

Online Appendix to the Paper:
The Long-Term Effects of the Printing Press
in Sub-Saharan Africa

Julia Cagé* and Valeria Rueda†

May 6, 2014

Contents

A Data	2
A.1 Data on Mission Centers Location and Investments	2
A.2 Population Data	6
A.3 Light Data	7
A.4 Afrobarometer Data	9
A.5 Controls	10
A.6 Newspaper Market	12
A.7 Descriptive Statistics	14
B Additional Results	17
B.1 Distance to the Printing Press and Newspaper Readership	17
B.2 Distance to the Printing Press and Economic Development	24
C Robustness Checks	28
C.1 Matching Strategy	28
C.2 Selection on Observables	33
C.3 Additional Robustness Checks	35

*Harvard University, Department of Economics. cage@fas.harvard.edu

†Sciences-Po Paris, Department of Economics. valeria.rueda@sciences-po.org

A Data

A.1 Data on Mission Centers Location and Investments

The *Geography Atlas of Protestant missions*, published in 1903, was the result of an extensive work of localizing all the missionary stations around the world and reporting all the activities they were investing in. Of all the reports, conducted in 1896, 1903, 1911 and 1925, this one (1903) is the most precise and extensive version as investments, denomination, number of students, teachers (both native and foreign) and missionaries are reported for each mission localized on the maps.

The Atlas contains maps of all the regions in the world and locates all the missions active in 1903. Figure A.2 provides an example of these maps. All the missions are uniquely identified in a statistical index providing detailed information on the type of infrastructure available and the number of workers and students. Figure A.1 shows a section of this statistical index.

List of variables available for each mission station

- **Latitude:** Latitude of the mission settlement. This variable was constructed from our geocoding work and is only available for sub-Saharan Africa
- **Longitude:** Longitude of the mission settlement. This variable was constructed from our geocoding work and is only available for sub-Saharan Africa
- **Society:** Society to which the mission is affiliated
- **Denomination** Denomination of the society, if relevant
- **Id:** Unique identifier per mission
- **Name:** Name of the mission's location as reported in the Atlas
- **Printing:** Number of printing presses in the mission
- **Date:** Date of arrival of the mission
- **Anti foot-binding society:** Binary variable equal to one if mission belonged to an anti-foot-binding society
- **Church building:** Number of Churches in the settlement.
- **Blind School:** Binary variable equal to one if mission had a school for the blind or conducted special work for the blind
- **Deafmute school:** Binary variable equal to one if mission had a school for the deaf-mute or conducted special work for the deaf-mute.
- **Female helper:** Number of female helper workers
- **Insane Asylum:** Binary variable equal to one if mission had an insane asylum
- **High School:** Binary variable equal to one if mission had a high school infrastructure.
- **Male Foreign Missionary:** Number of male foreign missionaries working in the mission.

- **Literary Worker:** Number of literary workers.
- **Orphanage:** Binary variable equal to one if mission had an orphanage
- **Sunday-School:** Binary variable equal to one if mission had a Sunday-school
- **Rendez-Vous for training new missionaries:** Binary variable equal to one if mission had a center for training new missionaries
- **Temperance Society:** Binary variable equal to one if mission belonged to a temperance society
- **Number of Students:** Number of students in the mission
- **Foundling asylum or nursery mission:** Binary variable equal to one if mission had a foundling asylum
- **Refuge for opium victims** Binary variable equal to one if mission had a refuge for opium victims
- **Theological school** Number of theological schools in the mission
- **Native male teachers** Number of native male teachers in the mission
- **Medical or nurses' class or school** Binary variable equal to one if mission taught medical or nurse's class
- **Bible society:** Binary variable equal to one if mission was a Bible society
- **Dispensary:** Number of dispensaries in the mission
- **Leper asylum:** Binary variable equal to one if mission had a leper asylum
- **Woman who is unmarried or a widow:** Number of unmarried or widow women in the mission (probably from the country of origin of the mission).
- **Society of young people:** Binary variable equal to one if mission belonged to a society of young people
- **Day School:** Number of elementary school infrastructure
- **College or university:** Number of college or university infrastructure
- **Wife of a missionary:** Total number of women who were the wife of a missionary
- **Hospital** Binary variable equal to one if mission had a hospital
- **Kindergarden** Number of kindergarden infrastructure
- **Anti-opium society** Binary variable equal to one if mission belonged to an anti-opium society

A

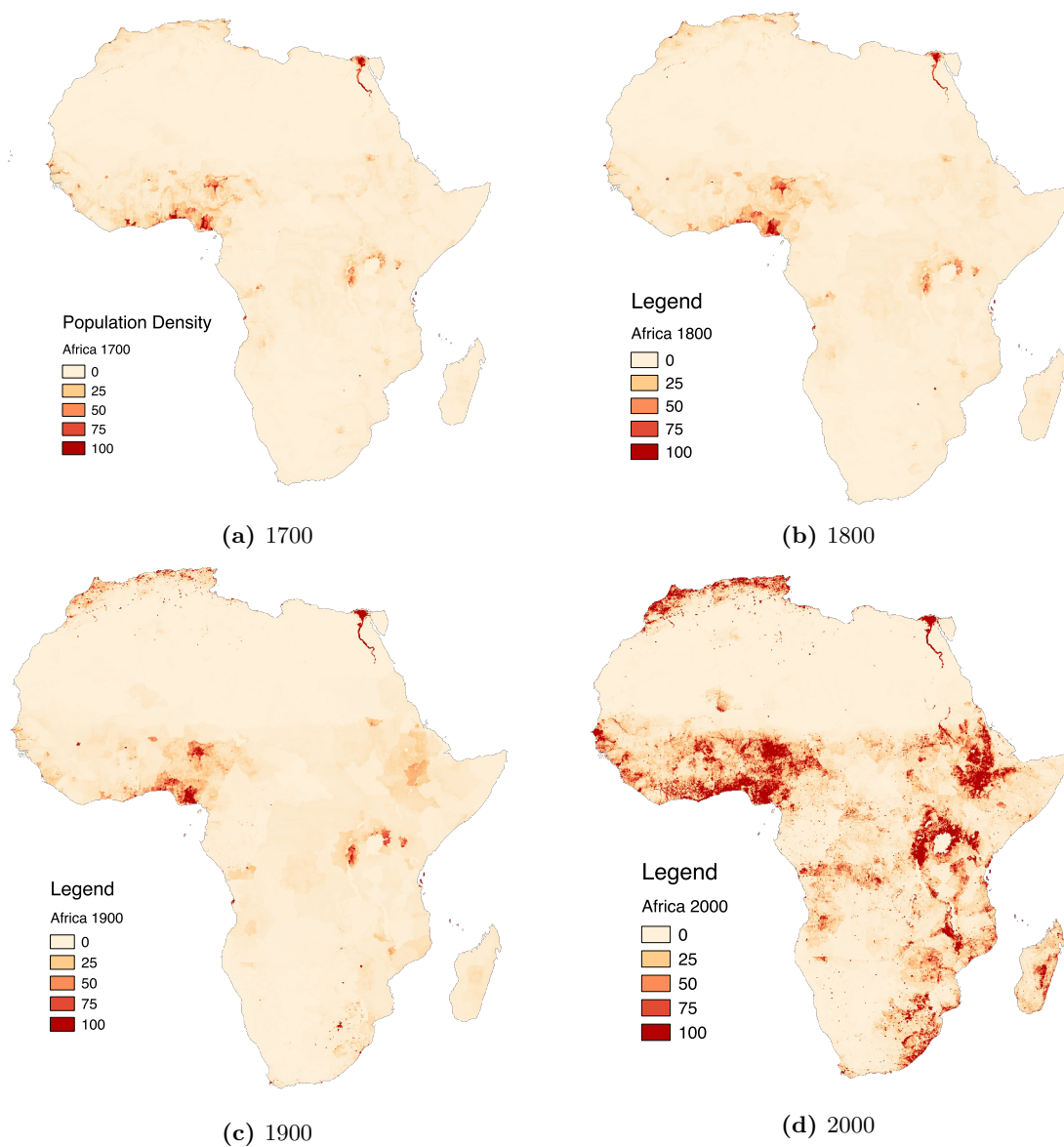
Abaco, see Great Abaco Island
Abadiyeh, 12-Palestine Inset. — **FFMA** (1899) m w 3-N
2-n 3-x=c s 2-v D
Abasa, ? Gold Coast. Plate 14. — **WMS** m 64-Nn 64-O
584-x=10-c 7-s 6-v
Abbotsbad, 11-D1. — **CEZMS** (1894) 3-W n=160 senana
pupils
CMS (1899) 2-m w (p) (1) N 9-x=D
Abetify, 14-D7. — **B** (1876) 2-m 2-w 20-N n 17-O 731-x=s
15-v K (Abetifi)
Abeth, very near Baakleen, 12-Palestine Inset. — See Mt.
Lebanon, work of **FM**
Abeokuta, 14-E7. — **CMS** (1846) 2-m w 2-W 24-N 5-n
1122-x=c 17-v T D L
SBC (1856) m w 2-N n 2-O 26-x=c s v
WMS 2-m 30-Nn 22-O 264-x=7-c 7-s 3-v
Aberdeen, 13-C1. — **BFSS** (1874) 2-m
Abetifi, see Abetify
Abo, see Abos
Abo, 14-F8 Inset. — **MGB** (1900) m=c v (Abo)
Abokobi, see Agbogba
Abors, corner of Tibet. China and Assam, north of Dibru-
garh, 10-H1. — **MPA** (1900) 2-m (1)=D
Aburah, 14-D7. — **WMS** m 49-Nn 75-O 545-x=10-c 10-s 5-v
Aburi, 14-D7. — **B** (1847) 3-m 2-w W 14-N 4-n 9-O
716-x=10-v K
WMS 3-m W 93-Nn 18-O 644-x=15-c 13-s 14-v h
Acca, see Acre
Accra, or Akra, 14-D7. — **MBC** (1900) m 50-x=c s
WMS 3-m 34-Nn 30-O 582-x=4-c 7-s 2-v
Achampatti, 12-E6. — **CMS** N
Achena, ? Ind. Ter., 3 and Inset. — **FNE** (1887) N
26-x=c s
Acre, Akka, or Accho, 12-Palestine Inset, 13-L7. — **CMS**
(1890) 2-m 2-w 4-W (p)=D H (Acra)
Acton Homes, 18-E5. — **WMSA** 29-N 12-O 297-x=2-c 3-s
2-v
Aleppo, 12-K9, 13-M6. — **PCEJ** (1895) 2-m (p) 3-N n=
s B v D
Alert Bay, 2-I6. — **CMS** (1878) 2-m 2-w 2-W N n 22-x
h i
Alessandria, 13-EF4. — **MEN** N O 61-x=s
Alexandria, 12-I10, 13-Ks, 15-G1. — **AJM** No statistic
BFBS 2-m 3-N=B
CSJ 4-m 3-w 3-W=4-v
DAK (1857) 13-W=H
EMS (1898) 7-m
IU (1898) m w=home of rest
LSPCJ No statistics
NAM (1892) 2-m 2-w 2-W N n=c v
UP (1857) 2-m w 2-W=B
WMS m 2-N 45-x=s
Algeciras, 13-B6, 7. — **SSM** (1890) m w N n=c v
Algiers, 13-D6, 14-E1. — **BFBS** (1882) m w 6-Nn=B
FSEJ m
NAM (1892) 2-m 2-w 2-W (1)=2-s
FB m w W
PCFI No statistics
SMF (1887) m w=c s
TM (1888) 5-W
Aligarh, 11-FG5. — **CMS** (1863) m w 3-W 21-N 10-n 3
82-x=c 10-v
MEN 4-m 2-w W 14-N 16-n 720-x=2-c 32-s S 15-v I
Aliwal North, 17-G7, 18-B7. — **FMMS** (1870) 2-m 2-w 5
9-O 1188-x=c s 2-v 2-T I t
Allahabad, 10-A2. — **BFBS** m=B
CMS (1859) 9-m 6-w 12-N n 300-x=c 5-v T
MEN 2-m w 9-N 9-n 67-x=c 27-s S 13-v f o
FN (1836) 4-m 2-w 5-W (p) 18-N 8-n 2-O 150-x=3-c
6-v 3-h D H
WU (1868) 2-W N 31-n=6-s B 14-v Rescue work, a
special services for beggars
YMCA (1900) m=Y
ZEM 6-W N 19-n=5-v h o 260 senana pupils
Allepie, see Alleppey
Allanaw 19-F6 — **WMS** (1894) — (1) 18-N 2-c 5-s

