant.

Figure B12: Refusals to Consent to Blood Test by DHS Cohort
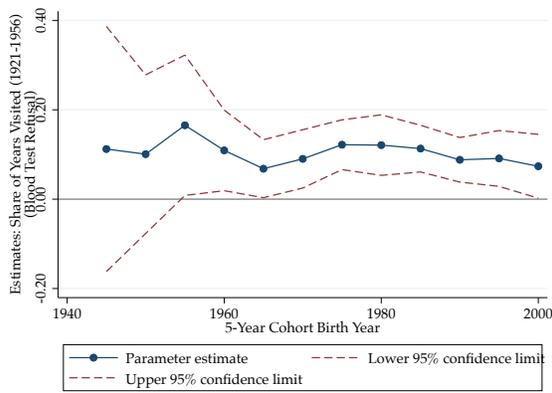
(a) No Controls Except Survey Fixed Effects
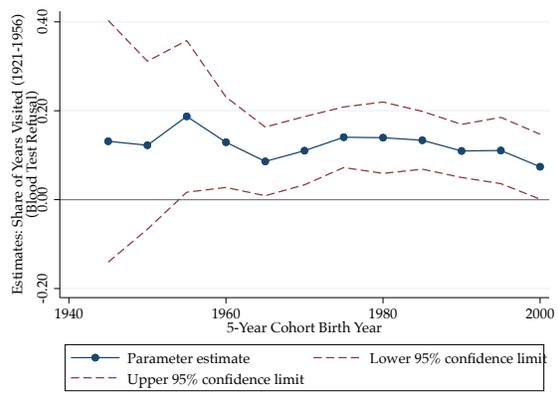
(b) Full Controls

B.7. Afrobarometer Figures

We plot the coefficients for each individual trust question available in the Afrobarometer. The AES coefficients are presented in Table 6.

Figure B13: Afrobarometer - Round 5 Trust Outcomes - Coefficient Plot
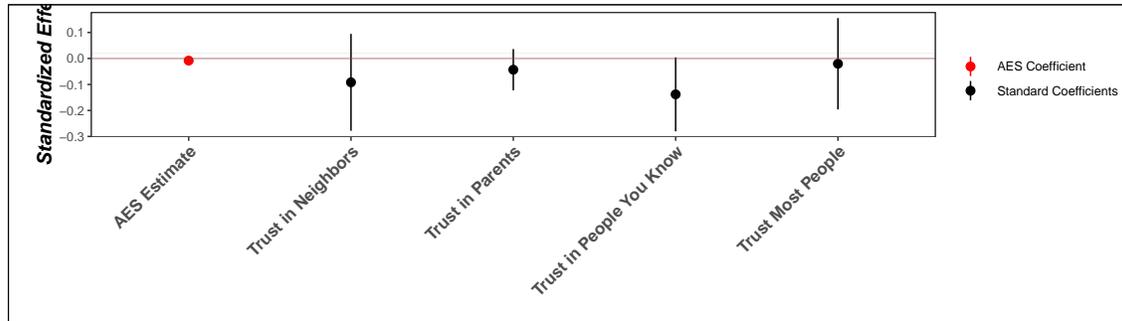
Figure B14: Afrobarometer - Round 6 Trust Outcomes - Coefficient Plot
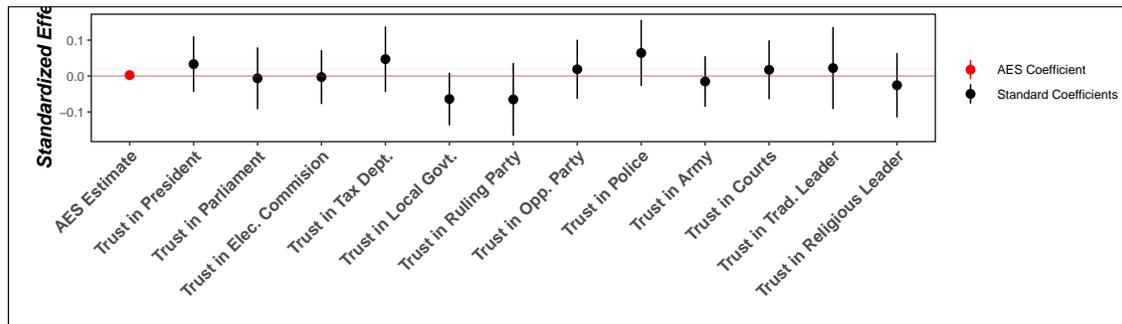
B.8. OLS Robustness

In Table B1 we present the vaccination results as we add each set of controls. In Table B2 we present the blood test refusal results as we add each set of controls. In Tables B3 and B4 we present the refusal results by hemoglobin test and HIV test separately. In Tables B5 and B6 we present the vaccination and refusal results including the pre-colonial controls. Note, not all ethnic groups have data for pre-colonial controls, so we interpolate values using the nearest neighbor within each country and include an indicator for whether or not the ethnic group had pre-colonial data in Murdock. Table B7 present the vaccination and blood test refusal results including Northern Chad, which was never visited by the campaigns and is not suitable for the tsetse fly. Table B8 presents OLS results with controls for all colonial medical report measures of sleeping sickness prevalence from 1921 and 1956. B9 uses data from the World Health Organization on sub-national immunization for Cameroon, Chad, Central Africa Republic, Gabon and Republic of Congo.

Table B1: OLS Estimates of Campaigns and Vaccination Rates - Robustness

	Dep. Var.: *DHS Vaccination Index*					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Years Visited (1921-1956)	0.0752	-0.0831	-0.0889	-0.0928	-0.116	-0.116
	(0.0516)	(0.0478)	(0.0478)	(0.0473)	(0.0442)	(0.0435)
Geography and Climate Controls	N	Y	Y	Y	Y	Y
Disease Suitability Controls	N	N	Y	Y	Y	Y
Colonial Controls	N	N	N	Y	Y	Y
Contemporary Controls	N	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	N	Y
Observations	50,773	50,773	50,773	50,773	50,668	50,668
Clusters	207	207	207	207	207	207
Mean Dep. Var.	0.532	0.532	0.532	0.532	0.532	0.532

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), CAR (1994) and Chad (1996, 2004, 2014). Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon, CAR, and Chad (2014), and at the district level for Congo and Chad (1996, 2004). *Vaccination Index* is the share of vaccines completed out of nine possible vaccines for children in the DHS. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. Regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

