

Online Appendix to the Paper
Sex and the Mission: The Conflicting Effects of Early Christian
investments on sub-Saharan Africa’s HIV Epidemic

Julia Cagé¹ and Valeria Rueda*²

¹Sciences Po Paris and CEPR

²School of Economics, University of Nottingham

September 24, 2019

Contents

A Data	2
A.1 List of the countries included in the analysis	2
A.2 Controls	3
B Additional figures	4
C Additional tables	10
C.1 Baseline analysis	10
C.2 Robustness	16
C.3 Channels	30
D Reverend Lattimer Fuller’s letter	39

*Correspondng author: [valeria\[dot\]rueda\[at\]nottingham\[dot\]ac\[dot\]uk](mailto:valeria[dot]rueda[at]nottingham[dot]ac[dot]uk)

A Data

A.1 List of the countries included in the analysis

This is the list of the countries present in our analysis. This list includes all the countries present in our sample. These are countries with both HIV biomarkers, GPS coordinates, and at least one historical missionary investment in health recorded in our sample.

1. Burundi in 2010.
2. Cameroon in 2004 and 2011.
3. Congo (Democratic Republic) in years 2007 and 2013.
4. Ethiopia in years 2005 and 2010.
5. Gabon in 2012
6. Guinea in years 2005 and 2012.
7. Kenya in years 2003 and 2008.
8. Lesotho in years 2004, 2009, and 2014.
9. Malawi in years 2010 and 2015.
10. Rwanda in years 2005 and 2010.
11. Senegal in years 2005, and 2010
12. Sierra Leone in 2008 and 2013
13. Swaziland in 2006.
14. Tanzania in 2003 and 2007.
15. Uganda in 2010.
16. Zambia in years 2007 and 2013.
17. Zimbabwe in 2005, 2010, and 2014.

A.2 Controls

Historical characteristics

- **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).
- **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).
- **Precolonial population:** 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:** 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).
- **Tsetse fly suitability index:** standardized value of the steady-state tsetse population from Alsan (2015).

Geographic characteristics Geographic characteristics are extracted from the DHS geocovariates files. The first edition of the documentation, which describes all the variables can be downloaded from the spatial data repository of the DHS website (Mayala et al., 2018).

B Additional figures

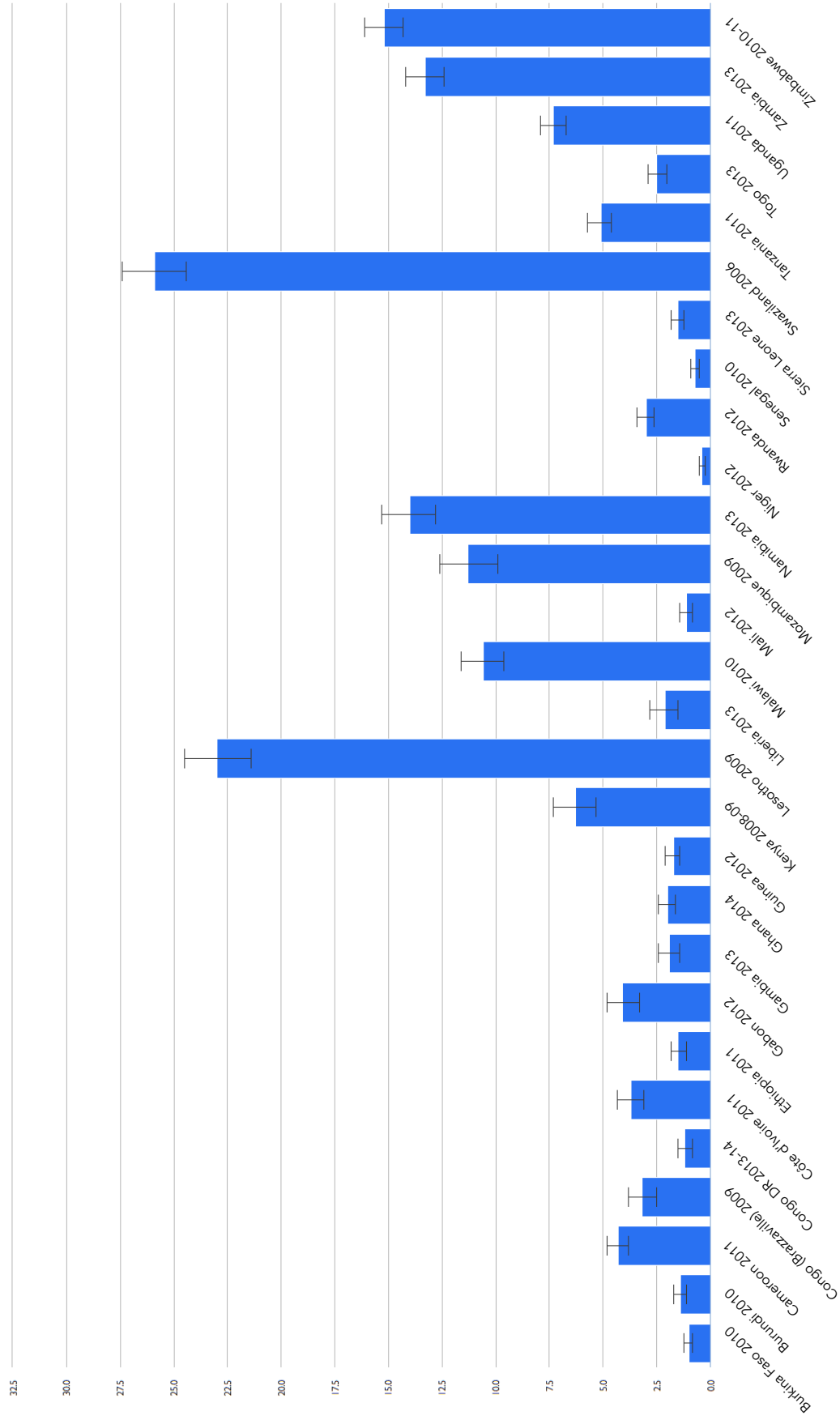
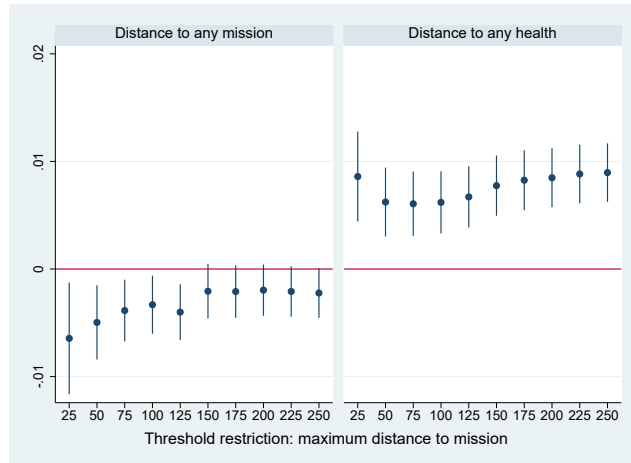
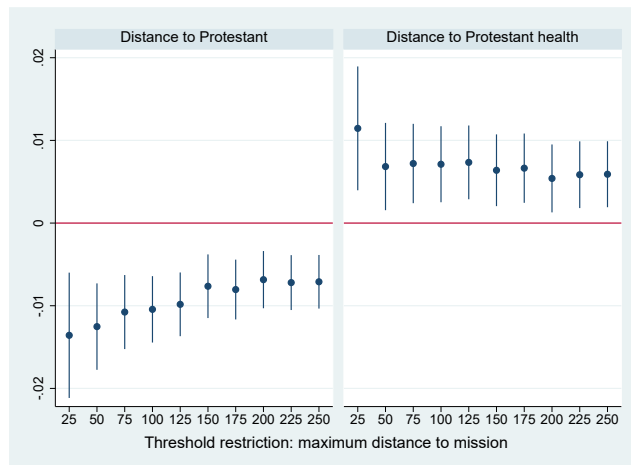