Figure A.1: Part of the Statistical index from the *Geography Atlas of Protestant missions*, 1903



Figure A.2: Plate Representing all the Missionary Stations in 1903, West Africa (*Geography Atlas of Protestant missions, 1903*)

A.2 Population Data



Notes: Population density in 1700 and 1800 varies between 0 and 20 in most of the sub-Saharan territory. Hence the low level of variation observed in the images.

Figure A.3: Population Density from HYDE 3.1

A.3 Light Data

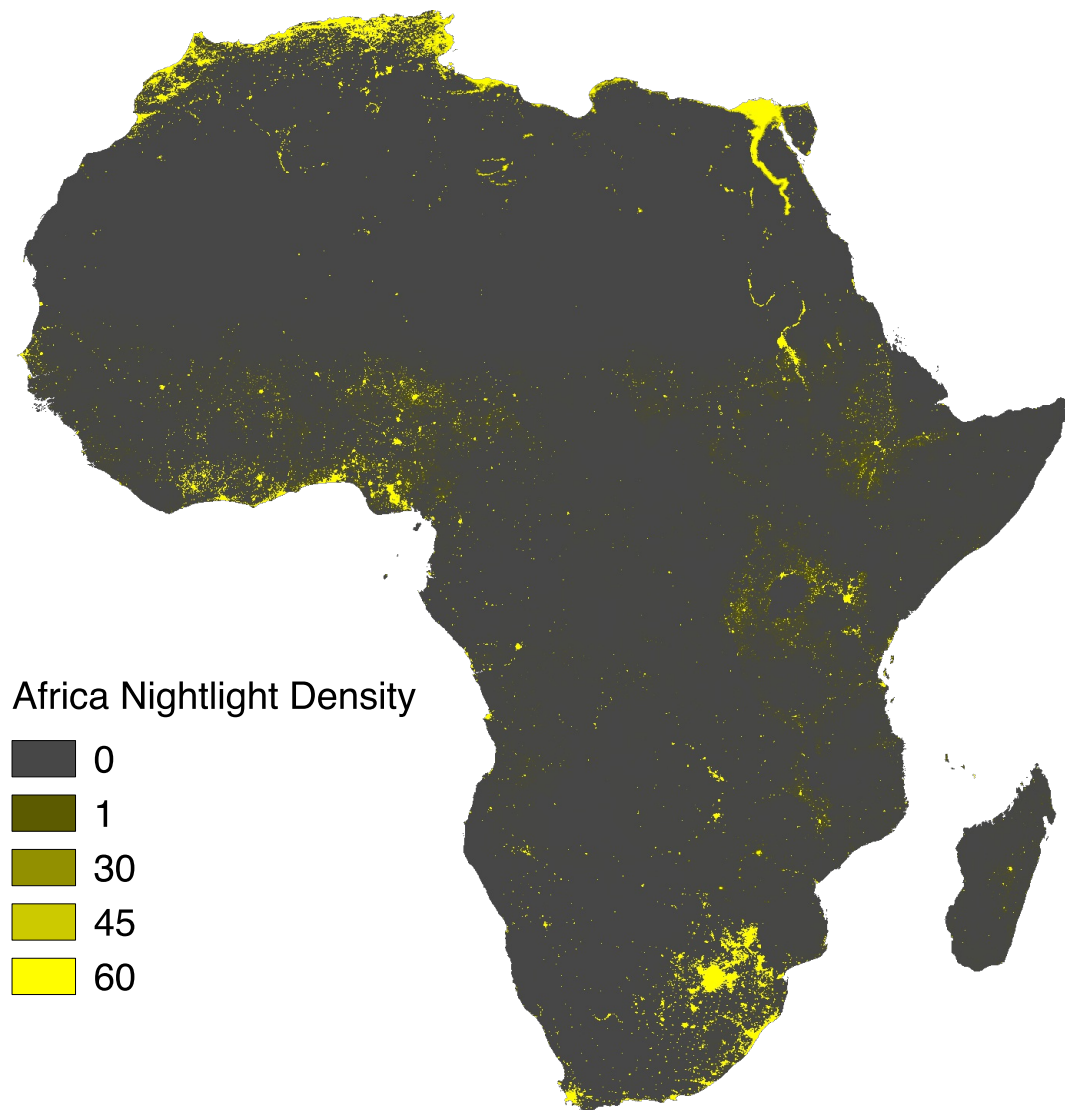
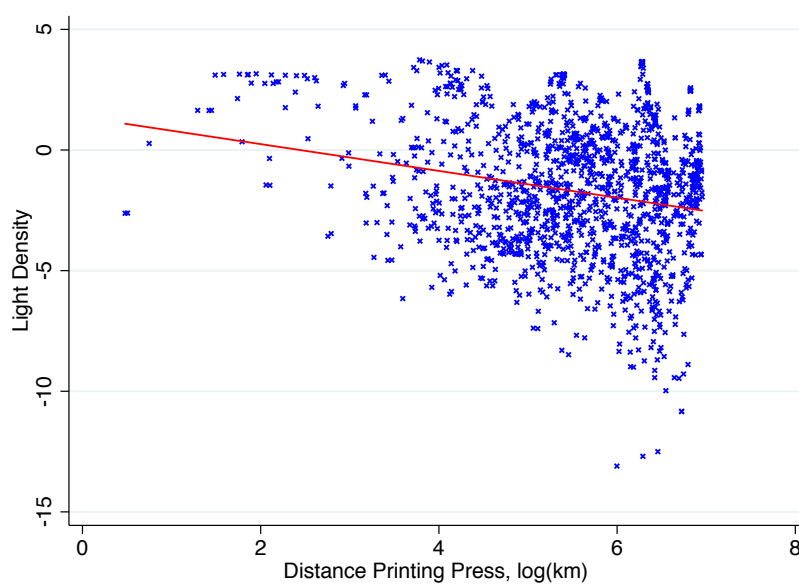


Figure A.4: Light Density at Night in Africa

Figure A.5: Proximity to a Printing Press and Light Density



Notes: The graph plots the average light density in a 10 km buffer around each town (in logs) over the distance to the printing press (in logs). The sample is restricted to towns located closer than 150 km away to a mission.

A.4 Afrobarometer Data

We use data from the 2005 Afrobarometer survey, which is a comparative series of national surveys on attitudes towards democracy, markets, and civil society in Africa. Surveys contain representative information at the individual level of attitudes towards political and social outcomes as well as individual evaluations of living standards. Table A.1 gives a brief description of all the variables from the Afrobarometer used in our analysis.

A precise description of the survey, variables and sampling methods is available on the Afrobarometer website, <http://www.afrobarometer.org>. Similarly, all the questionnaires used for the construction of the dataset are available on the <http://www.afrobarometer.org/survey-and-methods/questionnaires> website.