Table B2: OLS Estimates of Campaigns and Blood Test Refusals - Robustness

	Dep. Var.: *Blood Test Refused*					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Years Visited (1921-1956)	0.0752	0.136	0.129	0.128	0.108	0.115
	(0.0270)	(0.0352)	(0.0344)	(0.0351)	(0.0284)	(0.0293)
Geography and Climate Controls	N	Y	Y	Y	Y	Y
Disease Suitability Controls	N	N	Y	Y	Y	Y
Colonial Controls	N	N	N	Y	Y	Y
Contemporary Controls	N	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	N	Y
Observations	71,052	71,052	71,052	71,052	71,001	71,001
Clusters	160	160	160	160	160	160
Mean Dep. Var.	0.048	0.048	0.048	0.048	0.048	0.048

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), and Chad (2014). Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon and Chad, and at the district level for Congo. *Blood Test Refused* is an indicator variable for refusing to consent to a blood test (for either HIV or anemia) in the DHS data. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. Regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

Table B3: OLS Estimates of Refusal of Hemoglobin Test

	Dep. Var.: *Hemoglobin Blood Test Refused*					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Years Visited (1921-1956)	0.0765	0.133	0.115	0.114	0.0981	0.119
	(0.0241)	(0.0295)	(0.0323)	(0.0341)	(0.0282)	(0.0237)
Geography and Climate Controls	N	Y	Y	Y	Y	Y
Disease Suitability Controls	N	N	Y	Y	Y	Y
Colonial Controls	N	N	N	Y	Y	Y
Contemporary Controls	N	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	N	Y
Observations	32,652	32,652	32,652	32,652	32,652	32,652
Clusters	138	138	138	138	138	138
Mean Dep. Var.	0.048	0.048	0.048	0.048	0.048	0.048

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), and Congo (2011). Standard errors are clustered at the ethnic group level for Cameroon, at the colonial Sub-District level for Gabon, and at the district level for Congo. *Hemoglobin Blood Test Refused* is an indicator variable for refusing to consent to a blood test for hemoglobin levels. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. Regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude of each region. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

Table B4: OLS Estimates of Refusal of HIV Test

	Dep. Var.: *HIV Blood Test Refused*					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Years Visited (1921-1956)	0.0756	0.123	0.121	0.122	0.0995	0.100
	(0.0278)	(0.0366)	(0.0353)	(0.0354)	(0.0284)	(0.0299)
Geography and Climate Controls	N	Y	Y	Y	Y	Y
Disease Suitability Controls	N	N	Y	Y	Y	Y
Colonial Controls	N	N	N	Y	Y	Y
Contemporary Controls	N	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	N	Y
Observations	53,266	53,266	53,266	53,266	53,215	53,215
Clusters	160	160	160	160	160	160
Mean Dep. Var.	0.044	0.044	0.044	0.044	0.044	0.044

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2009) and Chad (2014). Standard errors are clustered at the ethnic group level for Cameroon, at the colonial Sub-District level for Gabon and Chad, and at the district level for Congo. *HIV Blood Test Refused* is an indicator variable for refusing to consent to a blood test for HIV. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. Regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude of each region. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

Table B5: OLS Estimates of Campaigns and Vaccination Rates - Pre-Colonial Controls

	Dep. Var.: *Vaccination Index*				
	(1)	(2)	(3)	(4)	(5)
Share of Years Visited (1921-1956)	-0.0624	-0.0695	-0.0753	-0.0933	-0.116
	(0.0564)	(0.0577)	(0.0579)	(0.0544)	(0.0435)
Pre-Colonial Controls	Y	Y	Y	Y	Y
Geography and Climate Controls	Y	Y	Y	Y	Y
Disease Suitability Controls	N	Y	Y	Y	Y
Colonial Controls	N	N	Y	Y	Y
Contemporary Controls	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	Y
Observations	50,773	50,773	50,773	50,668	50,668
Clusters	207	207	207	207	207
Mean Dep. Var.	0.532	0.532	0.532	0.532	0.532

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), CAR (1994) and Chad (1996, 2004, 2014). Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon, CAR, and Chad (2014), and at the district level for Congo and Chad (1996, 2004). *Vaccination Index* is the share of vaccines completed out of nine possible vaccines for children in the DHS. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. Regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Pre-Colonial Controls* include level of centralization, use of plow, whether indigenous slavery was practiced, whether agriculture was practiced for each main ethnic group in a region, and indicator variable equal to 1 if the observation is not missing. We assign a value of negative one to missing values for each *Pre-Colonial Control*. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

Table B6: OLS Estimates of Campaigns and Blood Test Refusals - Pre-Colonial Controls

	Dep. Var.: *Blood Test Refused*				
	(1)	(2)	(3)	(4)	(5)
Share of Years Visited (1921-1956)	0.110	0.106	0.109	0.0942	0.0716
	(0.0258)	(0.0273)	(0.0286)	(0.0246)	(0.0238)
Pre-Colonial Controls	Y	Y	Y	Y	Y
Geography and Climate Controls	Y	Y	Y	Y	Y
Disease Suitability Controls	N	Y	Y	Y	Y
Colonial Controls	N	N	Y	Y	Y
Contemporary Controls	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	Y
Observations	71,052	71,052	71,052	71,001	71,001
Clusters	160	160	160	160	160
Mean Dep. Var.	0.048	0.048	0.048	0.048	0.048

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), and Chad (2014). Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon and Chad, and at the district level for Congo. *Blood Test Refused* is an indicator variable for refusing to consent to a blood test (for either HIV or anemia) in the DHS data. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. Regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Pre-Colonial Controls* include level of centralization, use of plow, whether indigenous slavery was practiced, whether agriculture was practiced for each main ethnic group in a region, and indicator variable equal to 1 if the observation is not missing. We assign a value of negative one to missing values for each *Pre-Colonial Control*. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

Table B7: OLS Estimates of Campaigns and Vaccination Rates and Blood Test Refusals - Including Northern Chad

	Dep. Var.: *Vaccination Index*			Dep. Var.: *Blood Test Refused*		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Years Visited (1921-1956)	-0.0728	-0.0997	-0.0832	0.139	0.120	0.126
	(0.0512)	(0.0474)	(0.0514)	(0.0347)	(0.0290)	(0.0325)
Geography and Climate Controls	Y	Y	Y	Y	Y	Y
Disease Suitability Controls	Y	Y	Y	Y	Y	Y
Colonial Controls	Y	Y	Y	Y	Y	Y
Contemporary Controls	N	Y	Y	N	Y	Y
Lasso-Selected Controls	N	N	Y	N	N	Y
Observations	60,494	60,353	60,353	75,964	75,887	75,887
Clusters	220	220	220	171	171	171
Mean Dep. Var.	0.491	0.491	0.491	0.051	0.051	0.051