Figure B.1: HIV prevalence rates by country

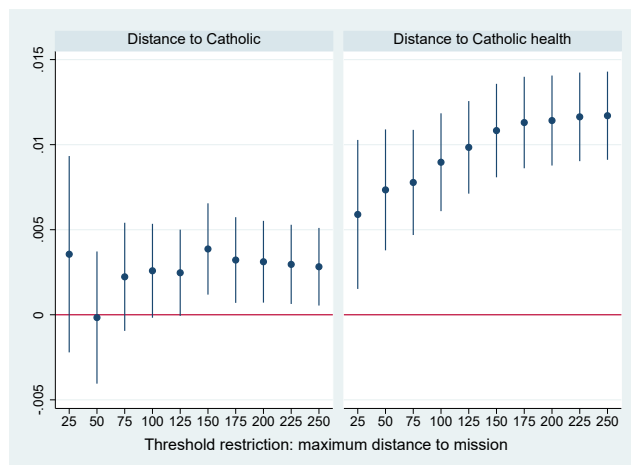
Notes: The figure shows the average prevalence rates from different sub-Saharan African countries. The estimates are those of the DHS using the most recent biomarker data for each country (extracted from statcompiler).



(B.1) Any Mission



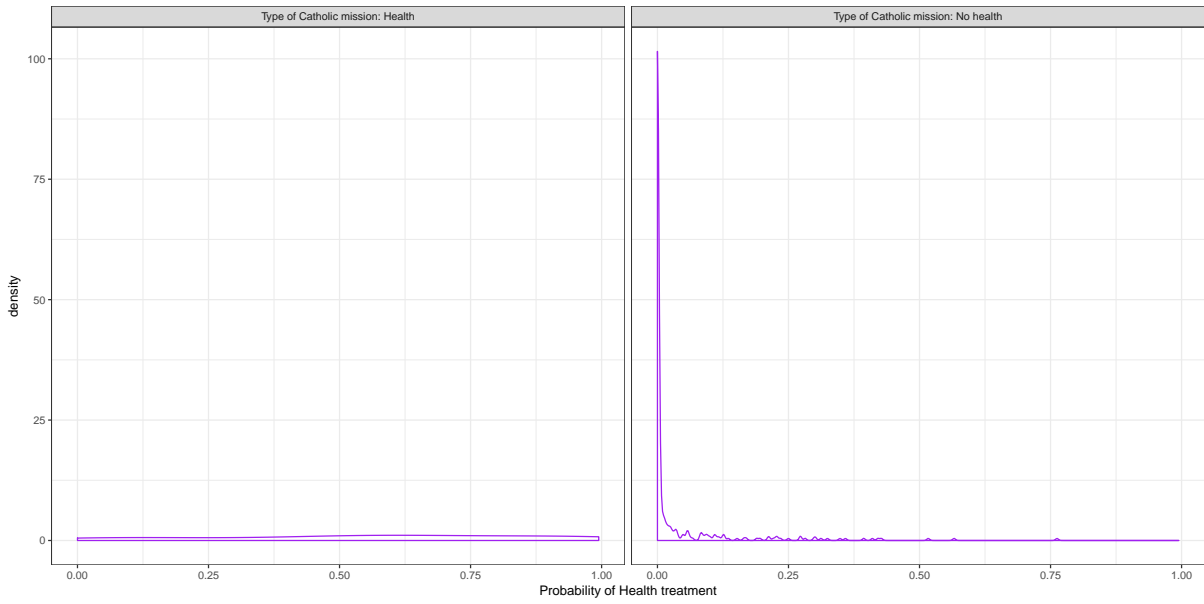
(B.2) Protestant



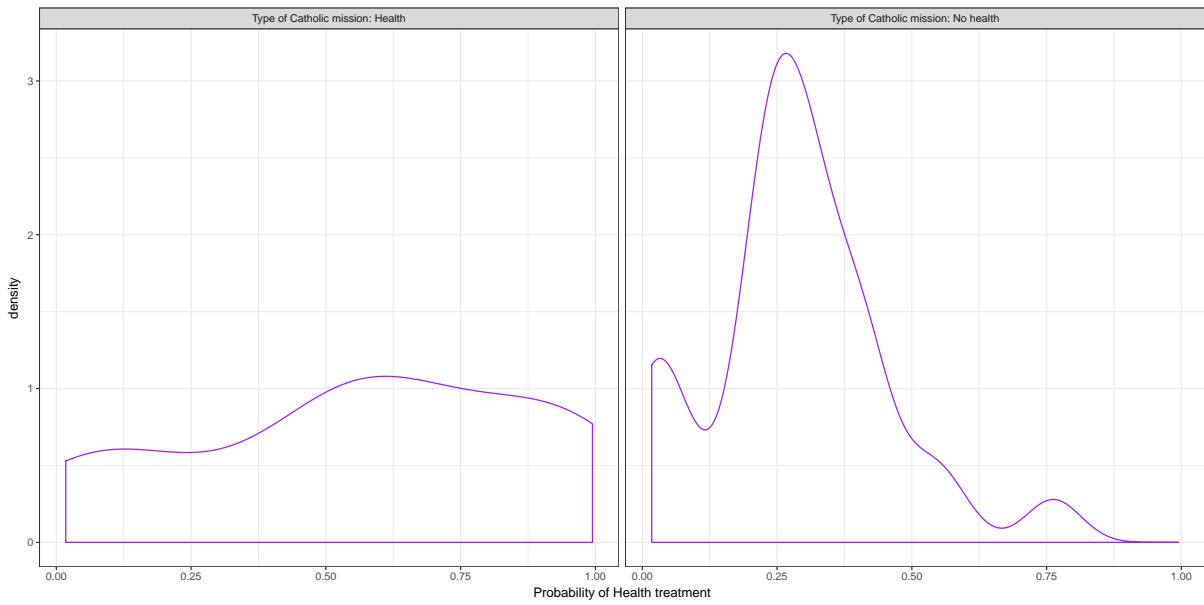
(B.3) Catholic

Figure B.2: OLS estimates of baseline specification using different distance thresholds

Notes: The figures reports OLS estimates. The unit of observation is the individual. The outcome variable is a binary variable equal to one if the individual is HIV positive. Standard errors are clustered at the town level. The x -axis shows different sample restrictions: each time, the sample is restricted to areas that are within a x km buffer around a mission. The vertical lines represent 90% confidence intervals.