Table A.1: Variables from Afrobarometer, Description

	Type	Description
Outcomes		
Register	BV	Respondent registered for previous elections
Turnout	BV	Respondent voted at previous elections
Actions	BV	Respondent takes actions as citizens (attends to local meetings, joins others to raise an issue or attends to marches)
Listen	BV	Respondent feels people listen to him for political discussions
Satisfaction	BV	Respondent is satisfied with the country's democratic environment
Read News	BV	Respondent reads news on a monthly basis
Listen News	BV	Respondent listens to the news on the radio at least once a month
Watch News	BV	Respondent watches the news on TV at least once a month
Controls		
Education	DI	Level of education from none (0) to graduate studies (9)
Cash Constraints	DI	How often respondent has gone without cash income in the year from never (0) to always (4).
Water Constraints	DI	How often respondent has gone without enough clean water for home use in the year from never (0) to always (4).
Protestant Today	BV	Individual is Protestant
Urban	BV	Individual lives in an urban area

Notes: DI refers to *Discrete Index*, BV to *Binary Variable*

A.5 Controls

Historical Characteristics

- **Cities 1400:** indicator variable that equals one if there was a city located on the land inhabited by each ethnic group. **Source:** Nunn and Wantchekon (2011) using data from Chandler (1987) on the location of African cities with more than 20,000 inhabitants in year 1400.
- **Distance to 1400 City:** the distance is computed by the authors using ArcGIS. **Source:** The information on the location of cities in 1400 is from Nunn and Wantchekon (2011).
- **Distance to 1800 City:** the distance is computed by the authors using ArcGIS. **Source:** The information on the location of cities in 1800 is from Nunn and Wantchekon (2011).
- **Distance to Catholic mission in 1889:** the distance is computed by the authors using ArcGIS. **Source:** The information on the location of missions in 1889 is from Nunn (2009). He geocoded the original map from Béthune (1889).
- **Explorer contact:** indicator variable that equals one if a European explorer traveled through land historically occupied by the ethnic group (the variable captures exploration routes between 1768 and 1894). **Source:** Nunn and Wantchekon (2011) using information on the location of railway lines in the first decade of the twentieth century from Company (1911).
- **Initial population density:** The average population density estimate in the 18th Century in a buffer of 10 km around each location. From the HYDE 3.1 database.
- **Malaria ecology of the land:** Malaria stability index. The index takes into account the prevalence and type of mosquitoes endemic to a region, their human biting rate, their daily survival rate, and their incubation period. It has been constructed for 0.5-degree-by-0.5-degree grid-cells globally. **Source:** Kiszewski et al. (2004).
- **Railway contact:** indicator variable that equals one if any part of the railway network was built on land historically inhabited by the ethnic group. **Source:** Nunn and Wantchekon (2011) using information on the location of railway lines in the first decade of the twentieth century from Company (1911).
- **Slave exports, per capita:** estimates of the number of slaves taken from each ethnic group disaggregated at the ethnicity level. These estimates were constructed by combining data on the total number of slaves shipped from all ports and regions of Africa with data on the slaves' ethnic identities. The estimates cover Africa's transatlantic and Indian Ocean slave trades between 1400 and 1900. **Source:** Nunn (2008).

Geographic Characteristics

- **Accumulated temperature in the year:** accumulated temperature or temperature sums are calculated for each grid cell by accumulating daily average temperatures for days when average temperature is above a certain threshold temperature. **Source:** Global Agro-Ecological Zones data.

- **Annual precipitation level:** the annual precipitation level data refers to total yearly precipitation. The model used to build this data has been applied considering the baseline period 1961-1990. A detailed description of calculation procedures used to build this data can be found in the GAEZ v.3.0 Global Agro Ecological Zones - Model Documentation (Fischer et al., 2000). **Source:** Global Agro-Ecological Zones data.
- **Average elevation:** in meters. **Source:** Nunn (2008) using data from Parker (1997).
- **Distance to the capital:** the distance is computed by the authors using ArcGIS.
- **Distance to the coast:** the distance to the coast is identified by the authors using the Proximity Utility in ArcGIS.
- **Number of agriculture growing days per year:** number of days during the year when temperature regime and moisture supply are conducive to crop growth and development. This period is also termed the “length of the growing period” (LGP). It is a proxy for the agroclimatic potential productivity of land. The LGP is determined based on prevailing temperatures and water balance calculations for a reference crop. **Source:** Global Agro-Ecological Zones data.
- **Population density in 2005:** population density of the respondent’s neighborhood. **Source:** Afrobarometer (2005).
- **Share of land within 10 km of water:** in percent. **Source:** Nunn (2010) using data from the Digital Chart of the World (DCW).
- **Suitability for rainfed crops:** rainfed suitability has been calculated for nine crop groups that are important to most farming systems in developing countries, namely cereals, fibre crops, oil crops, pulses, roots and tubers, stimulants, sugar crops, tree fruits and vegetables. The algorithm examines in each gridcell all the crop types belonging to a particular crop group. Among these it determines the crop type that maximizes agronomic suitability. **Source:** Global Agro-Ecological Zones data.

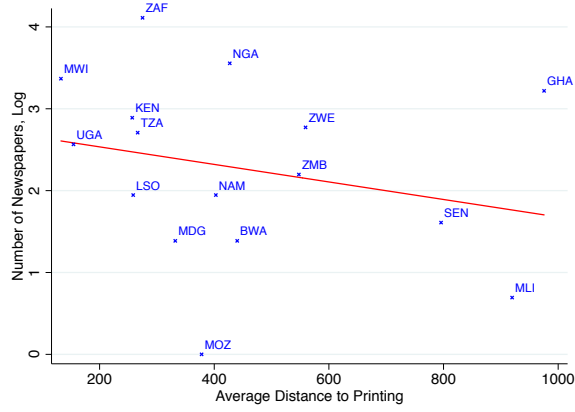
A.6 Newspaper Market

Newspaper Data The cross-country newspaper dataset we build relies on various sources:

- **Tudesq (1995)**: *Feuilles d’Afrique: étude de la presse de l’Afrique sub-saharienne*.
- **Daubert (2009)**: *La presse écrite d’Afrique francophone en question: essai nourri par l’essor de la presse française*.
- The **Standing Conference on Library Material on Africa (SCOLMA)**: it provides an inventory of the available African newspapers in archives collected in 1973.
- The *Directory of African Media* (Maja-Pearce, 1996): digitized by the authors.
- The *African book world and press: a directory* (Zell, 1980): digitized by the authors.
- The *Willings press guide* (Redman and Group, 1993, 2003, 2012): digitized by the authors (one issue every ten years).
- **ICON**: using a web-spider, we collect all the data from the *Icon database* which provides information (first and last date published, title and frequency) on newspapers around the world. This information comes from all the records available in a precise set of institutions: the British Library, the Center for Research Libraries, Harvard University, the Library of Congress, the Library and Archives of Canada, the New York Public Library, the New York State Library, the Online Computer Library Center Inc., the University of Florida, the University of Illinois at Urbana-Champaign, the University of North Carolina at Chapel Hill, the University of Southern California, the University of Washington.
- **Readex**: We collect the summary data from the *Readex database*. Over several decades, Readex has published many of the most widely used collections of primary source research materials in academic libraries, first as Readex Microprint Corporation and since 1984 as a division of NewsBank. These digital collections include information for 40 African newspapers from 1800 to 1922.¹

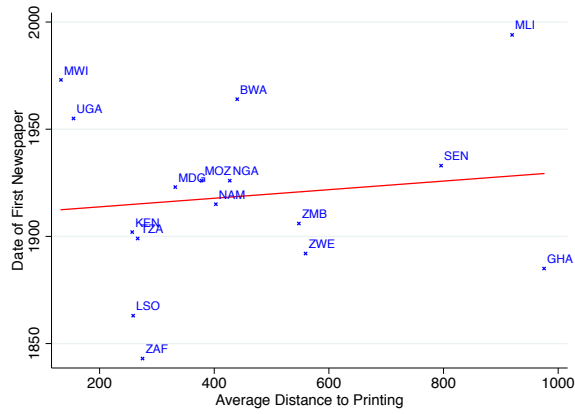
We merge the information from all these datasets/books together to build a consistent database of African newspapers through history. This dataset is a panel from 1800 to 2012 with the number of newspapers published each year in each sub-Saharan country and their date of creation. One caveat of this dataset is that some newspapers (especially small local newspapers) may be missing. However we think that it provides an interesting overview of the state of the newspaper market and of its evolution over time.

¹<http://www.readex.com>



Notes: The x-axis represents the country average distance of Afrobarometer to the closest mission with printing press. The y-axis represents the country total number of newspapers in the dataset. Historical newspaper supply data is described in the Appendix, section A.

Figure A.6: Distance to the Printing Press and Number of Newspapers



Notes: The x-axis represents the country average distance of the villages in the Afrobarometer to the closest mission with a printing press. The y-axis represents the publication date of the first newspaper in the country. Historical newspaper supply data is described in the Appendix, section A

Figure A.7: Distance to the Printing Press and Publication Date of the First Newspaper

A.7 Descriptive Statistics

Table A.2: Summary Statistics, Afrobarometer Data

	All Sample		150km	
	Mean	sd	Mean	sd
Read News	0.331	0.471	0.355	0.479
Listen News	0.858	0.349	0.857	0.350
Watch News	0.451	0.498	0.443	0.497
Education	2.997	1.998	3.176	1.919
Female	0.500	0.500	0.501	0.500
Age	36.519	14.680	36.465	14.698
Cash Constraints	2.099	1.345	2.065	1.350
Water Constraints	1.168	1.397	1.107	1.365
Observations	18469		13286	

Table A.3: Historical and Geographical Characteristics of Printing Presses With and Without a Publication Record

	No Publication	Publication	Diff/se
Geographic Characteristics			
Accumulated Temp /1000	15.320	11.582	3.738 (5.989)
Annual Precipitation/ 1000	1.317	1.434	-0.117 (0.306)
Suitability for Rainfed Crops	5.750	4.000	1.750* (0.704)
Number of Growing Days / 100	1.876	2.429	-0.553 (0.344)
Distance to Capital, 100 km	2.497	2.773	-0.275 (0.855)
Distance to the Coast, 100 km	1.307	3.828	-2.521* (1.102)
Malaria Ecology	7.618	12.302	-4.685 (3.609)
Historical Characteristics			
Slave Exports, per capita	0.058	0.032	0.026 (0.047)
Railway Contact	0.083	0.125	-0.042 (0.122)
Explorer Contact	0.250	0.312	-0.062 (0.179)
Initial Population Density	17.611	10.872	6.739 (8.418)
Distance to 1400 City, 100 km	6.097	7.624	-1.527 (1.678)
Distance to 1800 City, 100 km	8.963	12.010	-3.047 (3.203)
Observations	28		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table compares geographical and historical characteristics of missions with a printing press depending on whether they had a publication record in 1923 or not. Column 1 presents the results for missions with no publication record. Column 2 presents the results for missions with a publication record. In Column 3 we perform a t-test on the equality of means (standard errors in parenthesis). Variables are described in the Online Appendix.