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), CAR (1994) and Chad (1996, 2004, 2014). Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon, CAR, and Chad (2014), and at the district level for Congo and Chad (1996, 2004). *Vaccination Index* is the share of vaccines completed out of nine possible vaccines for children in the DHS. *Blood Test Refused* is an indicator variable for refusing to consent to a blood test (for either HIV or anemia) in the DHS data. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. All regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

Table B8: OLS Estimates of Campaigns and Vaccination Rates - Including All Colonial Sleeping Sickness Prevalence Measures

	Dep. Var.: *Vaccination Index*			Dep. Var.: *Blood Test Refused*		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Years Visited (1921-1956)	-0.118	-0.138	-0.116	0.109	0.0582	0.0721
	(0.0436)	(0.0519)	(0.0435)	(0.0273)	(0.0297)	(0.0235)
Geography and Climate Controls	Y	Y	Y	Y	Y	Y
Disease Suitability Controls	Y	Y	Y	Y	Y	Y
Colonial Controls	Y	Y	Y	Y	Y	Y
Contemporary Controls	Y	Y	Y	N	Y	Y
Lasso-Selected Controls	N	N	Y	N	N	Y
All Prevalence Measures	N	Y	Y	N	Y	Y
Observations	50,750	50,750	50,750	71,001	71,001	71,001
Clusters	207	207	207	160	160	160
Mean Dep. Var.	0.533	0.533	0.533	0.048	0.048	0.048

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), CAR (1994) and Chad (1996, 2004, 2014). Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon, CAR, and Chad (2014), and at the district level for Congo and Chad (1996, 2004). *Vaccination Index* is the share of vaccines completed out of nine possible vaccines for children in the DHS. *Blood Test Refused* is an indicator variable for refusing to consent to a blood test (for either HIV or anemia) in the DHS data. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. *Relative Suitability: Cassava vs. Millet* is the logarithm of the differences in suitabilities for cassava and millet. All regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability, and initial sleeping sickness prevalence. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects. *All Prevalence Measures* includes all yearly measures of sleeping sickness prevalence available in the colonial reports from 1921 to 1956.

Table B9: OLS Estimates of Campaigns and Vaccination Rates - WHO Data

			Dep. Var.: *WHO Vaccination Index*			
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Years Visited (1921-1956)	-0.00884	-0.00804	-0.00815	-0.00815	-0.00739	-0.00820
	(0.00341)	(0.00351)	(0.00354)	(0.00354)	(0.00344)	(0.00324)
Geography and Climate Controls	N	Y	Y	Y	Y	Y
Disease Suitability Controls	N	N	Y	Y	Y	Y
Colonial Controls	N	N	N	Y	Y	Y
Contemporary Controls	N	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	N	Y
Observations	701	701	701	701	701	701
Clusters	146	146	146	146	146	146
Mean Dep. Var.	0.779	0.779	0.779	0.779	0.779	0.779

Notes: Data are from the WHO sub-national immunization coverage data. The unit of observation is the sub-national unit. The data is for Cameroon, Chad, CAR, Gabon and Republic of Congo. Standard errors are clustered at the sub-district level. *WHO Vaccination Index* is the average of the vaccination rates for each sub-national unit with reported coverage. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. Regressions include country and year (either 2016 or 2017) fixed effects. Estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Pre-Colonial Controls* include level of centralization, use of plow, whether indigenous slavery was practiced, and whether agriculture was practiced for each main ethnic group in a region. *Contemporary Controls* include reported population for each sub-district. See Section A.7 for more information.

B.9. IV Robustness

Table B10 examines the geographical and historical correlates of the number of colonial medicine campaign visits. In Table B11 we show that the instrument predicts sleeping sickness for across sub-Saharan Africa. As additional robustness tests for the IV estimates, in Table B12 we present the IV results including Northern Chad. Note that the results are imprecise for this sample as the instrument only has predictive power for regions suitable for agriculture. Tables B13 and B14 present the results with pre-colonial controls. The results are robust for blood test refusals but imprecise for vaccinations due to a number of ethnic groups missing pre-colonial data. Figure B9 shows the distribution of the relative log suitabilities as well as the relative suitability measure. Table B15 presents the IV results with controls for all colonial medical report measures of sleeping sickness prevalence from 1921 and 1956.

Table B10: Correlates of Sleeping Sickness Campaign Visits

		Dep. Var.: *Share of Years Visited*				
	(1)	(2)	(3)	(4)	(5)	(6)
Area	0.008					-0.729
	(1.073)					(0.939)
Latitude	-0.027					-0.040
	(0.004)					(0.008)
Longitude	0.013					0.010
	(0.006)					(0.006)
Elevation	-0.131					0.099
	(0.055)					(0.189)
Temperature		0.008				0.043
		(0.014)				(0.038)
Rainfall		0.222				-0.054
		(0.437)				(0.445)
Land Suitability		-0.434				-0.161
		(0.132)				(0.101)
Malaria Ecology Index			-0.001			0.006
			(0.005)			(0.005)
Tsetse Fly Suitability Index			0.184			-0.025
			(0.042)			(0.058)
Initial Prevalence			0.289			0.324
			(0.157)			(0.124)
Atlantic Slave Trade Exports				0.054		-0.006
				(0.008)		(0.013)
Number of Missions				0.067		-0.015
				(0.032)		(0.021)
Rural					-0.056	-0.062
					(0.033)	(0.016)
Observations	2,746	2,746	2,746	2,746	2,746	2,746
R-squared	0.437	0.166	0.243	0.145	0.018	0.506
Clusters	207	207	207	207	207	207
Mean Dep. Var.	0.405	0.405	0.405	0.405	0.405	0.405

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), CAR (1994) and Chad (1996, 2004, 2014). The unit of observation is at the DHS cluster level. Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon, CAR, and Chad (2014), and at the district level for Congo and Chad (1996, 2004). *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. Regressions examine the relationship between the *Share of Years Visited* and: mean surface area, centroid latitude, centroid longitude, mean altitude, mean temperature, mean precipitation, mean land suitability, mean malaria ecology index, tsetse fly suitability, initial sleeping sickness prevalence, total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade, number of missions for each main ethnic group in a region, and urban-rural status.

Table B11: Relationship between Instrument and Sleeping Sickness Prevalence Across Africa

	Dep. Var.: *Share of Group with Sleeping Sickness Present During Colonial Era*			
	(1)	(2)	(3)	(4)
Relative Suitability: Cassava vs. Millet	0.0422	0.0261	0.0240	0.0167
	(0.0126)	(0.0131)	(0.0135)	(0.0106)
Geography and Climate Controls	Y	Y	Y	Y
Disease Suitability Controls	N	Y	Y	Y
Colonial Controls	N	N	Y	Y
Lasso-Selected Controls	N	N	N	Y
Observations	346	346	346	346
Mean Dep. Var.	0.552	0.552	0.552	0.552

Notes: Data are from Alsan (2015) at the Murdock ethnic group level. We present robust standard errors in parentheses. *Share of Group with Sleeping Sickness Incidence* is constructed from Ford and Katondo (1977), who provide a binary measure of presence of sleeping sickness based on colonial surveillance records on the tsetse fly and observed sleeping sickness, which Alsan (2015) aggregates to the share of an ethnic group that had sleeping sickness. *Relative Suitability: Cassava vs. Millet* is the difference in log suitabilities for cassava and millet. All regressions include country fixed effects. Estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index and tsetse fly suitability index. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region.