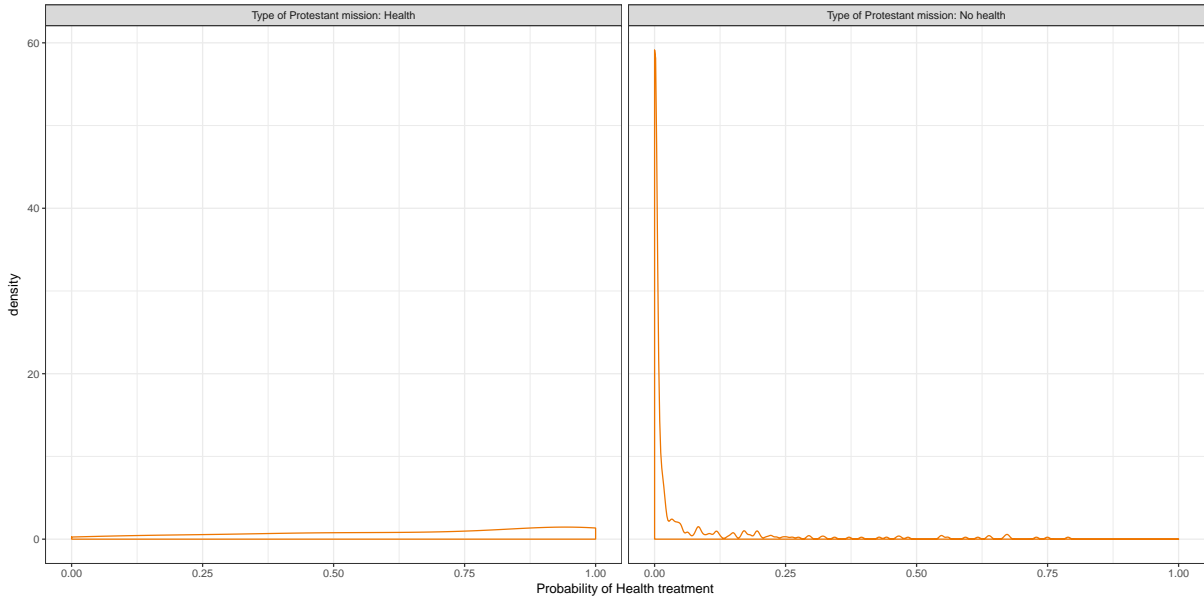
(B.1) Before matching



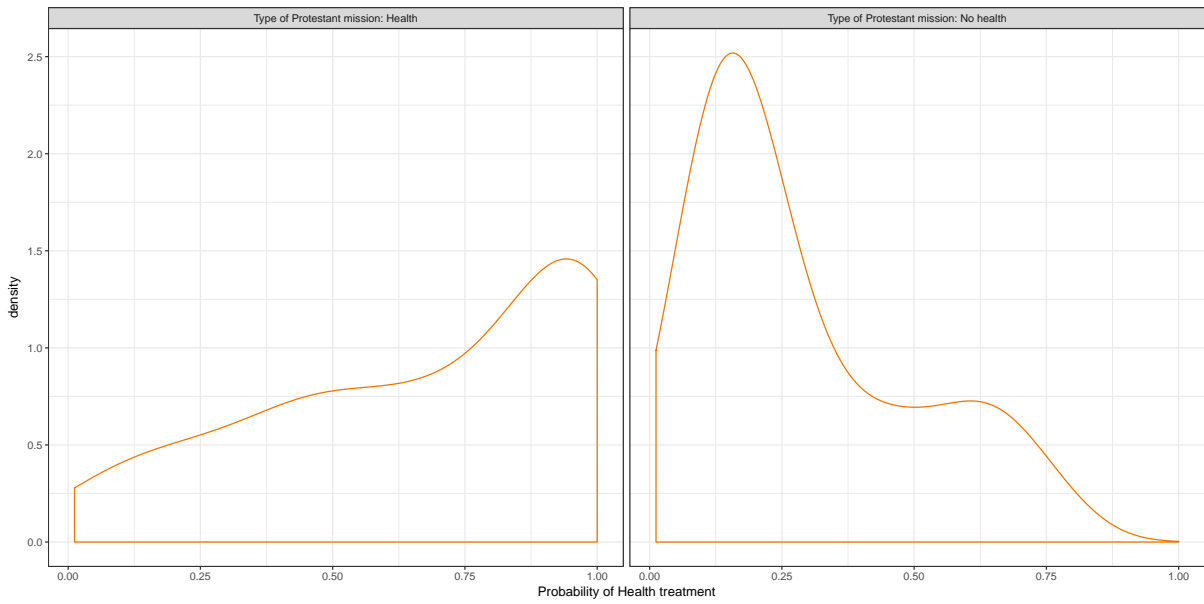
(B.2) After matching

Figure B.3: Density of propensity score before and after matching for Catholic missions

Notes: The figures show the distribution of propensity scores before and after matching for Catholic missions



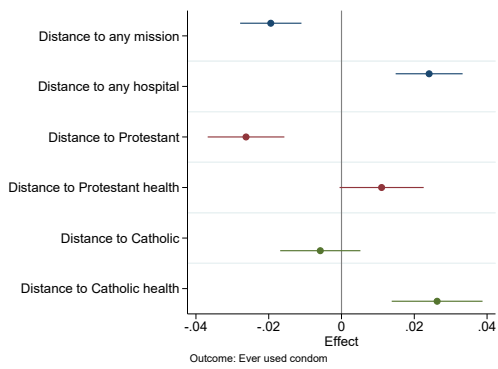
(B.1) Before matching



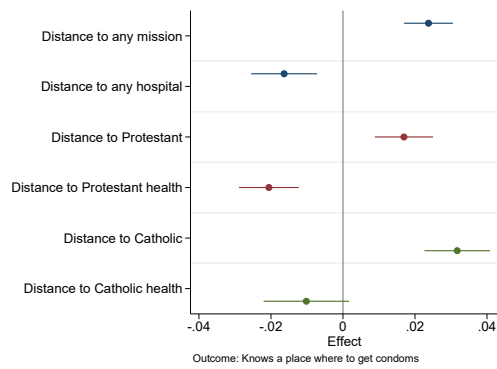
(B.2) After matching

Figure B.4: Density of propensity score before and after matching for Protestant missions