Table A.4: Investments of Printing Presses With and Without a Publication Record

	No Publication	Publication	Diff/se
Mission Characteristics			
Arrival Date	1866	1876	-10 (10)
Bible Society	0.58	0.19	0.40 (0.23)
Number of Native Workers	2.58	1.19	1.40 (2.62)
Total Population	431	400	31 (230)
Investment in Education			
Schools	2.33	1.12	1.21 (0.94)
Number of Students	408	389	19 (228)
Schools per Student (%)	2.52	0.47	2.05 (1.44)
Teachers per Student (%)	12.44	24.00	-11.57 (15.44)
Investments in Health			
Health Facilities	1.67	1.31	0.35 (0.68)
Physicians per Capita (%)	0.04	1.63	-1.59 (1.81)
Health Facilities per Capita (%)	0.50	3.66	-3.16 (3.60)
Observations	28		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table compares investments⁴ of missions with a printing press depending on whether they had a publication record in 1923 or not. Column 1 presents the results for missions with no publication record. Column 2 presents the results for missions with a publication record. In Column 3 we perform a t-test on the equality of means (standard errors in parenthesis). Variables are described in the Online Appendix.

B Additional Results

B.1 Distance to the Printing Press and Newspaper Readership

Table B.1: Distance to a Printing Press and Newspaper Readership, Baseline OLS Estimation, Reporting all the Covariates

	(1)	(2)	(3)	(4)	(5)
	All	All	200km	150km	100 km
	b/se	b/se	b/se	b/se	b/se
Distance Printing Press	-0.014** (0.007)	-0.015** (0.007)	-0.013* (0.007)	-0.015** (0.007)	-0.013* (0.008)
Distance Mission		0.004 (0.005)	0.005 (0.006)	0.000 (0.006)	0.001 (0.007)
Education	0.065*** (0.003)	0.065*** (0.002)	0.066*** (0.003)	0.063*** (0.003)	0.062*** (0.003)
Urban	0.105*** (0.010)	0.105*** (0.010)	0.110*** (0.012)	0.110*** (0.012)	0.099*** (0.013)
Clinic in Town	0.028*** (0.008)	0.028*** (0.008)	0.033*** (0.009)	0.034*** (0.010)	0.042*** (0.011)
Protestant Today	-0.001 (0.009)	-0.001 (0.009)	-0.002 (0.010)	0.007 (0.010)	0.003 (0.011)
Catholic Now	-0.009 (0.009)	-0.009 (0.009)	-0.007 (0.010)	0.004 (0.010)	0.000 (0.011)
Water Constraints	0.005* (0.003)	0.005* (0.003)	0.003 (0.003)	0.002 (0.003)	0.001 (0.004)
Cash Constraints	-0.020*** (0.003)	-0.020*** (0.003)	-0.020*** (0.003)	-0.020*** (0.003)	-0.018*** (0.004)
Age	-0.002* (0.001)	-0.002* (0.001)	-0.002 (0.001)	-0.002 (0.001)	-0.002* (0.001)
Age Squared	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Female	-0.070*** (0.006)	-0.070*** (0.006)	-0.072*** (0.007)	-0.072*** (0.007)	-0.077*** (0.008)
Log Distance Capital City	-0.007 (0.005)	-0.007 (0.005)	-0.008 (0.005)	-0.008 (0.005)	-0.010* (0.006)
Ethnic Pop density 2005	-0.005 (0.009)	-0.006 (0.009)	-0.009 (0.010)	-0.007 (0.011)	-0.010 (0.012)
Slave Export	-0.003 (0.007)	-0.002 (0.007)	0.016* (0.008)	0.014 (0.009)	0.015 (0.010)
Distance to the Coast	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Initial Population Density	0.006 (0.007)	0.006 (0.006)	0.007 (0.007)	0.001 (0.008)	0.002 (0.009)
Pop Density 2005	0.001 (0.004)	0.002 (0.004)	0.002 (0.005)	0.001 (0.005)	0.000 (0.006)
Ethnic Initial Pop Density	-0.001 (0.006)	-0.001 (0.006)	0.002 (0.008)	0.000 (0.009)	0.003 (0.009)
Malaria	0.000 (0.001)	0.000 (0.001)	0.000 (0.002)	-0.000 (0.002)	-0.000 (0.002)
Elevation, over 100	0.005** (0.002)	0.005** (0.002)	0.005* (0.002)	0.004* (0.003)	0.005* (0.003)
Observations	15086	15086	12405	10970	9383
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Additional Controls	No	No	No	No	No
Clusters	1809	1809	1456	1315	1136
R2	0.360	0.360	0.357	0.353	0.357
F-Statistic	205.5	201.6	185.4	174.9	158.3

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports OLS estimates. The unit of observation is an individual. The dependent variable is newspaper readership. Standard errors in parentheses are clustered by village. The baseline controls are the individual-, village-, and ethnicity-level controls described in the text. All specifications include country fixed effects. In columns 1 and 2 we present results for the entire sample. In columns 3 to 5 the sample is sequentially restricted to individuals living 200 km (column 3), 150 km (column 4) and 100 km (column 5) away from a historical mission settlement.

Table B.2: Distance to a Printing Press and Newspaper Readership, Baseline OLS Estimation, Standardized Coefficients (beta Coefficients), Reporting all the Covariates

	(1) All beta/se	(2) All beta/se	(3) 200km beta/se	(4) 150km beta/se	(5) 100 km beta/se
Distance Printing Press	-0.032** (0.007)	-0.034** (0.007)	-0.028* (0.007)	-0.034** (0.007)	-0.031* (0.008)
Distance Mission		0.010 (0.005)	0.012 (0.006)	0.001 (0.006)	0.003 (0.007)
Education	0.274*** (0.003)	0.274*** (0.002)	0.263*** (0.003)	0.250*** (0.003)	0.243*** (0.003)
Urban	0.107*** (0.010)	0.107*** (0.010)	0.110*** (0.012)	0.110*** (0.012)	0.100*** (0.013)
Clinic in Town	0.029*** (0.008)	0.029*** (0.008)	0.034*** (0.009)	0.036*** (0.010)	0.044*** (0.011)
Protestant Today	-0.001 (0.009)	-0.001 (0.009)	-0.002 (0.010)	0.007 (0.010)	0.003 (0.011)
Catholic Now	-0.009 (0.009)	-0.008 (0.009)	-0.007 (0.010)	0.004 (0.010)	0.000 (0.011)
Water Constraints	0.013* (0.003)	0.013* (0.003)	0.008 (0.003)	0.005 (0.003)	0.001 (0.004)
Cash Constraints	-0.056*** (0.003)	-0.056*** (0.003)	-0.056*** (0.003)	-0.057*** (0.003)	-0.050*** (0.004)
Age	-0.055* (0.001)	-0.055* (0.001)	-0.059 (0.001)	-0.057 (0.001)	-0.071* (0.001)
Age Squared	0.042 (0.000)	0.041 (0.000)	0.044 (0.000)	0.043 (0.000)	0.059 (0.000)
Female	-0.074*** (0.006)	-0.074*** (0.006)	-0.075*** (0.007)	-0.075*** (0.007)	-0.080*** (0.008)
Log Distance Capital City	-0.022 (0.005)	-0.022 (0.005)	-0.022 (0.005)	-0.024 (0.005)	-0.031* (0.006)
Ethnic Pop density 2005	-0.014 (0.009)	-0.014 (0.009)	-0.021 (0.010)	-0.018 (0.011)	-0.023 (0.012)
Slave Export	-0.007 (0.007)	-0.005 (0.007)	0.036* (0.008)	0.032 (0.009)	0.035 (0.010)
Distance to the Coast	0.015 (0.000)	0.010 (0.000)	-0.001 (0.000)	-0.004 (0.000)	-0.023 (0.000)
Initial Population Density	0.019 (0.007)	0.018 (0.006)	0.021 (0.007)	0.001 (0.008)	0.006 (0.009)
Pop Density 2005	0.005 (0.004)	0.007 (0.004)	0.008 (0.005)	0.003 (0.005)	0.002 (0.006)
Ethnic Initial Pop Density	-0.004 (0.006)	-0.004 (0.006)	0.006 (0.008)	0.000 (0.009)	0.006 (0.009)
Malaria	0.004 (0.001)	0.003 (0.001)	0.000 (0.002)	-0.002 (0.002)	-0.009 (0.002)
Elevation, over 100	0.054** (0.002)	0.055** (0.002)	0.057* (0.002)	0.052* (0.003)	0.055* (0.003)
Observations	15086	15086	12405	10970	9383
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Additional Controls	No	No	No	No	No
Clusters	1809	1809	1456	1315	1136
R2	0.360	0.360	0.357	0.353	0.357
F-Statistic	205.5	201.6	185.4	174.9	158.3

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports OLS estimates. The unit of observation is an individual. The dependent variable is newspaper readership. The baseline controls are the individual-, village-, and ethnicity-level controls described in the text. All specifications include country fixed effects. In columns 1 and 2 we present results for the entire sample. In columns 3 to 5 the sample is sequentially restricted to individuals living 200 km (column 3), 150 km (column 4) and 100 km (column 5) away from a historical mission settlement.

Table B.3: Missions Characteristics in sub-Saharan Africa and “Society PP”, OLS Estimation

	(1)	
	Society PP	
	b	se
Geographic Characteristics		
Accumulated Temp /1000	-0.000	(0.000)
Annual Precipitation/ 1000	-0.014	(0.033)
Suitability for Rainfed Crops	-0.004	(0.006)
Number of Growing Days / 100	0.001	(0.003)
Distance to Capital, 100 km	-0.008	(0.007)
Distance to the Coast, 100 km	-0.004	(0.004)
Malaria Ecology	0.005	(0.004)
Historical Characteristics		
Slave Exports, per capita	-0.012	(0.019)
Railway Contact	-0.036	(0.034)
Explorer Contact	0.006	(0.023)
Initial Population Density	0.000	(0.000)
Distance to 1400 City, 100 km	-0.006	(0.006)
Distance to 1800 City, 100 km	0.002	(0.002)
Mission Characteristics		
Arrival Date	0.001	(0.001)
Bible Society	-0.076	(0.072)
Number of Native Workers	0.000	(0.002)
Total Population	-0.000	(0.002)
Investment in Education		
Schools	0.030	(0.029)
Number of Students	0.000	(0.002)
Schools per Student (%)	0.011	(0.011)
Teachers per Student (%)	-0.000	(0.000)
Investments in Health		
Health Facilities	0.009	(0.012)
Physicians per Capita (%)	0.000	(0.003)
Health Facilities per Capita (%)	-0.004	(0.004)
Observations	371	
Country FE	Yes	
Clusters (Society)	48	
R2	0.209	

Notes: * p<0.10, ** p<0.05, *** p<0.01. This table reports OLS estimates of mission-level regression of the share of missions from the mission’s society equipped with printing presses in all the regions of the world outside sub-Saharan Africa on mission characteristics, investments, geographical and historical characteristics. The unit of observation is a mission. Standard errors are clustered by society.