Table B12: First and Second Stage IV Estimates - Including Northern Chad

			Panel A: First-Stage Estimates			
			Share of Years Visited			
	(1)	(2)	(3)	(4)	(5)	(6)
Relative Suitability: Cassava vs. Millet	0.0117	0.0118	0.0129	0.0377	0.0369	0.0296
	(0.0108)	(0.0105)	(0.0108)	(0.0205)	(0.0197)	(0.0138)
Geography and Climate Controls	Y	Y	Y	Y	Y	Y
Disease Suitability Controls	Y	Y	Y	Y	Y	Y
Colonial Controls	Y	Y	Y	Y	Y	Y
Contemporary Controls	N	Y	Y	N	Y	Y
Lasso-Selected Controls	N	N	Y	N	N	Y
Observations	60,494	60,353	60,353	75,964	75,887	75,887
Clusters	220	220	220	171	171	171
Mean Dep. Var.	0.312	0.312	0.0263	0.389	0.389	0.389
F-Stat of Excluded Instrument	1.188	1.252	–	3.379	3.524	–

			Panel B: Second-Stage 2SLS Estimates			
	Dep. Var.: Vaccination Index			Dep. Var.: Blood Test Refused		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Years Visited (1921-1956)	0.532	0.0361	-0.935	0.0870	0.0462	0.405
	(0.775)	(0.559)	(0.786)	(0.167)	(0.157)	(0.159)
Geography and Climate Controls	Y	Y	Y	Y	Y	Y
Disease Suitability Controls	Y	Y	Y	Y	Y	Y
Colonial Controls	Y	Y	Y	Y	Y	Y
Contemporary Controls	N	Y	Y	N	Y	Y
Lasso-Selected Controls	N	N	Y	N	N	Y
Observations	60,494	60,353	60,353	75,964	75,887	75,887
Clusters	220	220	220	171	171	171
Mean Dep. Var.	0.491	0.491	0.491	0.051	0.051	0.051

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), CAR (1994) and Chad (1996, 2004, 2014). Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon, CAR, and Chad (2014), and at the district level for Congo and Chad (1996, 2004). *Vaccination Index* is the share of vaccines completed out of nine possible vaccines for children in the DHS. *Blood Test Refused* is an indicator variable for refusing to consent to a blood test (for either HIV or anemia) in the DHS data. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. *Relative Suitability: Cassava vs. Millet* is the logarithm of the difference in log suitabilities for cassava and millet. All regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude, mean altitude, and whether the region of northern Chad is suitable for cassava. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability, and initial sleeping sickness prevalence. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

Table B13: IV Estimates of Campaigns and Vaccination Rates - Pre-Colonial Controls

	Second-Stage 2SLS Estimates				
	Dep. Var.: *Vaccination Index*				
	(1)	(2)	(3)	(4)	(5)
Share of Years Visited (1921-1956)	0.243	0.0950	0.0986	-0.0600	-0.297
	(0.178)	(0.180)	(0.192)	(0.173)	(0.167)
Pre-Colonial Controls	Y	Y	Y	Y	Y
Geography and Climate Controls	Y	Y	Y	Y	Y
Disease Suitability Controls	N	Y	Y	Y	Y
Colonial Controls	N	N	Y	Y	Y
Contemporary Controls	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	Y
Observations	50,773	50,773	50,773	50,668	50,668
Clusters	207	207	207	207	207
Mean Dep. Var.	0.532	0.532	0.532	0.532	0.532
F Stat. Excluded Instrument	13.30	10.26	10.06	9.933	–

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), CAR (1994) and Chad (1996, 2004, 2014). Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon, CAR, and Chad (2014), and at the district level for Congo and Chad (1996, 2004). *Vaccination Index* is the share of vaccines completed out of nine possible vaccines for children in the DHS. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. *Relative Suitability: Cassava vs. Millet* is the instrument and is defined as the difference in log suitabilities for cassava and millet. Regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Pre-Colonial Controls* include level of centralization, use of plow, whether indigenous slavery was practiced, whether agriculture was practiced for each main ethnic group in a region, and indicator variable equal to 1 if the observation is not missing. We assign a value of negative one to missing values for each *Pre-Colonial Control*. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

Table B14: IV Estimates of Campaigns and Blood Test Refusals - Pre-Colonial Controls

	Second-Stage 2SLS Estimates				
	Dep. Var.: *Blood Test Refused*				
	(1)	(2)	(3)	(4)	(5)
Share of Years Visited (1921-1956)	0.175	0.119	0.132	0.109	0.124
	(0.0662)	(0.0792)	(0.0800)	(0.0716)	(0.0548)
Pre-Colonial Controls	Y	Y	Y	Y	Y
Geography and Climate Controls	Y	Y	Y	Y	Y
Disease Suitability Controls	N	Y	Y	Y	Y
Colonial Controls	N	N	Y	Y	Y
Contemporary Controls	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	Y
Observations	71,052	71,052	71,052	71,001	71,001
Clusters	160	160	160	160	160
Mean Dep. Var.	0.048	0.048	0.048	0.048	0.048
F Stat. Excluded Instrument	18.20	13.97	16.39	16.55	–

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), and Chad (2014). Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon and Chad, and at the district level for Congo. *Blood Test Refused* is an indicator variable for refusing to consent to a blood test (for either HIV or anemia) in the DHS data. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. *Relative Suitability: Cassava vs. Millet* is the instrument and is defined as the difference in log suitabilities for cassava and millet. Regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Pre-Colonial Controls* include level of centralization, use of plow, whether indigenous slavery was practiced, whether agriculture was practiced for each main ethnic group in a region, and indicator variable equal to 1 if the observation is not missing. We assign a value of negative one to missing values for each *Pre-Colonial Control*. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

Table B15: IV Estimates of Campaigns and Vaccination Rates - Including All Colonial Sleeping Sickness Prevalence Measures