Notes: The figures show the distribution of propensity scores before and after matching for Protestant missions



(B.1) Ever used condoms



(B.2) Knows where to find condoms

Figure B.5: Missionary investments and condom use

Notes: The figures report OLS estimates. The point represents the estimated effect and the lines represent the 90% confidence interval. The unit of observation is an individual. The dependent variables vary for each graph and are reported in the legend. Standard errors are clustered at town level. Controls are described in the text. All regressions restrict the analysis to regions at least 100 km close to missions

C Additional tables

C.1 Baseline analysis

Table C.1: Descriptive statistics: Distances

	All			Less than 100 km		
	(1)			(2)		
	Mean	Median	sd	Mean	Median	sd
Distance to any mission	79.1	47.0	99.5	37.4	31.7	25.9
Distance to Protestant mission	242.9	171.7	232.9	191.9	106.8	215.7
Distance to Catholic mission	96.0	65.9	104.4	56.6	45.9	48.8
Distance to any health	324.1	244.4	274.3	254.7	187.1	242.8
Distance to Protestant health	425.5	325.8	319.9	362.6	248.9	315.3
Distance to Catholic health	545.9	486.0	395.6	476.5	403.7	369.8
Observations	12465			9236		

Notes: The Table reports descriptive statistics. The unit of observation is the town. Column (1) reports the statistics for all the towns in the DHS included in our sample. Column (2) reports similar statistics but the sub-sample of towns located less than 100 km away from a mission. Variables are described in the text.

Table C.2: Descriptive statistics: HIV & Anemia prevalence

	All			Less than 100 km		
	(1)			(2)		
	Mean	Median	sd	Mean	Median	sd
HIV prevalence	0.069	0.021	0.105	0.076	0.030	0.112
Anemia prevalence	0.043	0.024	0.061	0.038	0.000	0.054
Observations	12465			9236		

Notes: The Table reports descriptive statistics. The unit of observation is the town. Column (1) reports the statistics for all the towns in the DHS included in our sample. Column (2) reports similar statistics but the sub-sample of towns located less than 100 km away from a mission. Variables are described in the text.

Table C.3: Descriptive statistics: DHS towns close and far to any mission

Variable	(1) Far from mission	(2) Close to mission	(3) Diff
Hours to large settlement	5.146 (4.601)	3.228 (3.293)	-1.918 (0.077)***
Distance to water	169.622 (115.972)	115.210 (103.294)	-54.411 (2.284)***
Growing Season Length	8.321 (2.826)	8.984 (2.530)	0.662 (0.056)***
TSI	0.153 (1.055)	-0.457 (1.338)	-0.609 (0.027)***
Malaria 2000	0.310 (0.185)	0.334 (0.179)	0.024 (0.004)***
Malaria 2005	0.277 (0.201)	0.258 (0.178)	-0.018 (0.004)***
Malaria 2010	0.203 (0.175)	0.225 (0.196)	0.022 (0.004)***
Malaria 2015	0.172 (0.142)	0.155 (0.124)	-0.018 (0.003)***
Slope	1.644 (1.962)	2.300 (2.256)	0.656 (0.048)***
Drought Episodes	4.779 (2.725)	5.899 (2.780)	1.120 (0.065)***
Rainfall 1985	1,187.675 (530.724)	1,290.490 (693.811)	102.815 (14.387)***
Rainfall 1995	1,092.507 (551.374)	1,170.528 (700.960)	78.021 (14.585)***
Rainfall 2005	1,086.308 (524.316)	1,086.607 (614.897)	0.299 (12.936)
Rainfall 2015	1,088.803 (526.387)	1,132.310 (644.644)	43.507 (13.479)***
Temperature November	23.535 (3.152)	22.111 (3.822)	-1.424 (0.080)***
Temperature February	23.688 (3.047)	22.210 (3.403)	-1.478 (0.072)***
Temperature May	23.299 (4.686)	20.231 (5.312)	-3.067 (0.112)***
Temperature August	21.641 (3.693)	19.345 (4.866)	-2.296 (0.101)***
Distance to railway	172.258 (153.354)	77.740 (86.378)	-94.518 (2.165)***
Distance to explorer route	229.551 (250.157)	185.274 (191.269)	-44.276 (4.284)***
Distance to city 1400	653.650 (316.124)	802.443 (358.361)	148.794 (7.342)***
Distance to city 1800	990.214 (627.455)	1,304.666 (945.856)	314.452 (18.681)***
Precolonial polygamous	0.925 (0.263)	0.989 (0.102)	0.064 (0.003)***
Precolonial nuclear family	0.168 (0.374)	0.062 (0.241)	-0.106 (0.006)***
Precolonial taboo on sex bf marriage (women)	0.189 (0.391)	0.211 (0.408)	0.022 (0.008)***
Observations	2,846	11,374	14,220

Notes: The Table reports descriptive statistics. The unit of observation is the town. Column (1) reports the statistics for towns located in a buffer of 100km around any mission. Column (2) reports similar statistics for those outside the buffer. In column (3), we perform a t-test on the equality of means (robust standard errors are in parentheses).

Table C.4: Missionary investments and HIV Prevalence: logit regression

	(1)	(2)	(3)
	b/se	b/se	b/se
HIV positive			
Distance to any mission	-0.078** (0.031)		
Distance to any health	0.132*** (0.030)		
Distance to Protestant		-0.161*** (0.034)	
Distance to Protestant health		0.175*** (0.036)	
Distance to Catholic			0.029 (0.039)
Distance to Catholic health			0.101** (0.049)
Observations	173,264	100,510	149,724
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	6,689	4,060	5,654
Pseudo R-sq	0.20	0.18	0.21
Mean DepVar	0.07	0.09	0.07
Sd DepVar	0.26	0.29	0.25

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports logit estimates. The unit of observation is the individual. The dependent variable is the a binary variable equal to one if the respondent is HIV positive. Standard errors in parentheses are clustered at the town level. Controls are described in the text.