Table B.4: Missions' Denomination

	No Printing	Printing	Total
Anglican	25.2	29.6	25.3
Baptist	2.2	11.1	2.5
Disciples of Christ	0.3	0.0	0.3
Free Church	1.5	3.7	1.6
Independent	5.7	0.0	5.5
Lutheran	27.5	14.8	27.0
Methodist	17.0	14.8	16.9
Multidenominational	4.6	0.0	4.4
Presbyterian	8.3	22.2	8.9
Quaker	0.6	0.0	0.6
United Brethern in Christ	4.6	3.7	4.6
Undenominational	1.2	0.0	1.2
United Church	1.2	0.0	1.2
Total	100.0	100.0	100.0

Notes: The table reports the share of the missions in our sample from each denomination. In Column 1 we report this share for the missions without printing press; in Column 2 for the missions with a printing press; and in Column 3 for all the missions taken together.

Table B.5: Determinants of the Probability of Having a Printing Press at the Mission Level, OLS Estimation

	(1)	
	Mission PP	
	b	se
Society PP	0.088**	(0.033)
Geographic Characteristics		
Accumulated Temp /1000	-0.001	(0.001)
Annual Precipitation/ 1000	0.154***	(0.054)
Suitability for Rainfed Crops	-0.009	(0.012)
Number of Growing Days / 100	-0.013***	(0.004)
Distance to Capital, 100 km	0.002	(0.006)
Distance to the Coast, 100 km	0.007	(0.004)
Malaria Ecology	0.005	(0.003)
Historical Characteristics		
Slave Exports, per capita	-0.034**	(0.015)
Railway Contact	-0.031	(0.025)
Explorer Contact	0.024	(0.043)
Initial Population Density	-0.000**	(0.000)
Distance to 1400 City, 100 km	0.004	(0.003)
Distance to 1800 City, 100 km	-0.000	(0.002)
Mission Characteristics		
Arrival Date	-0.001	(0.001)
Bible Society	0.005	(0.057)
Number of Native Workers	-0.014**	(0.006)
Total Population	0.014**	(0.006)
Investment in Education		
Schools	0.015	(0.024)
Number of Students	-0.014**	(0.006)
Schools per Student (%)	-0.007*	(0.004)
Teachers per Student (%)	0.000	(0.001)
Other Investments		
Health Facilities	0.034	(0.032)
Physicians per Capita (%)	0.006	(0.013)
Health Facilities per Capita (%)	0.005	(0.004)
Observations	371	
Country FE	Yes	
Clusters (Society)	48	
R2	0.373	

Notes: * p<0.10, ** p<0.05, *** p<0.01. This table reports OLS estimates of mission-level regression of a binary variable indicating whether the mission has a printing press on mission characteristics, investments, geographical and historical characteristics as well as on the share of missions from the mission's society equipped with printing presses in all the regions of the world outside sub-Saharan Africa. The unit of observation is a mission. Standard errors are clustered by society.

Table B.6: Impact of the Proximity to a Printing Press, IV Estimation. 40 km Radius

	(1) 200km	(2) 150km	(3) 100km
Panel A: Dependent Variable is Distance Printing Press			
Village Printing (40)	-0.609** (0.199)	-0.591** (0.219)	-0.444* (0.233)
Panel B: Dependent Variable is News Readership			
Distance Printing Press	-0.050** (0.019)	-0.046** (0.019)	-0.049** (0.022)
Observations	11925	10583	9059
Baseline Controls	Yes	Yes	Yes
Clusters	1401	1267	1093
R2 First Stage	0.629	0.627	0.625
F First Stage	51.390	52.433	49.811

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The unit of observation is an individual. Standard errors in parentheses are clustered by village. The table reports the first and second stage of the IV estimation of the impact of the distance to a printing press on newspaper readership. The distance to a printing press is instrumented by Village Printing_j(40). The construction of the variables "Village Printing_j(50)" and "Denomination PP_m" is described in more details in the text. The controls are the individual-, village-, ethnicity- and mission-level controls described in the text.

Table B.7: Impact of the Proximity to a Printing Press, IV Estimation. 60 km Radius

	(1) 200km	(2) 150km	(3) 100km
Panel A: Dependent Variable is Distance Printing Press			
Village Printing (60)	-0.498*** (0.150)	-0.479** (0.166)	-0.386** (0.181)
Panel B: Dependent Variable is News Readership			
Distance Printing Press	-0.041** (0.019)	-0.037* (0.019)	-0.038* (0.022)
Observations	11925	10583	9059
Baseline Controls	Yes	Yes	Yes
Clusters	1401	1267	1093
R2 First Stage	0.639	0.635	0.630
F First Stage	58.725	62.539	58.945

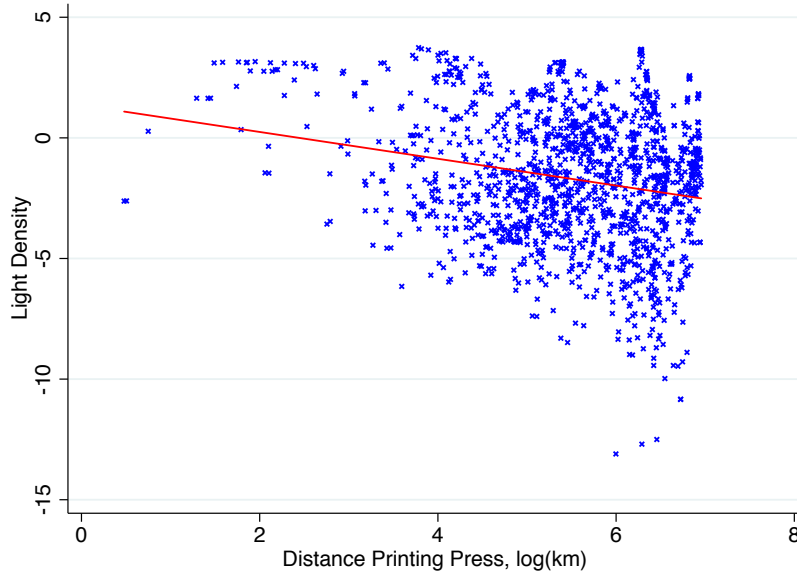
Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The unit of observation is an individual. Standard errors in parentheses are clustered by village. The table reports the first and second stage of the IV estimation of the impact of the distance to a printing press on newspaper readership. The distance to a printing press is instrumented by Village Printing_j(60). The construction of the variables "Village Printing_j(50)" and "Denomination PP_m" is described in more details in the text. The controls are the individual-, village-, ethnicity- and mission-level controls described in the text.

B.2 Distance to the Printing Press and Economic Development

In this section, we develop an alternative method to explore the channels explaining the long-term relationship between distance to the printing press and economic development. We aim at determining whether proximity to the printing press increased economic development through political participation, historical or contemporary development of the newspaper industry. Because of the direct effect of distance to the printing press on newspaper readership, the results from Table B.8 are to be interpreted carefully.

The printing press can affect contemporary economic development through the development of the media, political participation or religiosity. Columns 2 to 8 of Table B.8 examine these potential channels. For each town j , we define a set of binary variables $\mathbf{1}\{\text{High } z\}$ equal to one whenever the value of z is higher than the median in town j . The only variables capturing the effect of the distance to the printing press on economic development are newspaper readership and the share of Protestants in town j (Columns 3, 7, and 8). Protestantism affects negatively economic development nowadays. The interaction term of distance to the printing press and newspaper readership is negative and statistically significant. It is almost identical to the effect in Column 1, suggesting that the development of the culture of information and the bourgeois sphere is the main channel explaining through which the printing press had a long-term impact on economic development.

Figure B.1: Proximity to a Printing Press and Light Density



Notes: The graph plots the average light density in a 10 km buffer around each town (in logs) over the distance to the printing press (in logs). The sample is restricted to towns located closer than 150 km away to a mission.

Table B.8: Proximity to a Printing Press and Light Density

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Distance Printing Press	-0.292** (0.136)	-0.098 (0.164)	-0.108 (0.165)	-0.173 (0.166)	-0.277** (0.136)	-0.280** (0.138)	-0.579*** (0.174)	-0.136 (0.258)
Distance PP with Publication		-0.358** (0.167)						-0.247 (0.177)
Distance PP x 1 {High Read News}			-0.244* (0.130)					-0.224** (0.109)
1 {High Read News}			1.674** (0.751)					1.544** (0.615)
Distance PP x 1 {High Education}				-0.160 (0.130)				-0.033 (0.109)
1 {High Education}				0.973 (0.763)				0.188 (0.626)
Distance PP x 1 {High Action}					-0.041 (0.096)			-0.086 (0.097)
1 {High Action}					0.303 (0.517)			0.532 (0.530)
Distance PP x 1 {High Listen}						-0.021 (0.095)		-0.051 (0.090)
1 {High Listen}						-0.008 (0.520)		0.145 (0.490)
Distance PP x 1 {High Protestant Today}							0.368** (0.149)	0.298* (0.168)
1 {High Protestant Today}							-1.807** (0.850)	-1.387 (0.958)
Observations (towns)	1307	1307	1307	1307	1307	1307	1307	1307
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.612	0.615	0.616	0.613	0.612	0.613	0.616	0.622
Clusters	320	320	320	320	320	320	320	320
F-Statistic	53.93	47.39	60.41	56.92	51.89	54.34	50.37	49.44

Notes: * p<0.10, ** p<0.05, *** p<0.01. The table reports OLS estimates. The unit of observation is the town. The controls are the average individual (per village)-, village-, ethnicity- and mission-level controls described in the text. The regressions are weighted by the population size of each town

Table B.9: Instrumental Variable: Distance to the Printing Press and Light Density - 40 km

	(1) 200km	(2) 150km	(3) 100km
Panel A: Dependent Variable is Distance Printing Press			
Village Printing (40)	-0.828*** (0.240)	-0.805*** (0.241)	-0.768*** (0.227)
Panel B: Dependent Variable is Light Density			
Distance Printing Press	-0.597** (0.263)	-0.681** (0.291)	-0.647** (0.304)
Observations (towns)	1428	1293	1118
Baseline Controls	Yes	Yes	Yes
Clusters	346	343	335
R2 First Stage	0.804	0.800	0.800
F First Stage	140.623	110.505	58.996

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports IV estimates. The unit of observation is the town. The controls are the average individual (per village)-, village-, ethnicity- and mission-level controls described in the text. The dependent variable is the log of average light density in a 10km buffer around the town.