	Dep. Var.: *Vaccination Index*			Dep. Var.: *Blood Test Refused*		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Years Visited (1921-1956)	-0.254	-0.213	-0.290	0.162	0.333	0.231
	(0.141)	(0.316)	(0.167)	(0.0572)	(0.144)	(0.0775)
Geography and Climate Controls	Y	Y	Y	Y	Y	Y
Disease Suitability Controls	Y	Y	Y	Y	Y	Y
Colonial Controls	Y	Y	Y	Y	Y	Y
Contemporary Controls	Y	Y	Y	N	Y	Y
Lasso-Selected Controls	N	N	Y	N	N	Y
All Prevalence Measures	N	Y	Y	N	Y	Y
Observations	50,750	50,750	50,750	71,001	71,001	71,001
Clusters	207	207	207	160	160	160
Mean Dep. Var.	0.533	0.533	0.533	0.048	0.048	0.048
F Stat Excluded Instrument	11.97	4.505	–	18.598	6.552	–

Notes: Data are from the DHS for Cameroon (2004 and 2011), Gabon (2012), Congo (2011), CAR (1994) and Chad (1996, 2004, 2014). Standard errors are clustered at the ethnic group-district level for Cameroon, at the colonial sub-district level for Gabon, CAR, and Chad (2014), and at the district level for Congo and Chad (1996, 2004). *Vaccination Index* is the share of vaccines completed out of nine possible vaccines for children in the DHS. *Blood Test Refused* is an indicator variable for refusing to consent to a blood test (for either HIV or anemia) in the DHS data. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. *Relative Suitability: Cassava vs. Millet* is the difference in log suitabilities for cassava and millet. All regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects; however, estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability, and initial sleeping sickness prevalence. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects. *All Prevalence Measures* includes all yearly measures of sleeping sickness prevalence available in the colonial reports from 1921 to 1956.

B.10. Reduced Form Estimates for Cameroon: Alternative Bandwidths

Table B16: Falsification Exercise - Reduced Form Estimates in Former British and French Cameroon - Alternative Bandwidths

	Falsification Test for Relative Suitability of Cassava vs. Millet							
	Panel A: 75 km Bandwidth							
	Dep Var.: *Vaccination Index*				Dep Var.: *Blood Test Refused*			
	Former British Cameroon		Former French Cameroon		Former British Cameroon		Former French Cameroon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative Suitability: Cassava vs. Millet	0.00548 (0.00499)	0.00441 (0.00487)	-0.0128 (0.00586)	-0.0122 (0.00528)	-0.00273 (0.00262)	-0.00279 (0.00262)	0.00478 (0.00298)	0.00466 (0.00299)
Geography and Climate Controls	Y	Y	Y	Y	Y	Y	Y	Y
Disease Suitability Controls	Y	Y	Y	Y	Y	Y	Y	Y
Colonial Controls	Y	Y	Y	Y	Y	Y	Y	Y
Contemporary Controls	N	Y	N	Y	N	Y	N	Y
p-value: Former British vs. French Coefficients	–	–	0.034	0.047	–	–	0.053	0.064
Bandwidth	75 km	75 km	75 km	75 km	75 km	75 km	75 km	75 km
Observations	2,396	2,396	3,998	3,998	5,693	5,693	8,683	8,683
Clusters	164	164	268	268	164	164	268	268
Mean Dep. Var.	0.771	0.771	0.732	0.732	0.0300	0.0300	0.0610	0.0610
	Panel B: 100 km Bandwidth							
	Dep Var.: *Vaccination Index*				Dep Var.: *Blood Test Refused*			
	Former British Cameroon		Former French Cameroon		Former British Cameroon		Former French Cameroon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative Suitability: Cassava vs. Millet	0.00825 (0.00557)	0.00758 (0.00559)	-0.0123 (0.00584)	-0.0115 (0.00526)	-0.00276 (0.00255)	-0.00282 (0.00255)	0.00399 (0.00293)	0.00402 (0.00296)
Geography and Climate Controls	Y	Y	Y	Y	Y	Y	Y	Y
Disease Suitability Controls	Y	Y	Y	Y	Y	Y	Y	Y
Colonial Controls	Y	Y	Y	Y	Y	Y	Y	Y
Contemporary Controls	N	Y	N	Y	N	Y	N	Y
p-value: Former British vs. French Coefficients	–	–	0.034	0.045	–	–	0.075	0.085
Bandwidth	100 km	100 km	100 km	100 km	100 km	100 km	100 km	100 km
Observations	2,491	2,491	4,122	4,122	5,853	5,853	8,909	8,909
Clusters	170	170	274	274	170	170	274	274
Mean Dep. Var.	0.766	0.766	0.731	0.731	0.0297	0.0297	0.0603	0.0603

Notes: Data are from the DHS for Cameroon (2004 and 2011). Standard errors are clustered at the DHS cluster level. *Vaccination Index* is the share of vaccines completed out of nine possible vaccines for children in the DHS. *Blood Test Refused* is an indicator variable for refusing to consent to a blood test (for either HIV or anemia) in the DHS data. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. *Relative Suitability: Cassava vs. Millet* is the difference in log suitabilities for cassava and millet at the DHS cluster. *Bandwidth* reports the distance used from the former British and French Cameroon border to restrict the sample. All regressions control for age, age squared, gender, urban-rural status, and include survey-year fixed effects. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index and tsetse fly suitability. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and distance to the nearest mission. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

B.11. Falsification Using Other Sub-Saharan African Countries with DHS Data

We conduct an additional falsification exercise for the proposed instrument (the log of suitability for cassava relative to millet). While the French and Belgian colonial governments organized a series of colonial medical campaigns, the British did not conduct equivalent campaigns in their colonies. Thus, a natural falsification exercise is to examine whether our instrument differentially predicts blood test refusals and lower vaccination rates in the former AEF colonies compared to former British colonies (where no campaigns occurred). This would be true if the exclusion restriction assumption of the instrument holds.

To conduct this exercise, we use data from the IPUMS-DHS Integrated sample for non-French and non-Belgian colonies that also include GPS coordinates. These data in our sample are for Ethiopia (2000, 2005, 2010), Ghana (1993, 1998, 2003, 2008), Kenya (2003, 2008), Lesotho (2004, 2009), Malawi (2000, 2004, 2010, 2016), Namibia (2000), Nigeria (1990, 2003), Sudan (1993, 1997, 2005, 2010), Tanzania (1999, 2010), and Uganda (2000 2006, 2011) IPUMS and ICF (1991-2014). We use the GPS coordinates to construct the instrument for these samples and the geographic and historical covariates included throughout our analysis. For former AEF French colonies, we continue to include: Cameroon (2004 and 2011), Gabon (2012), Congo (2011), Central African Republic (1994) and Chad (1996, 2004, 2014).