Table C.5: Missionary investments and HIV Prevalence: Standardized (beta) coefficients

	(1)	(2)	(3)
	beta/se	beta/se	beta/se
Distance to any mission	-0.011*** (0.001)		
Distance to any health	0.028*** (0.001)		
Distance to Protestant		-0.034*** (0.002)	
Distance to Protestant health		0.027*** (0.002)	
Distance to Catholic			0.008** (0.001)
Distance to Catholic health			0.041*** (0.001)
Observations	170,153	96,616	149,672
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
R-sq	0.10	0.10	0.10
Mean DepVar	0.07	0.09	0.07
Sd DepVar	0.26	0.29	0.25

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is a binary variable equal to one if the respondent is HIV positive. Controls are described in the text.

Table C.6: Missionary Investments and HIV Prevalence: Baseline Specification with bad controls

	All sample			Mission \leq 100 km		
	(1)	(2)	(3)	(4)	(5)	(6)
	b/se	b/se	b/se	b/se	b/se	b/se
Distance to any mission	-0.002 (0.001)			-0.003* (0.002)		
Distance to any health	0.009*** (0.002)			0.006*** (0.002)		
Distance to Protestant		-0.003* (0.002)			-0.012*** (0.002)	
Distance to Protestant health		0.006*** (0.002)			0.007*** (0.003)	
Distance to Catholic			0.003** (0.001)			0.003 (0.002)
Distance to Catholic health			0.012*** (0.002)			0.009*** (0.002)
Observations	240,122	240,122	240,122	173,263	100,509	149,723
Country and Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	8,873	8,873	8,873	6,689	4,060	5,654
R-sq	0.10	0.10	0.10	0.10	0.10	0.10
Mean DepVar	0.07	0.07	0.07	0.07	0.09	0.07
Sd DepVar	0.26	0.26	0.26	0.26	0.29	0.25

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is a binary variable equal to one if the respondent is HIV positive. Standard errors in parentheses are clustered at the town level. Controls are the ones described in the text, adding binary indicators for the respondent's level of education (none, primary, or high school), as well as binary indicators of wealth percentiles provided by the DHS.

Table C.7: Matching regression for Catholic missions

	<i>Dependent variable:</i>
	Health
Growing days	0.034** (0.016)
Precipitation	-0.005 (0.004)
Malaria ecology within 40 km	-0.238** (0.104)
Distance to city	-1.164 (1.274)
Distance to coast	0.397 (0.880)
Distance to rail	1.132 (1.182)
Distance to explorer route	-1.261 (1.324)
Precolonial population density	-1.094 (1.359)
Distance city 1400	0.001 (0.003)
Distance City 1800	-1.577 (1.263)
Catholic administrator's residence	6.832** (3.192)
Archbishop's residence	4.715* (2.565)
archbish_1	0.866 (2.092)
Country FE	Yes
Precolonial FE	Yes
Observations	376
Log Likelihood	-41.240
Akaike Inf. Crit.	198.480

Note:

*p<0.1; **p<0.05; ***p<0.01

Notes: * p<0.10, ** p<0.05, *** p<0.01. The Table reports logit estimates. The unit of observation is the Catholic mission. The dependent variable is a binary variable equal to one if the mission invested in health.

C.2 Robustness

Table C.8: Matching regression for Protestant missions

	<i>Dependent variable:</i>
	Health
Rainfed crop suitability	0.233 (0.164)
Av. precipitation per day	0.101 (0.195)
Malaria ecology within 40km	-0.015 (0.054)
Percent of growing days per year	0.001 (0.016)
Distance to capital	0.109 (0.221)
Distance to coast	-0.285 (0.178)
Railway contact	-1.041 (0.930)
Explorer contact	0.439 (0.685)
Initial pop density	0.009 (0.009)
Distance to city 1400	0.205 (0.154)
Distance to city 1800	-0.283* (0.158)
Bible society	2.100 (1.703)
Native workers	-6.498 (634.102)
Mission population	0.463*** (0.095)
Number of students	-0.462*** (0.095)
Printing press	0.542 (1.019)
School	0.208 (0.485)
Schools per Student (pct)	-1.041** (0.488)
Country FE	Yes
Precolonial FE	Yes
Denomination FE	Yes
Observations	463
Log Likelihood	-74.291
Akaike Inf. Crit.	250.582

Note:

*p<0.1; **p<0.05; ***p<0.01

Notes: * p<0.10, ** p<0.05, *** p<0.01. The Table reports logit estimates. The unit of observation is the Catholic mission. The dependent variable is a binary variable equal to one if the mission invested in health.

Table C.9: Balancing table after matching procedure, Catholic missions

Variable	(1) No health	(2) Health	(3) Diff
Growing days	246.342 (93.929)	204.880 (87.054)	-41.462 (24.202)*
Precipitation	1,263.129 (475.727)	1,087.107 (458.419)	-176.022 (124.852)
Malaria ecology within 40km	9.745 (8.149)	9.105 (6.980)	-0.640 (2.028)
Distance to city	0.844 (0.370)	0.754 (0.531)	-0.090 (0.122)
Distance to coast	1.432 (0.967)	1.178 (1.022)	-0.254 (0.266)
Distance to rail	0.591 (0.579)	0.437 (0.545)	-0.154 (0.150)
Distance to explorer route	0.669 (0.576)	0.633 (0.491)	-0.035 (0.143)
Precolonial population density	27.882 (123.145)	59.648 (203.010)	31.766 (44.872)
Distance to city 1400	6.416 (0.707)	6.027 (1.114)	-0.389 (0.249)
Distance to city 1800	7.275 (0.326)	7.170 (0.417)	-0.105 (0.100)
Catholic administrator's residence	0.036 (0.189)	0.107 (0.315)	0.071 (0.069)
Archbishop's residence	0.036 (0.189)	0.071 (0.262)	0.036 (0.061)
Propensity score	0.288 (0.168)	0.566 (0.326)	0.278 (0.069)***
Precolonial polygamy	1.000 (0.000)	1.000 (0.000)	0.000 (0.000)***
Observations	28	28	56

Notes: The Table reports summary statistics for the matched control group (column (1)), and the treatment group (column (2)). Column (3) reports the differences, and the significance of the t-test (* p<0.10, ** p<0.05, *** p<0.01).