Table B.10: Instrumental Variable: Distance to the Printing Press and Light Density - 50 km

	(1) 200km	(2) 150km	(3) 100km
Panel A: Dependent Variable is Distance Printing Press			
Village Printing (50)	-0.864*** (0.199)	-0.860*** (0.199)	-0.836*** (0.191)
Panel B: Dependent Variable is Light Density			
Distance Printing Press	-0.545** (0.252)	-0.608** (0.279)	-0.570* (0.294)
Observations (towns)	1428	1293	1118
Baseline Controls	Yes	Yes	Yes
Clusters	346	343	335
R2 First Stage	0.806	0.802	0.802
F First Stage	142.030	111.244	91.390

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports IV estimates. The unit of observation is the town. The controls are the average individual (per village)-, village-, ethnicity- and mission-level controls described in the text. The dependent variable is the log of average light density in a 10km buffer around the town.

Table B.11: Instrumental Variable: Distance to the Printing Press and Light Density - 60 km

	(1) 200km	(2) 150km	(3) 100km
Panel A: Dependent Variable is Distance Printing Press			
Village Printing (60)	-0.851*** (0.189)	-0.879*** (0.189)	-0.874*** (0.194)
Panel B: Dependent Variable is Light Density			
Distance Printing Press	-0.517** (0.245)	-0.557** (0.269)	-0.511* (0.285)
Observations (towns)	1428	1293	1118
Baseline Controls	Yes	Yes	Yes
Clusters	346	343	335
R2 First Stage	0.807	0.803	0.803
F First Stage	143.773	117.279	66.379

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports IV estimates. The unit of observation is the town. The controls are the average individual (per village)-, village-, ethnicity- and mission-level controls described in the text. The dependent variable is the log of average light density in a 10km buffer around the town.

C Robustness Checks

C.1 Matching Strategy

Using a Logit model, we regress the binary variable indicating whether missions are endowed with a printing press on all the observable characteristics available at the mission level (these observables correspond to the variables reported in Tables 3 and 4). From this regression we compute the propensity score, which is the estimated probability of having a printing press. We then match each mission with a printing press to the mission with the closest propensity score using a one-to-one matching. The missions matched are extremely similar to the missions with a printing press but had not imported it in 1903. We call the sample of missions similar to those with a printing press but without a printing press \overline{PP} . The sample of missions with a printing press is called PP . Tables C.1 and C.2 show respectively the results of the regression from which we compute the propensity score and the balance check comparing samples PP and \overline{PP} . There are no significant differences between the two samples.

Each city from the Afrobarometer is then associated to the closest mission in the sample $\{\overline{PP} \cup PP\}$. We define a binary variable “Treat” equal to one if a city is associated to a mission that did import the printing press and to zero otherwise. We construct the treatment area as follows: we decompose the map of Africa as a Voronoi diagram using the missions in $\{\overline{PP} \cup PP\}$ as generators (Figure C.1). A Voronoi Diagram is a way of dividing space into a number of regions (Thiessen Polygons). A set of points (called seeds, sites, or generators) is specified beforehand and for each seed there is a corresponding region consisting of all the points closer to that seed than to any other. For each Thiessen Polygon, the binary variable “Printing Treatment” indicates whether the generator point (mission) was indeed endowed with a printing press.

Equation (1) describes the identification equation:

$$\begin{aligned} \text{News}_{ijec} = & \delta_1 \text{Distance Mission}_j + \delta_2 \text{Treat}_j \\ & + \delta_3 \text{Treat}_j \times \text{Distance Mission}_j \\ & + \mathbf{X}'_i \beta_2 + \mathbf{Y}'_j \beta_3 + \mathbf{Z}'_e \beta_4 + \delta_c + u_{ijec} \end{aligned} \quad (1)$$

“Distance Mission” is the distance from town j to the closest mission station in the sample $\{\overline{PP} \cup PP\}$. “Treat * Distance Mission” is the crossed effect of “Treat” and “Distance Mission”. The controls are the same as in Table 5(b).

Table C.3 gives the results of the estimation of equation (1). The effect of proximity to the closest mission from $\{\overline{PP} \cup PP\}$ is also more robust if the closest mission did invest in a printing press. Interestingly, the magnitude of the effect is comparable to the one in the previous section.

Table C.1: Estimation of the Propensity Score: Importing the Printing Press and Missions Characteristics, Logit

	(1)	
	Mission PP	
	b	se
Annual Precipitation level	0.000**	(0.000)
Suitability for Rainfed Crops	-0.016	(0.216)
Number of agriculture growing days	-0.000**	(0.000)
Accumulated Temperature	-0.000	(0.000)
Historical Characteristics		
Distance to 1400 City, 100 km	0.046	(0.073)
Distance to 1800 City, 100 km	-0.048	(0.040)
Mission Characteristics		
Bible Society	-0.196	(0.728)
Number of Native Workers	-0.213**	(0.084)
Total Population	0.157***	(0.056)
Investment in Education		
Schools	0.261	(0.327)
Number of Students	-0.157***	(0.056)
Schools per Student (%)	0.129	(0.110)
Other Investments		
Health Facilities	0.323	(0.301)
Physicians per Capita (%)	0.065**	(0.029)
Health Facilities per Capita (%)	0.010	(0.040)
Observations	508	
Country FE	Yes	Yes
Clusters (Society)	64	
Pseudo R2	0.417	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports the first stage of the matching strategy. This first stage estimates the propensity score at the mission-level. The propensity score is the estimated probability of importing the printing press from the logit regression of the binary variable indicating whether a mission imported the printing press on mission's historical and geographical characteristics as well as mission investments.

Table C.2: Propensity Score Matching: Balancing Test

	No printing	Printing	Diff/se
Annual Precipitation level	1235.559	1484.559	-249.000 (243.329)
Suitability for Rainfed Crops	4.294	4.353	-0.059 (0.586)
Number of agriculture growing days	196.412	240.000	-43.588 (24.257)
Accumulated Temperature	16147.588	11601.765	4545.824 (5797.209)
Historical Characteristics			
Distance to 1400 City, 100 km	7.764	6.886	0.878 (1.825)
Distance to 1800 City, 100 km	12.401	9.555	2.846 (3.107)
Mission Characteristics			
Bible Society	0.353	0.471	-0.118 (0.228)
Number of Native Workers	0.000	2.941	-2.941 (2.078)
Total Population	473.000	507.059	-34.059 (301.932)
Investment in Education			
Schools	1.412	1.941	-0.529 (0.926)
Number of Students	460.882	489.765	-28.882 (300.122)
Schools per Student (%)	2.716	1.951	0.766 (1.731)
Other Investments			
Health Facilities	1.294	1.647	-0.353 (0.505)
Physicians per Capita (%)	1.596	1.533	0.063 (2.074)
Health Facilities per Capita (%)	4.481	3.438	1.044 (3.941)
Observations	34		

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table compares geographical and historical characteristics of missions with a printing press and the missions matched using the propensity score. Column 1 presents the results for missions without a printing press. Column 2 presents the results for missions with a printing press. In Column 3 we perform a t-test on the equality of means (standard errors in parenthesis). Variables are described in the Online Appendix.

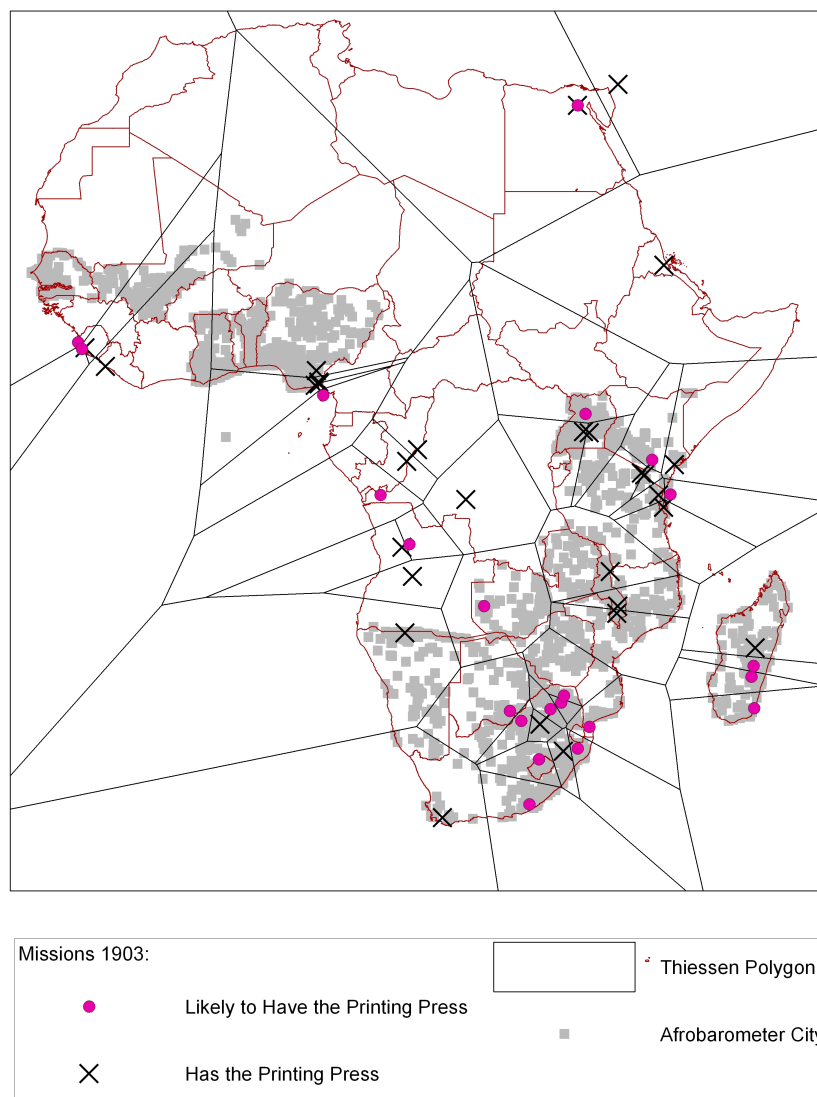


Figure C.1: Voronoi Diagram of Africa using Historical Mission Settlements with the Printing Press and Similar to those with the Printing Press as Generators