Table B17 presents the reduced form estimates for former non-French (and non-Belgian) colonies in columns (1)-(4), for former AEF colonies in columns (5)-(8), and the p-value for the difference in coefficients across the two regions for each reduced form estimate. The results suggest that the instrument has a differential effect on blood test refusal and, to a lesser extent, vaccination rates for the former AEF colonies compared to former non-French colonies, consistent with the exclusion restriction assumption for the instrument.[35]

Table B17: Falsification Exercise with Broader Sample

	Falsification Test for Relative Suitability of Cassava vs. Millet							
	Dep Var.: *DHS Vaccination Index*				Dep Var.: *Blood Test Refused*			
	Non-French Colonies		Former AEF Colonies		Non-French Colonies		Former AEF Colonies	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative Suitability:								
Cassava vs. Millet	0.0137	0.00810	-0.00924	-0.0197	-0.00462	-0.00415	0.0168	0.0133
	(0.00702)	(0.00598)	(0.00889)	(0.00860)	(0.00527)	(0.00516)	(0.00785)	(0.00634)
Geography and Climate Controls	Y	Y	Y	Y	Y	Y	Y	Y
Disease Suitability Controls	Y	Y	Y	Y	Y	Y	Y	Y
Colonial Controls	Y	Y	Y	Y	Y	Y	Y	Y
Contemporary Controls	N	Y	N	Y	N	Y	N	Y
p-value: Former								
Non-French vs. AEF Coefficients	–	–	0.041	0.013	–	–	0.039	0.047
Observations	366,912	362,197	50,773	50,668	186,982	186,982	71,052	71,001
Clusters	180	180	207	207	148	148	160	160
Mean Dep. Var.	0.463	0.464	0.532	0.532	0.128	0.128	0.0483	0.0483

Notes: Data are from the IPUMS-DHS Integrated sample for non-French and non-Belgian colonies with gps coordinates: Ethiopia (2000, 2005, 2010), Ghana (1993, 1998, 2003, 2008), Kenya (2003, 2008), Lesotho (2004, 2009), Malawi (2000, 2004, 2010, 2016), Namibia (2000), Nigeria (1990, 2003), Sudan (1993, 1997, 2005, 2010), Tanzania (1999, 2010), and Uganda (2000 2006, 2011). For former AEF French colonies, the data include: Cameroon (2004 and 2011), Gabon (2012), Congo (2011), Central African Republic (1994) and Chad (1996, 2004, 2014). Standard errors are clustered at the DHS cluster level. *Vaccination Index* is the share of vaccines completed out of nine possible vaccines for children in the DHS. *Blood Test Refused* is an indicator variable for refusing to consent to a blood test (for either HIV or anemia) in the DHS data. *Relative Suitability: Cassava vs. Millet* is the difference in log suitabilities for cassava and millet at the DHS cluster. All regressions control for age, age squared, urban-rural status, and include survey-year fixed effects. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, centroid latitude, centroid longitude and mean altitude. *Disease Suitability Controls* include mean malaria ecology index and tsetse fly suitability. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions. *Contemporary Controls* include educational attainment fixed effects and wealth index fixed effects.

B.12. *World Bank and Afrobarometer Tables and Figures*

Figure B15 presents the binscatters of times visited during the colonial medical campaigns and the project outcome. The first set of figures are for health projects only and the second set of

---

[35]Note that blood test refusals are a better suited and more comparable outcome across countries, especially in this exercise, as the DHS conducts the blood tests rather than local governments. Thus, blood test refusals allow us to abstract from differences in supply and differences in health institutions across former colonial legacies.

figures for non-health projects. We present the binscatters with (i) only country fixed effects and then with (ii) a full set of controls which includes geographic controls, climate controls, disease controls, colonial controls, and contemporary controls. The contemporary control is the total funds committed to each project. These control variables are described in the notes of the figure.

The binscatters reveal striking correlations; for health projects, there is a strong negative correlation between number of times visited and the outcome score assigned to the project. Health projects in areas that had greater exposure the colonial medical campaigns receive lower outcome ratings, as shown in Figures B15a and B15b. However, this is not the case for projects that are in other sectors, as shown in Figures B15c and B15d. In the non-health sectors, projects are on average more successful. This suggests that these places are not somehow less successful at development projects on average, but that they perform poorly specifically in the health sector.

Figure B15: Times Prospected and World Bank Project Outcomes for Health and Non-health Projects
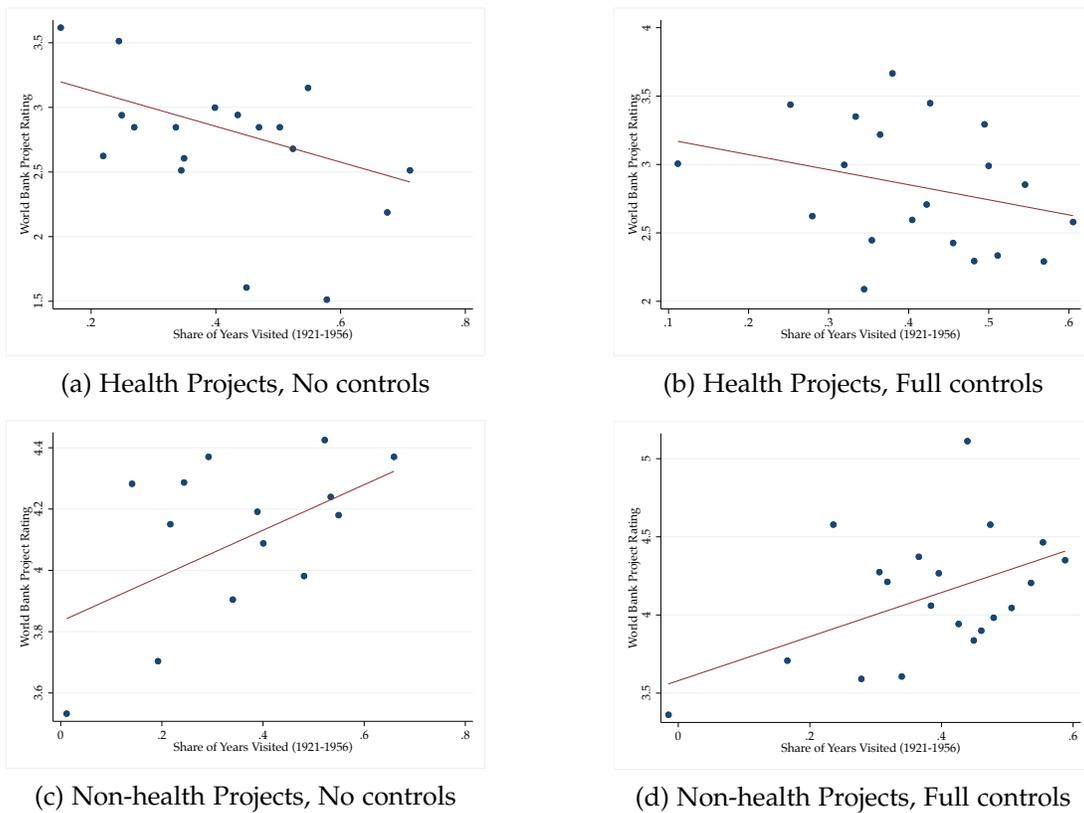
(a) Health Projects, No controls