Table C.10: Balancing table after matching procedure, Protestant missions

Variable	(1) No health	(2) Health	(3) Diff
Rainfed crop suitability	5.353 (1.673)	5.147 (1.730)	-0.206 (0.292)
Av. precipitation per day	3.090 (2.083)	3.019 (1.830)	-0.070 (0.336)
Malaria ecology within 40km	5.822 (6.509)	8.256 (7.760)	2.433 (1.228)**
Percent of growing days per year	53.582 (27.892)	53.457 (24.622)	-0.125 (4.512)
Distance to capital	2.707 (1.663)	2.767 (1.696)	0.059 (0.288)
Distance to coast	2.804 (3.483)	3.435 (3.909)	0.631 (0.635)
Railway contact	0.088 (0.286)	0.191 (0.396)	0.103 (0.059)*
Explorer contact	0.162 (0.371)	0.294 (0.459)	0.132 (0.072)*
Initial pop density	9.375 (26.602)	14.678 (52.773)	5.304 (7.167)
Distance to city 1400	8.356 (3.877)	7.436 (4.089)	-0.921 (0.683)
Distance to city 1800	16.000 (7.571)	14.841 (6.729)	-1.159 (1.228)
Bible society	0.044 (0.207)	0.118 (0.325)	0.074 (0.047)
Mission population	325.529 (516.143)	393.618 (762.333)	68.088 (111.642)
Number of students	321.603 (515.158)	385.603 (760.035)	64.000 (111.345)
Printing press	0.059 (0.237)	0.132 (0.341)	0.074 (0.050)
School	0.353 (0.686)	0.647 (0.894)	0.294 (0.137)**
Schools per Student (pct)	0.299 (0.891)	0.319 (0.704)	0.020 (0.138)
Propensity score	0.287 (0.208)	0.677 (0.308)	0.389 (0.045)***
Precolonial polygamy	1.000 (0.000)	1.000 (0.000)	0.000 (0.000)***
Observations	68	68	136

Notes: The Table reports summary statistics for the matched control group (column (1)), and the treatment group (column (2)). Column (3) reports the differences, and the significance of the t-test (* p<0.10, ** p<0.05, *** p<0.01).

Table C.11: Selection on observables compared to selection on unobservables

	Any hospital	Protestant Hospital	Catholic Hospital
No Controls	0.69	0.49	0.59
Year and Country FE	1.31	1.59	1.24

Notes: The table reports the absolute values for the estimated selection on observables compared to selection on unobservables for the coefficient on distance to a health facility. These are computed using the stata module `?` by Oster (2016). The first specification (first line) does not include any unrelated controls. The second specification adds year and country fixed effects. The R-max chosen is 30% higher than the R-squared in our regressions (0.13).

Table C.12: Missionary investments and refusal to test

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.0031 (0.0025)		
Distance to any hospital	-0.0001 (0.0025)		
Distance to Protestant		0.0004 (0.0032)	
Distance to Protestant health		0.0012 (0.0030)	
Distance to Catholic			-0.0027 (0.0030)
Distance to Catholic health			0.0023 (0.0029)
Observations	118,145	78,437	101,143
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	7,263	5,074	6,096
R-sq	0.04	0.04	0.04
Mean DepVar	0.064	0.071	0.066
Sd DepVar	0.245	0.257	0.248

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is a binary variable equal to one if the respondent in the female sample refused the HIV test. Standard errors in parentheses are clustered at the town level. All regressions restrict the analysis to regions at least 100 km close to missions. Controls are described in the text.

Table C.13: Missionary investments and education

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.321*** (0.031)		
Distance to any hospital	-0.015 (0.030)		
Distance to Protestant		-0.194*** (0.033)	
Distance to Protestant health		0.015 (0.032)	
Distance to Catholic			-0.339*** (0.035)
Distance to Catholic health			0.040 (0.034)
Observations	344,170	218,880	297,927
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	8,963	6,030	7,729
R-sq	0.39	0.39	0.40
Mean DepVar	-0.16	0.44	-0.23
Sd DepVar	3.77	3.40	3.82

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is the respondent's log years of education. Standard errors in parentheses are clustered at the town level. All regressions restrict the analysis to regions at least 100 km close to missions. Controls are described in the text.

Table C.14: Missionary investments and medical prenatal care

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.0121** (0.0055)		
Distance to any hospital	-0.0150** (0.0059)		
Distance to Protestant		-0.0148* (0.0080)	
Distance to Protestant health		0.0222 (0.0136)	
Distance to Catholic			-0.0136** (0.0066)
Distance to Catholic health			-0.0207*** (0.0069)
Observations	111,454	73,264	93,881
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	8,626	6,000	7,404
R-sq	0.27	0.24	0.26
Mean DepVar	0.759	0.790	0.755
Sd DepVar	0.428	0.407	0.430

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual (women only). The dependent variable is a binary variable equal to one if the respondent (mothers only) received prenatal care from a medically trained professional (nurse or doctor). Standard errors in parentheses are clustered at the town level. All regressions restrict the analysis to regions at least 100 km close to missions. Controls are described in the text.

Table C.15: Missionary investments and severe anemia (women)

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.0004* (0.0002)		
Distance to any hospital	0.0001 (0.0003)		
Distance to Protestant		-0.0002 (0.0003)	
Distance to Protestant health		-0.0002 (0.0004)	
Distance to Catholic			-0.0001 (0.0002)
Distance to Catholic health			0.0001 (0.0003)
Observations	234,187	150,851	202,819
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	8,927	6,009	7,703
R-sq	0.00	0.00	0.00
Mean DepVar	0.003	0.003	0.004
Sd DepVar	0.058	0.057	0.060

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual (women only). The dependent variable is a binary variable equal to one if the respondent is anemic. Standard errors in parentheses are clustered at the town level. All regressions restrict the analysis to regions at least 100 km close to missions. Controls are described in the text.

Table C.16: Missionary investments and anemia prevalence (women)

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	0.0003 (0.0009)		
Distance to any hospital	0.0005 (0.0009)		
Distance to Protestant		-0.0016 (0.0013)	
Distance to Protestant health		0.0014 (0.0016)	
Distance to Catholic			0.0027*** (0.0010)
Distance to Catholic health			-0.0004 (0.0010)
Observations	229,505	148,189	198,546
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	8,926	6,009	7,702
R-sq	0.02	0.02	0.02
Mean DepVar	0.040	0.042	0.040
Sd DepVar	0.195	0.200	0.196

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual (women only). The dependent variable is a binary variable equal to one if the respondent is anemic. Standard errors in parentheses are clustered at the town level. All regressions restrict the analysis to regions at least 100 km close to missions. Controls are described in the text.

Table C.17: Missionary investments and stunted growth (women)

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	0.0004 (0.0005)		
Distance to any hospital	-0.0001 (0.0004)		
Distance to Protestant		0.0012 (0.0008)	
Distance to Protestant health		-0.0006 (0.0010)	
Distance to Catholic			0.0004 (0.0006)
Distance to Catholic health			-0.0002 (0.0005)
Observations	234,187	150,851	202,819
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	8,927	6,009	7,703
R-sq	0.02	0.02	0.01
Mean DepVar	0.008	0.008	0.008
Sd DepVar	0.090	0.088	0.088

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual (women only). The dependent variable is a binary variable equal to one if the respondent has stunted growth. Standard errors in parentheses are clustered at the town level. All regressions restrict the analysis to regions at least 100 km close to missions. Controls are described in the text.