Table C.3: Printing Press and Newspaper Readership, Controlling for Observables, Matching Estimation

	(1)	(2)	(3)
	200 km	150 km	100 km
	b/se	b/se	b/se
Treat * Distance Mission	-0.018*	-0.020*	-0.015
	(0.011)	(0.011)	(0.013)
Printing Treatment	0.111*	0.108*	0.080
	(0.060)	(0.062)	(0.067)
Distance Mission	0.001	-0.005	-0.008
	(0.010)	(0.011)	(0.012)
Observations	12405	10970	9383
Baseline Controls	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes
Clusters	1456	1315	1136
R2	0.360	0.357	0.363
F-Statistic	126.6	118.4	108.5

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports OLS estimates. The unit of observation is an individual. The dependent variable is newspaper readership. Standard errors are clustered by village. The baseline controls are the individual-, village- and ethnicity-level controls described in the text and the distance to the closest mission without a printing press. All specifications include country fixed effects. The additional controls are the determinants of the location of mission stations, geographic characteristics, mission characteristics and investments and the distance to the closest mission settlements with different kinds of investments described in the text. In Columns 1 to 3 the sample is sequentially restricted to individuals living 200 km (Column 1), 150 km (Column 2) and 100 km (Column 3) away from a historical mission settlement.

C.2 Selection on Observables

Suppose there is a set of unobservable explanatory variables W' . Since these variables are unobserved, they are not included in regression (1). The *proportional selection assumption* (PSA) states that $\delta C_{WX} = \frac{C_{W'X}}{V_{W'}}$. X is the treatment variable (here distance to the printing press), W is the set of observed covariates, $C_{W'X}$ is the covariance of W' and X and $V_{W'}$ is the variance of W' . δ is a measure of the relationship between C_{WX} and $C_{W'X}$. The PSA assumption is key in the approach as it states that the relationship between the observed covariates W and the treatment X is informative about the relationship between the unobserved covariates W' and X , from which the bias is coming.

Consider the following three regressions:

$$\text{News} = \gamma X + W + W' + \varepsilon_{\max} \quad (\text{M-max})$$

$$\text{News} = \xi X + M + \varepsilon_1 \quad (\text{M-1})$$

$$\text{News} = \Lambda X + W + \varepsilon_2 \quad (\text{M-2})$$

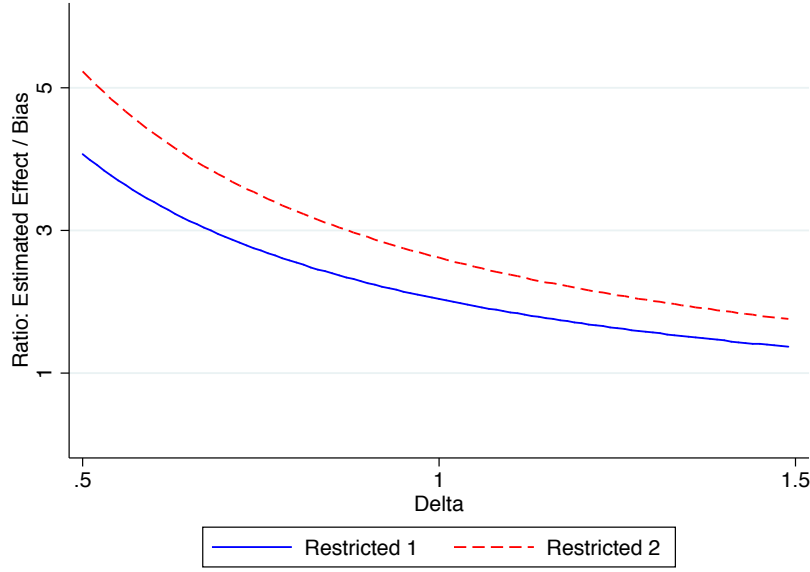
Let R_{\max} be the R-squared of the full model regression (M-max). R_2 is the R-squared of the regression (M-2) including all the observed covariates. R_1 is the R-squared of regression (M-1) including only a restricted set of covariates M . M is a set of observed controls that do not have a related unobserved component and are orthogonal to W and W' (Oster, 2013).

According to Oster (2013), under the PSA and when δ is close to one, $B(\delta) = \delta \frac{(\xi - \Lambda)(R_{\max} - R_2)}{R_2 - R_1}$ is (i) equal to the unobserved bias if $\delta = 1$; (ii) a close upper bound on the bias if $\delta < 1$; (iii) and a close lower bound on the bias if $\delta > 1$.²

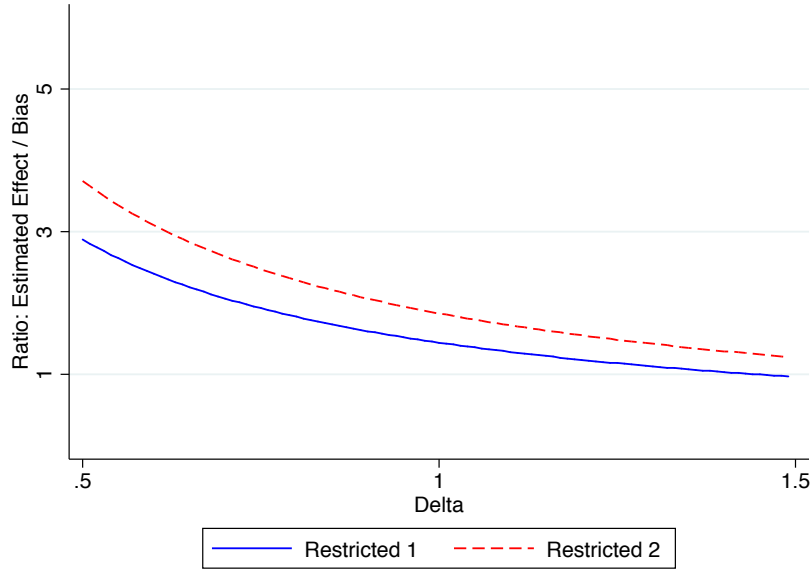
We can the bias due to unobserved variables from the movements in the treatment effect due to the added controls using the ratio $B(\delta)$. However, to compute the ratio it would be necessary to know the true value of R_{\max} . Because there is probably some randomness in the movements of the outcome, it is unlikely that R_{\max} is equal to one. Oster (2013) uses the R-squared from different randomized experiments as a measure of R_{\max} . We cannot use the same approach because we analyze historical events. In the literature of the long-term consequences of historical events, in particular of Protestantism (Becker and Woessmann, 2009) and the diffusion of the printing press (Dittmar, 2011) the R-squared rarely exceed 0.65 and are usually close to 0.5. In our regressions, the R-squared never exceed 0.5. Thus, in our computations we choose 0.6 and 0.7, two conservative values of R_{\max} .

Figure C.2(a) (respectively C.2(b)) plots the ratio of the treatment as given in Table 5(b) over the bias $B(\delta)$ for different values of δ with $R_{\max} = 0.6$ (respectively $R_{\max} = 0.7$). In each graph, two different restricted sets of controls M are chosen. The first set only includes country fixed effects; the second one includes country fixed effects, age, age squared, gender, distance to the closest mission, and distance to the capital city. All the regressions restrict the sample to cities located in 150-km radius of a mission. In all the specifications, for all the values of δ and the two values of R_{\max} , the ratio of the treatment over the bias $B(\delta)$ is higher than one. These results make it unlikely that the entire estimated effect of the distance to the printing press is driven by unobserved variables.

²If W is selected randomly from $\{W, W'\}$, then $\delta = 1$. If W is the most important set of controls from $\{W, W'\}$ then $\delta < 1$.



(a) $R_{\max} = 0.6$



(b) $R_{\max} = 0.7$

Notes: Each graph plots the ratio of the treatment over the bias $B(\delta)$ using two different sets of restricted controls M . The first set (“Restricted 1”) only includes country fixed effects. The second set (“Restricted 2”) includes country fixed effects, age, age squared, gender, distance to the closest mission, and distance to the capital city. Figure C.2(a) (respectively C.2(b)) uses the value $R_{\max} = 0.6$ (respectively $R_{\max} = 0.7$) to compute $B(\delta)$.

Figure C.2: Magnitude of the Treatment Relative to the Bias for Different Values of δ

C.3 Additional Robustness Checks

Table C.4: Printing Press and Newspaper Readership, All Distance Controls

	(1)	(2)	(3)
	200km	150km	100km
	b/se	b/se	b/se
Distance Printing Press	-0.018** (0.009)	-0.023*** (0.009)	-0.021** (0.009)
Distance Mission	0.008 (0.007)	0.003 (0.007)	0.002 (0.008)
Distance Capital City	-0.008 (0.005)	-0.008 (0.005)	-0.014** (0.006)
Distance Coast	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Distance Health Facility	-0.013* (0.008)	-0.011 (0.008)	-0.012 (0.008)
Distance Any Schools	-0.014 (0.012)	-0.016 (0.012)	-0.019 (0.012)
Observations	12405	10970	9383
Baseline Controls	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes
Clusters	1456	1315	1136
R2	0.359	0.356	0.361
F-Statistic	137.6	130.2	120.0

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports OLS estimates. The unit of observation is an individual. The dependent variable is newspaper readership. Standard errors in parentheses are clustered by village. The baseline controls are the individual-, village- and ethnicity-level controls described in the text and the distance to the closest mission without a printing press. All specifications include country fixed effects. The additional controls are the determinants of the location of mission stations, geographic characteristics, mission characteristics and investments and the distance to the closest mission settlements with different kinds of investments described in more detail in the text. In Columns 1 to 3 the sample is sequentially restricted to individuals living 200 km (Column 1), 150 km (Column 2) and 100 km (Column 3) away from an historical mission settlement.