(b) Health Projects, Full controls

(c) Non-health Projects, No controls

(d) Non-health Projects, Full controls

Notes: Data are from AidData for World Bank aid projects. *World Bank Project Rating* is variable ranging from 1 to 5, where 1=a project was rated as highly unsatisfactory, 2=unsatisfactory, 3=moderately unsatisfactory, 4=moderately satisfactory, and 5=satisfactory. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. *Health Project* is an indicator variable equal to 1 if the project was labeled a "health" sector project by the world bank in the sector designations for a project. All regressions control for country fixed effects and country by health project fixed effects. Full controls include the following set of controls: *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, latitude, longitude and mean altitude of each cluster. *Disease Suitability Controls* include mean malaria ecology index and tsetse fly suitability. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and distance to the nearest mission. *Contemporary Controls* include the total funds committed for each project.

Next, we plot the standardized effect size of historical exposure to medical campaigns and project success by project sector in Figure B16. For the health sector, projects receive a 0.2 standard deviation lower score. For the other sectors, the estimated coefficient is positive, though not

consistently significant.[36] This suggests the negative effects of the campaigns is specific to health projects.

Figure B16: Standardized Effect by Sector
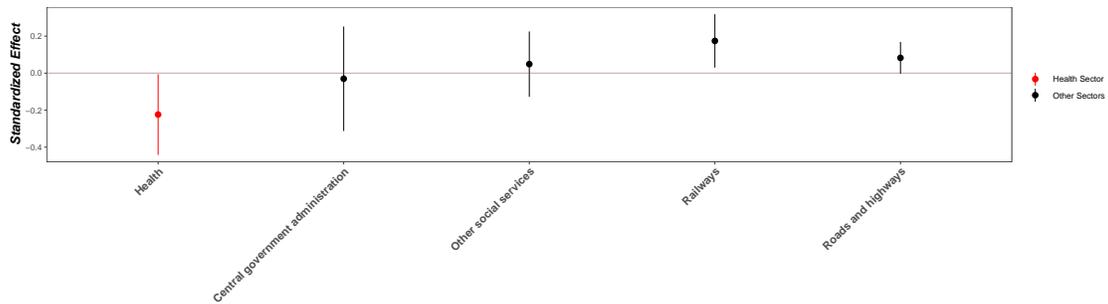
Notes: Data are from AidData for World Bank aid projects. *World Bank Project Rating* is variable ranging from 1 to 5, where 1=a project was rated as highly unsatisfactory, 2=unsatisfactory, 3=moderately unsatisfactory, 4=moderately satisfactory, and 5=satisfactory. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. Projects are classified as being part of a sector if a project includes the sector in any of the 5 world bank sector designations for a project. All regressions control for country fixed effects and country by health project fixed effects, include the following set of controls: *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, latitude, longitude and mean altitude of each project location. *Disease Suitability Controls* include mean malaria ecology index and tsetse fly suitability. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and distance to the nearest mission. *Contemporary Controls* include the total funds committed for each project.

---

[36]We define other sectors the same manner which we defined health projects, where we define a project as being part of one of the listed World Bank sector categories if one of it's five sector categories corresponds to that World Bank sector category. We present the coefficients only for World Bank sectors where we have enough project-sector observations to estimate the full specification presented in Figures B15b and B15d.

Table B18: Colonial Medical Campaigns and Received World Bank Project, Received Health Project, and Project Received Rating

	Panel A: Dep. Var.: *Any World Bank Project*				
	(1)	(2)	(3)	(4)	
Share of Years Visited	0.119	0.136	0.128	0.112	
	(0.182)	(0.186)	(0.193)	(0.183)	
Geography and Climate Controls	Y	Y	Y	Y	
Disease Suitability Controls	N	Y	Y	Y	
Colonial Controls	N	N	Y	Y	
Lasso-Selected Controls	N	N	N	Y	
Observations	280	280	280	280	
Clusters	280	280	280	280	
Mean Dep. Var.	0.471	0.471	0.471	0.471	

	Panel B: Dep. Var.: *Health Project*				
	(1)	(2)	(3)	(4)	(5)
Share of Years Visited	-0.0229	-0.0199	-0.0188	-0.0178	-0.0586
	(0.116)	(0.114)	(0.110)	(0.107)	(0.112)
Geography and Climate Controls	Y	Y	Y	Y	Y
Disease Suitability Controls	N	Y	Y	Y	Y
Colonial Controls	N	N	Y	Y	Y
Contemporary Controls	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	Y
Observations	595	595	595	595	595
Clusters	120	120	120	120	120
Mean Dep. Var.	0.212	0.212	0.212	0.212	0.212

	Panel C: Dep. Var.: *Project Received Rating*				
	(1)	(2)	(3)	(4)	(5)
Share of Years Visited	0.0517	0.0458	0.0304	-0.00414	0.0464
	(0.0986)	(0.104)	(0.107)	(0.108)	(0.0989)
Geography and Climate Controls	Y	Y	Y	Y	Y
Disease Suitability Controls	N	Y	Y	Y	Y
Colonial Controls	N	N	Y	Y	Y
Contemporary Controls	N	N	N	Y	Y
Lasso-Selected Controls	N	N	N	N	Y
Observations	595	595	595	595	595
Clusters	120	120	120	120	120
Mean Dep. Var.	0.361	0.361	0.361	0.361	0.361

Notes: Data are from AidData for World Bank aid projects. Standard errors are clustered at the ethnic group level for Cameroon, at the colonial Sub-District level for Gabon, and at the district level for Congo and Chad. *Any World Bank Project* is an indicator variable equal to one if an area received any World Bank project. *Health Project* is an indicator variable equal to one if an area received a World Bank project in the health sector. *Project Received Rating* is an indicator variable equal to one if an area received a World Bank project that received a rating. *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. *Health Project* is an indicator variable equal to 1 if the project was labeled a "health" sector project by the world bank in the sector designations for a project. All regressions control for country fixed effects. Estimates with *Lasso-Selected Controls* use lasso methods from Belloni et al. (2014a) to select controls from the full set of controls. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, centered longitude and mean altitude of each cluster. *Disease Suitability Controls* include mean malaria ecology index and tsetse fly suitability. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and distance to the nearest mission. *Contemporary Controls* include the total funds committed for each project. Panel A does not include contemporary controls, since if there is no World Bank project, there is no information on total funds committed.