Table C.18: Missionary investments and HIV : the DRC and the rest

	Non Central Africa			Central Africa		
	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se
Distance to any mission	-0.004** (0.002)			-0.002 (0.002)		
Distance to any hospital	0.005** (0.002)			0.000 (0.003)		
Distance to Protestant		-0.016*** (0.003)			-0.001 (0.006)	
Distance to Protestant health		0.010*** (0.003)			0.008 (0.009)	
Distance to Catholic			-0.000 (0.002)			-0.001 (0.003)
Distance to Catholic health			0.011*** (0.002)			-0.001 (0.003)
Observations	131,596	71,023	114,850	10,139	5,245	9,744
Country and Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	5,069	2,868	4,376	314	155	302
R-sq	0.10	0.10	0.10	0.01	0.02	0.01
Mean DepVar	0.07	0.09	0.06	0.01	0.01	0.01
Sd DepVar	0.25	0.28	0.24	0.11	0.11	0.12

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is a binary variable equal to one if the respondent is HIV positive. Standard errors in parentheses are clustered at the town level. All regressions restrict the analysis to regions at least 100 km close to missions. Controls are described in the text. Columns (1)-(3) exclude the DRC. Columns (4)-(6) restrict the analysis to the DRC.

Table C.19: Missionary investments and urbanization

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	19.013*** (2.106)		
Distance to any hospital	18.881*** (2.233)		
Distance to Protestant		14.538*** (2.722)	
Distance to Protestant health		15.592*** (2.785)	
Distance to Catholic			37.644*** (2.752)
Distance to Catholic health			-6.243 (3.987)
Observations	7,232	4,456	4,456
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
R-sq	0.45	0.57	0.57
Mean DepVar	182.74	177.77	177.77
Sd DepVar	158.47	151.93	151.93

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the town. The dependent variable is the average travel time in minutes to a urban center. Standard errors are in parentheses Controls are described in the text.

Table C.20: Missionary investments and HIV in urbanized areas

	Low Travel time		
	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.009*		
	(0.005)		
Distance to any hospital	0.012***		
	(0.004)		
Distance to Protestant		-0.013***	
		(0.005)	
Distance to Protestant health		0.017**	
		(0.007)	
Distance to Catholic			-0.009**
			(0.005)
Distance to Catholic health			0.008
			(0.005)
Observations	21,485	21,485	21,485
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	739	739	739
R-sq	0.11	0.11	0.11
Mean DepVar	0.10	0.10	0.10
Sd DepVar	0.30	0.30	0.30

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is a binary variable equal to one if the respondent is HIV positive. Standard errors in parentheses are clustered at the town level. Controls are described in the text. The regressions restrict the analysis to the 10% best connected areas to a urban center (with the lowest travel time to a urban center). This corresponds to regions within a 2 km (urban) or 10 km (rural) buffer for which the estimated travel time to a settlement of 50,000 or more is less than 8 minutes.

Table C.21: Missionary investments and HIV Prevalence: No contemporary town-level controls

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.010*** (0.001)		
Distance to any health	0.007*** (0.001)		
Distance to Protestant		-0.015*** (0.001)	
Distance to Protestant health		0.009*** (0.002)	
Distance to Catholic			-0.003*** (0.001)
Distance to Catholic health			0.007*** (0.001)
Observations	259,488	138,173	228,923
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	10,566	6,023	9,289
R-sq	0.08	0.07	0.08
Mean DepVar	0.08	0.12	0.07
Sd DepVar	0.27	0.32	0.26

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is a binary variable equal to one if the individual is HIV positive. Standard errors in parentheses are clustered at the town level. Controls are described in the text. All regressions restrict the analysis to towns within a 100 km buffer of a mission. regions at least 100 km close to missions.

C.3 Channels

Table C.22: Missionary investments and consumption of sex worker services in the year (men)

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.0020 (0.0016)		
Distance to any hospital	0.0096*** (0.0018)		
Distance to Protestant		-0.0026 (0.0026)	
Distance to Protestant health		0.0012 (0.0034)	
Distance to Catholic			-0.0013 (0.0020)
Distance to Catholic health			0.0130*** (0.0021)
Observations	68,085	43,807	59,178
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	6,702	4,581	5,734
R-sq	0.02	0.02	0.02
Mean DepVar	0.045	0.049	0.043
Sd DepVar	0.207	0.216	0.204

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the male individual. . The dependent variable is a binary variable equal to one if the individual reports having paid for sex from sex workers in the year. Standard errors in parentheses are clustered at the town level. Controls are described in the text. All regressions restrict the analysis to towns within a 100 km buffer of a mission.

Table C.23: Missionary investments and consumption of sex worker services in life (men)

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.0006 (0.0032)		
Distance to any hospital	0.0261*** (0.0044)		
Distance to Protestant		-0.0099** (0.0048)	
Distance to Protestant health		0.0201*** (0.0067)	
Distance to Catholic			0.0007 (0.0038)
Distance to Catholic health			0.0304*** (0.0048)
Observations	58,676	33,840	52,694
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	5,758	3,402	5,170
R-sq	0.08	0.07	0.09
Mean DepVar	0.136	0.158	0.134
Sd DepVar	0.342	0.365	0.340

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the male individual. The dependent variable is a binary variable equal to one if the individual reports having paid for sex from sex workers ever in his life. Standard errors in parentheses are clustered at the town level. Controls are described in the text. All regressions restrict the analysis to towns within a 100 km buffer of a mission.

Table C.24: Missionary investments and monogamy

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.0175*** (0.0024)		
Distance to any hospital	-0.0048* (0.0028)		
Distance to Protestant		-0.0190*** (0.0032)	
Distance to Protestant health		0.0149*** (0.0037)	
Distance to Catholic			-0.0260*** (0.0031)
Distance to Catholic health			0.0059* (0.0036)
Observations	187,279	115,169	160,300
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	8,898	5,980	7,674
R-sq	0.15	0.15	0.16
Mean DepVar	0.827	0.829	0.828
Sd DepVar	0.379	0.377	0.377

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is a binary variable equal to one if the individual reports being in a monogamous relationship. Standard errors in parentheses are clustered at the town level. Controls are described in the text. All regressions restrict the analysis to towns within a 100 km buffer of a mission.

Table C.25: Missionary investments and number of sexual partners in the year, other than spouse

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.0200*** (0.0030)		
Distance to any hospital	0.0149*** (0.0036)		
Distance to Protestant		-0.0183*** (0.0044)	
Distance to Protestant health		0.0125*** (0.0044)	
Distance to Catholic			-0.0164*** (0.0042)
Distance to Catholic health			0.0189*** (0.0046)
Observations	331,838	214,139	286,439
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	8,779	5,936	7,563
R-sq	0.07	0.06	0.07
Mean DepVar	0.232	0.283	0.233
Sd DepVar	0.737	0.797	0.749

Notes: * p<0.10, ** p<0.05, *** p<0.01. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is the declared number of sexual partners in the year, other than the spouse. Standard errors in parentheses are clustered at the town level. Controls are described in the text. All regressions restrict the analysis to towns within a 100 km buffer of a mission.