Table C.5: Printing Press and Newspaper Readership, Controlling for Observables, Probit Estimation

	(1)	(2)	(3)
	200km	150km	100km
	b/se	b/se	b/se
Read News			
Distance Printing Press	-0.055* (0.030)	-0.072** (0.030)	-0.065** (0.032)
Observations	12405	10970	9383
Baseline Controls	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes
Clusters	1456	1315	1136
Pseudo R2	0.329	0.325	0.330

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports Probit estimates. The unit of observation is an individual. The dependent variable is newspaper readership nowadays. Standard errors in parentheses are clustered by village. The baseline controls are the individual-, village- and ethnicity-level controls described in the text and the distance to the closest mission without a printing press. All specifications include country fixed effects. The additional controls are the determinants of the location of mission stations, geographic characteristics, mission characteristics and investments and the distance to the closest mission settlements with different kinds of investments described in more details in the text. In Columns 1 to 3 the sample is sequentially restricted to individuals living 200 km (Column 1), 150 km (Column 2) and 100 km (Column 3) away from a historical mission settlement.

Table C.6: Printing Press and Newspaper Readership, Controlling for Observables but not for Potential “Bad Controls”

	(1) 200km b/se	(2) 150km b/se	(3) 100km b/se
Distance Printing Press	-0.028*** (0.011)	-0.034*** (0.011)	-0.033*** (0.012)
Observations	12508	11056	9449
Baseline Controls; Reduced	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes
Clusters	1456	1315	1136
R2	0.220	0.222	0.226
F-Statistic	68.67	64.30	58.67

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports OLS estimates. The unit of observation is an individual. The dependent variable is newspaper readership. Standard errors in parentheses are clustered by village. The baseline controls do not include “bad controls” (namely contemporary possible outcomes: education, religion, water constraints, cash constraints, watching news on TV and listening the news on radio). All specifications include country fixed effects. In Columns 1 to 3 the sample is sequentially restricted to individuals living 200 km (Column 1), 150 km (Column 2) and 100 km (Column 3) away from an historical mission settlement.

Table C.7: Robustness Check: Distance to the Largest Missions

	200km		150km		100km	
	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se
Distance Large Mission	-0.008 (0.009)	-0.009 (0.009)	-0.006 (0.009)	-0.008 (0.009)	-0.002 (0.010)	-0.003 (0.010)
Distance Printing Press		-0.019** (0.009)		-0.024*** (0.009)		-0.021** (0.009)
Observations	12405	12405	10970	10970	9383	9383
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	1456	1456	1315	1315	1136	1136
R2	0.359	0.359	0.355	0.356	0.361	0.361
F-Statistic	137.9	135.3	130.8	128.1	120.1	117.8

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports OLS estimates. The unit of observation is an individual. The dependent variable is newspaper readership. Standard errors in parentheses are clustered by village. The baseline controls are the individual-, village- and ethnicity-level controls described in the text and the distance to the closest mission without a printing press. All specifications include country fixed effects. The additional controls are the determinants of the location of mission stations, geographic characteristics, mission characteristics and investments and the distance to the closest mission settlements with different kinds of investments described in more detail in the text. In Columns 1 to 6 the sample is sequentially restricted to individuals living 200 km (Columns 1 and 2), 150 km (Columns 3 and 4) and 100 km (Columns 5 and 6) away from a historical mission settlement.

Table C.8: Printing Press and Newspaper Readership, Two-Way Clustering at the Mission Level

	(1)	(2)	(3)
	200km	150km	100km
	b/se	b/se	b/se
Distance Printing Press	-0.018*	-0.023**	-0.021**
	(0.010)	(0.010)	(0.011)
Observations	12405	10970	9383
Baseline Controls	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes
Cluster 1: Village	1456	1315	1136
Cluster 2: Mission	358	356	348
R2	0.359	0.356	0.361
F-Statistic	78.17	83.36	94.56

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports OLS estimates. The unit of observation is an individual. The dependent variable is newspaper readership. Standard errors are two-way clustered at the closest mission- and village-levels. The baseline controls are the individual-, village- and ethnicity-level controls described in the text and the distance to the closest mission without a printing press. All specifications include country fixed effects. The additional controls are the determinants of the location of mission stations, geographic characteristics, mission characteristics and investments and the distance to the closest mission settlements with different kinds of investments described in more detail in the text. In Columns 1 to 3 the sample is sequentially restricted to individuals living 200 km (Column 1), 150 km (Column 3) and 100 km (Column 3) away from a historical mission settlement.

Table C.9: Printing Press and Newspaper Readership, Former British Colonies

	(1)	(2)	(3)
	200km	150km	100km
	b/se	b/se	b/se
Distance Printing Press	-0.031***	-0.032***	-0.034***
	(0.010)	(0.009)	(0.010)
Observations	8760	7900	6948
Baseline Controls	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes
Clusters	1025	938	839
R2	0.331	0.326	0.327
F-Statistic	132.3	116.3	107.0

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports OLS estimates. The unit of observation is an individual. The dependent variable is newspaper readership. The sample of countries is reduced to former British colonies. Standard errors in parentheses are clustered by village. The baseline controls are the individual-, village- and ethnicity-level controls described in the text and the distance to the closest mission without a printing press. All specifications include country fixed effects. The additional controls are the determinants of the location of mission stations, geographic characteristics, mission characteristics and investments and the distance to the closest mission settlements with different kinds of investments described in more detail in the text. In Columns 1 to 3 the sample is sequentially restricted to individuals living 200 km (Column 1), 150 km (Column 3) and 100 km (Column 3) away from a historical mission settlement.

Table C.10: Falsification Test: Radio and Television

	Radio News			TV News		
	(1)	(2)	(3)	(4)	(5)	(6)
	200km	150km	100km	200km	150km	100km
	b/se	b/se	b/se	b/se	b/se	b/se
Distance Printing Press	0.003 (0.005)	0.003 (0.006)	0.004 (0.006)	0.009 (0.010)	0.011 (0.010)	0.012 (0.010)
Observations	12405	10970	9383	12405	10970	9383
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	1456	1315	1136	1456	1315	1136
R2	0.111	0.118	0.124	0.423	0.431	0.433
F-Statistic	15.74	16.52	15.34	159.4	189.5	186.9

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports OLS estimates. The unit of observation is an individual. In Columns 1 to 3 the dependent variable is listening to the news on the radio at least once a month. In Columns 4 to 6 the dependent variable is watching the news on TV at least once a month. Standard errors in parentheses are clustered by village. The baseline controls are the individual-, village- and ethnicity-level controls described in the text and the distance to the closest mission without a printing press. All specifications include country fixed effects. The additional controls are the determinants of the location of mission stations, geographic characteristics, mission characteristics and investments and the distance to the closest mission settlements with different kinds of investments described in more detail in the text. The sample is sequentially restricted to individuals living 200 km (Columns 1 and 4), 150 km (Columns 2 and 5) and 100 km (Columns 3 and 6) away from a historical mission settlement.

References

- Becker, Sascha O and Ludger Woessmann**, “Was Weber Wrong? A Human Capital Theory of Protestant Economic History,” *Quarterly Journal of Economics*, 2009, 124 (2), 531–596.
- Béthune, Léon**, *Les Missions Catholiques d’Afrique*, Desclee, 1889.
- Chandler, Thomas**, *Four Thousand Years of Urban Growth: An Historical Census*, St. David’s University Press, 1987.
- Company, Century**, *The Century Atlas: Africa [map]*, Buffalo, NY: Matthews-Northryp, 1911.
- Daubert, Pierre**, *La presse écrite d’Afrique francophone en question: essai nourri par l’essor de la presse française* Etudes Africaines, Paris: l’Harmattan, 2009.
- Dittmar, Jeremiah E**, “Information Technology and Economic Change: The Impact of The Printing Press,” *The Quarterly Journal of Economics*, 2011, 126 (3), 1133–1172.
- Fischer, G, H van Velthuisen, and F O Nachtergaele**, *Global Agro-Ecological Zones Assessment: Methodology and Results*, Vol. IR-00-064, Laxenburg, Austria: IIASA - International Institute for Applied Systems Analysis, November 2000.
- Kiszewski, A, A Mellinger, A Spielman, P Malaney, S E Sachs, and J Sachs**, “A global index representing the stability of malaria transmission,” *American Journal of Tropical Medicine and Hygiene*, 2004, 70 (5), 486–498.
- Maja-Pearce, Adewale**, *Directory of African Media*, International Federation of Journalists, 1996.
- Nunn, Nathan**, “The Long Term Effects of Africa’s Slave Trade,” *Quarterly Journal of Economics*, 2008, 123 (1), 139–176.
- , “Christians in Colonial Africa,” Mimeo, Harvard University 2009.
- , “Religious Conversion in Colonial Africa,” *American Economic Review Papers and Proceedings*, May 2010, 100 (2), 147–152.
- **and Leonard Wantchekon**, “The Slave Trade and the Origins of Mistrust in Africa,” *American Economic Review*, December 2011, 101 (7), 3221–3252.
- Oster, Emily**, “Unobservable Selection and Coefficient Stability: Theory and Validation,” Technical Report w19054, NBER Working Paper 2013.
- Parker, P**, *National cultures of the world: a statistical reference* ABC-Clio ebook, GREENWOOD Publishing Group Incorporated, 1997.
- Redman, C and Gale Group**, *Willings Press Guide 1993* Willing’s Press Guide, Gale/Cengage Learning, 1993.
- **and –** , *Willings Press Guide 2003* Willing’s Press Guide, Gale / Cengage Learning, 2003.
- **and –** , *Willings Press Guide: 2012*, Cision UK, 2012.

Tudesq, André-Jean., *Feuilles d'Afrique: Etude de la presse de l'Afrique sub-saharienne* Publications de la Maison des sciences de l'homme d'Aquitaine, Editions de la maison des sciences de l'homme d'Aquitaine, 1995.

Zell, Hans M., *The African book world and press : a directory*, 2nd ed., Zell London, 1980.