Table B19: Afrobarometer Results

| | Afrobarometer Health Questions: | | | |
| | Health Clinic in PSU | Difficulty to Obtain Treatment | No Contact w/ Clinic/Hospital | Freq. without Medicines/Treatment |
	(1)	(2)	(3)	(4)
Share of Years Visited (1921-1956)	-0.00151	-0.00480	0.00438	0.0198
	(0.00511)	(0.00657)	(0.00148)	(0.00978)
Geography and Climate Controls	Y	Y	Y	Y
Disease Suitability Controls	Y	Y	Y	Y
Colonial Controls	Y	Y	Y	Y
Contemporary Controls	Y	Y	Y	Y
Observations	2,772	1,984	2,772	2,763
Clusters	70	70	70	70
Mean Dep. Var.	0.827	2.526	0.273	1.601

Notes: Data are from the Afrobarometer for Cameroon (Round 5 and Round 6) and Gabon (Round 6). Standard errors are clustered at the ethnic group-district level for Cameroon and at the colonial sub-district level for Gabon. *Health Clinic in PSU* is an indicator variable equal to 1 if the respondent lives in a primary sampling unit (PSU) that has at least one health clinic. *Difficulty to Obtain Treatment* is the difficulty a respondent faced in receiving medical treatment if he/she did have contact with a public clinic or hospital in the previous year (where from 1=Very easy, 2=Easy, 3=Difficult, 4=Very difficult). *No Contact w/ Clinic/Hospital* is an indicator variable equal to 1 if the respondent had no contact with a public clinic or hospital in the previous year. *Freq. without Medicines or Treatment* is the frequency with which a respondent went without medicines or medical treatment in the previous year (where 1=About once every two or three months, 2= About once a month, 3=Two or three times a month, 4= About once a week, 5= Several times a week, and 6= Every day). *Share of Years Visited* measures the share of years the mobile medical teams visited a region for sleeping sickness treatment between 1921 and 1956. All regressions control for age, age squared, gender, urban-rural status and include survey round fixed effects. *Geography and Climate Controls* include mean temperature, mean precipitation, mean land suitability, the mean surface area, centroid latitude, and centroid longitude of each cluster. *Disease Suitability Controls* include mean malaria ecology index, tsetse fly suitability and initial sleeping sickness prevalence. *Colonial Controls* include total number of slaves taken from each main ethnic group in a region during the Atlantic slave trade and number of missions in each main ethnic group in a region. *Contemporary Controls* include educational attainment fixed effects.

Appendix References

Afrobarometer, *Afrobarometer Data [Gabon, Cameroon] [Rounds 5 and 6] [Datasets]*, http://www.afrobarometer.org: Afrobarometer [Distributer], 2013, 2015.

AidData, *WorldBank Geocoded Research Release Level1 v1.4.2 geocoded dataset [Dataset]*, Williamsburg, VA and Washington, DC: AidData [Distributer], 2017. http://aiddata.org/research-datasets.

Alsan, Marcella, "The Effect of the TseTse Fly on African Development," *American Economic Review*, January 2015, *105* (1), 382–410.

__ , "Replication Data for: The Effect of the TseTse Fly on African Development," Inter-university Consortium for Political and Social Research [Distributor] 2015.

BenYishay, A., R. Rotberg, J. Wells, Z. Lv, S. Goodman, L. Kovacevic, and D Runfola, "Geocoding Afrobarometer Rounds 1 - 6: Methodology & Data Quality. AidData.," 2017.

Fischer, Gunther, Harrij van Velthuizen, Freddy Nachtergaele, and Serge Medow, *Global Agro-Ecological Zones [Datasets]*, Rome, Italy: FAO/IIASA [Distributer], 2000.

GADM, *GADM Version 2.5 (various) [Datasets]*, gadm.org [Distributer], 2015. www.gadm.org.

__ , *GADM Version 2.8 (various) [Datasets]*, gadm.org [Distributer], 2015. www.gadm.org.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis, "Very High Resolution Interpolated Climate Surfaces for Global Land Areas," *International Journal of Climatology*, 2005, *25*, 1965–1978.

ICF, *Demographic and Health Surveys (various) [Datasets]*, Rockville, Maryland: ICF [Distributer], 1991-2014. Funded by USAID.

IPUMS and ICF, *IPUMS Demographic and Health Surveys (IPUMS DHS) (various) [Datasets]*, Rockville, Maryland: IPUMS and ICF [Distributer], 1991-2014.

Kiszewski, Anthony, Andrew Mellinger, Andrew Spielman, Pia Malaney, Sonia Ehrlich Sachs, and Jeffrey Sachs, "A Global Index Representing the Stability of Malaria Transmission," *The American Journal of Tropical Medicine and Hygiene*, 2004, *70* (5), 486–498.

Lehner, B., K. Verdin, and A. Jarvis, "New Global Hydrography Derived From Spaceborne Elevation Data," *EOS, Transactions*, 2008, *89* (10), 93–94. www.hydrosheds.org.

Lowes, Sara and Eduardo Montero, "Data and Code for: The Legacy of Colonial Medicine in Central Africa," 2020. American Economic Association (publisher), Inter-university Consortium for Political and Social Research (distributor).

Michalopoulos, Stelios, "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 2012, *102* (4), 1508–39.

Murdock, George Peter, *Africa: It's Peoples and Their Cultural History*, New York, NY: McGraw-Hill Book Company, 1959.

_ , *Ethnographic Atlas*, Pittsburgh, PA: University of Pittsburgh Press, 1967.

Nunn, Nathan, "Religious Conversion in Colonial Africa," *American Economic Review Papers and Proceedings*, 2010, *100* (2), 147–152.

_ **and Leonard Wantchekon**, "Replication Data for: The Slave Trade and the Origins of Mistrust in Africa," Inter-university Consortium for Political and Social Research [Distributor] 2011.

_ **and** _ , "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 2011, *101* (7), 3221–3252.

Ramankutty, Navin, Jonathan A. Foley, John Norman, and Kevin McSweeney, "The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible Climate Change," *Global Ecology and Biogeography*, 2002, *11*, 377–392.

Roome, William, *Ethnographic Survey of Africa: Showing the Tribes and Languages; Also the Stations of Missionary Societies [map]*, Vol. 1:5 1924.

World Health Organization, *Subnational immunization coverage data reported by Member States to WHO and UNICEF through the joint annual data collection process [Dataset]*, WHO [Distributer], 2018.