Table C.26: Missionary investments and number of sexual partners during lifetime

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.1154*** (0.0232)		
Distance to any hospital	0.2158*** (0.0245)		
Distance to Protestant		-0.1505*** (0.0320)	
Distance to Protestant health		0.1961*** (0.0329)	
Distance to Catholic			-0.1448*** (0.0318)
Distance to Catholic health			0.3009*** (0.0340)
Observations	229,127	155,788	192,246
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	7,347	5,134	6,179
R-sq	0.16	0.16	0.16
Mean DepVar	3.092	3.258	3.134
Sd DepVar	4.763	4.821	4.883

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is the declared number of sexual partners during lifetimes. Standard errors in parentheses are clustered at the town level. Controls are described in the text. All regressions restrict the analysis to towns within a 100 km buffer of a mission.

Table C.27: Missionary investments and number of sexual partners during lifetime, per year sexually active

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.0140*** (0.0033)		
Distance to any hospital	0.0092** (0.0038)		
Distance to Protestant		-0.0164*** (0.0039)	
Distance to Protestant health		0.0077* (0.0042)	
Distance to Catholic			-0.0114** (0.0050)
Distance to Catholic health			0.0223*** (0.0054)
Observations	176,504	131,384	145,358
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	7,323	5,133	6,155
R-sq	0.23	0.25	0.23
Mean DepVar	0.453	0.455	0.462
Sd DepVar	0.735	0.716	0.748

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the individual. The dependent variable is the declared number of sexual partners during lifetimes, per year sexually active. Standard errors in parentheses are clustered at the town level. Controls are described in the text. All regressions restrict the analysis to towns within a 100 km buffer of a mission.

Table C.28: Missionary presence and Christianity

	(1)	(2)	(3)
	b/se	b/se	b/se
Distance to any mission	-0.0150*** (0.0048)		
Distance to any hospital	-0.0104** (0.0050)		
Distance to Protestant		-0.0207*** (0.0062)	
Distance to Protestant health		-0.0126* (0.0069)	
Distance to Catholic			-0.0223*** (0.0056)
Distance to Catholic health			0.0182*** (0.0051)
Observations	320,230	198,958	274,701
Country and Wave FE	Yes	Yes	Yes
Contemporary Controls	Yes	Yes	Yes
Historical and Geo controls	Yes	Yes	Yes
Clusters	8,173	5,347	6,965
R-sq	0.55	0.49	0.58
Mean DepVar	0.661	0.702	0.653
Sd DepVar	0.473	0.457	0.476

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Table reports OLS estimates. The unit of observation is the town. The dependent variable is a binary variable equal to one if the respondent is Christian. Standard errors in parentheses are clustered at the closest mission level. Controls are described in the text. All regressions restrict the analysis to regions at least 100 km close to missions.

Table C.29: Descriptive statistics: HIV and behaviors among Christians and the rest of the population, urban areas only

	(1)	(2)	(3)
	Not Christian	Christian	Diff
HIV positive	0.032	0.086	-0.054*** (0.001)
Abstinent before married	0.656	0.588	0.068*** (0.001)
Age at marriage	19.413	20.628	-1.216*** (0.021)
Nb of sex partners in life	2.768	3.546	-0.778*** (0.018)
Sex worker	0.099	0.129	-0.029*** (0.002)
Sex worker (year)	0.031	0.042	-0.011*** (0.001)
Age gap bw woman and man	-10.572	-6.419	-4.153*** (0.031)
Knows condom lowers chances of AIDS	0.740	0.819	-0.080*** (0.001)
Knows a place where to get condoms	0.416	0.193	0.223*** (0.001)
Catholic religion	0.000	0.306	-0.306*** (0.001)
Rural Area	0.000	0.000	0.000 (0.000)
Years of Schooling	3.738	6.983	-3.245*** (0.012)
Female	0.658	0.666	-0.008*** (0.001)
Observations	195823	335858	531681

Notes: The Table reports descriptive statistics for the sample of individuals residing in urban areas. The unit of observation is the individual. Column (1) reports the statistics for non Christian individuals . Column (2) reports similar statistics for Christians. Column (3) describes all the sample

D Reverend Lattimer Fuller's letter

“The Usual medical estimate of the health of the million Natives in this diocese is that seventy five percent of them suffer more or less from syphilis, and not less than twenty five percent from incipient or advanced phthisis. Personally, I believe that the latter percentage is rapidly on the increase.

The syphilis here is largely inherited and very virulent because it is a new disease to the people and thus finds a suitable nidus. It appears constantly as an affection of the throat on the approach of puberty; at this stage it can be easily and successfully treated, only there is no one to treat it.

The phthisis is largely due to the fact that the people have practically no idea of ordinary rules of health and pay no attention to a cough until the disease has got firm hold. This also being a new disease attacks the people with great force.

In addition to the above two diseases the natives here are I think quite as susceptible to disease as the Europeans for they are rapidly losing their old savage robustness and have as yet no traditional care of health.

Here on the reef we have lately been thinking what method of medical work would be most likely to success and we think that nothing would be really of much use in such densely populated a district except something in the way of a cottage hospital where two or three native girls would be trained to look after the sick, and where cases could be treated as in or as out-patients and from where a nurse could go and visit along the reef.

Again our idea is that for ordinary districts a nurse could do very effective work, as people are there in reach of doctors. But further away from civilization, as for example in the Zoutpansberg (Pietsburg) there ought certainly to be a resident doctor with a hospital.

I only wish it were possible to contemplate such work being begun.

References

- Alsan, Marcella**, “The Effect of the TseTse Fly on African Development,” *American Economic Review*, jan 2015, *105* (1), 382–410.
- Company, Century**, *The Century Atlas: Africa [map]*, Buffalo, NY: Matthews-Northryp, 1911.
- 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.
- Mayala, Benjamin, Thomas D Fish, David Eitelberg, and Trinadh Dontamsetti**, “The DHS Program Geospatial Covariate Datasets Manual,” Technical Report, USAID 2018.
- Nunn, Nathan and Leonard Wantchekon**, “The Slave Trade and the Origins of Mistrust in {Africa},” *American Economic Review*, dec 2011, *101* (7), 3221–3252.
- Oster, Emily**, “Unobservable Selection and Coefficient Stability: Theory and Validation,” *Journal of Business Economics and Statistics*, 2016, *Forthcomin*.