Long-Term Effects of Access to Health Care:
Medical Missions in Colonial India

Rossella Calvi[§1] and Federico G. Mantovanelli[♣2]

[1]Department of Economics, Boston College

[2]Analysis Group

This Version: August, 2015

(First Draft: August, 2014)

Abstract

We study the long-term effect of access to health care on individuals' health status by investigating the relationship between the proximity to a Protestant medical mission in colonial India and current health outcomes. We use individuals' anthropometric indicators to measure health status and geocoding tools to calculate the distance between the location of individuals today and Protestant health facilities founded in the nineteenth century. We exploit variation in activities of missionary societies and use an instrumental variable approach to show that proximity to a Protestant medical mission has a causal effect on individuals' health status. We find that a 50 percent reduction in the distance from a historical medical facility increases current individuals' body mass index by 0.4. We investigate some potential transmission channels and we find that the long-run effect of access to health care is not driven by persistence of infrastructure, but by improvements in individuals' health potential and changes in hygiene and health habits.

Keywords: India, health, body mass index, historical persistence, Protestant missions.

JEL codes: I15, N35, O12, O15, Z12, Z13.

1 Introduction

According to the most recent World Health Organization (WHO) estimates, India is ranked 150th in the world in terms of life expectancy and one out of three Indian adults is underweight. These country-level figures hide a considerable regional variation in health conditions: while South Indian states are doing relatively well in terms of health performance, the same cannot be said for some of the most populous states, such as Uttar Pradesh and Bihar, whose health indicators are worse than those of many Sub-Saharan African countries.[1] Moreover, recent evidence suggests that differences in health performance persist even at the sub-state level and across areas within the same district.[2]

This paper studies one of the possible sources of this geographical heterogeneity in health conditions by examining the long-term effect of access to historical health facilities on current individual health outcomes. We exploit information about the diffusion, location and activities of Protestant medical missionaries in colonial India.[3] In the late nineteenth century, missionaries became aware of the role that medicine could have in facilitating evangelization among non-Christian populations and the number of doctors and medical personnel sent to India by Protestant societies increased dramatically.[4]

We investigate whether the proximity to a Protestant medical mission in colonial India predicts individual health outcomes today. To this aim, we construct a novel and unique dataset that combines contemporary individual-level data with historical information on Protestant medical missions.[5] Consistent with the previous literature, we measure health status with individuals' anthropometric indicators.[6] We use geocoding tools to calculate the distance between the location of individuals today and that of Protestant health facilities in nineteenth century India. We find that the proximity to a Protestant medical mission is positively and significantly associated with current individuals' body mass index (BMI) and other health indicators.

To give a causal interpretation to our estimates, we account for two types of selection bias. First, historical and geographic characteristics may have driven the missionaries' location decisions. Second, Protestant missions may have endogenously selected themselves into building

---

[1]See Pal and Ghosh (2007).

[2]See Srinivisan (2010). India comprises 28 states and 7 union territories. These are divided into 593 districts (Census 2001).

[3]Our decision of focusing on Protestant medical missions as opposed to Catholic missions is mainly data driven. To the best of our knowledge, there is no document that systematically records medical activities of Catholic missions. Moreover, India was the preferred destination of the Protestant missionary venture due to its British element. Protestant were especially committed to operate among non-elites and socially disadvantaged groups, while Catholic missionaries, at least until the Second Vatican Council in 1965, strategically decided to focus their efforts on the elite groups of society (Mathew (1999)).

[4]See Richter (1908) and Fitzgerald (2001).

[5]Throughout the paper we use the terms medical missions, medical facilities, medical settlements, health facilities, and hospitals, interchangeably.

[6]See Schultz (2002, 2003).

a health facility and into providing health care. We show that these potential biases are not the the driving force behind our results. To correct for possible systematic differences between regions with and without Protestant missions, we restrict our sample of analysis to individuals living within a certain distance from a Protestant mission. By limiting the analysis to more concentrated areas we minimize the risk of our findings being driven by unobserved geographical differences. In addition, we control for the minimum distance from generic Protestant and Catholic missions. These variables should capture factors that also determined the location decisions of Protestant medical missionaries. If our results were purely driven by unobserved heterogeneity across locations, we would find a significant correlation between these variables and individuals' current health outcomes. We address the second type of selection bias by developing an instrumental variable approach in the spirit of Cagé and Rueda (2014). We exploit the fact that Protestant missionary societies differ in their inclination to undertake medical activities. We use these differences to compute each society's share of medical missions in all regions of the world *outside* of India and then use those shares to construct our instrument. We show that our results are robust to the instrumental variable approach. We find that halving the distance from the location of a Protestant medical facility increases current individuals' BMI by 0.4.

We analyze three potential transmission channels: persistence of infrastructure, improvements in health potential (health status of previous generations and nutrition), and improvements in health habits (hygiene and health awareness). First, consistent with the previous literature, we show that proximity to current health facilities does not play a role.[7] Second, we find that the long-term effect of historical medical missions is partially, but not entirely, driven by improvements in health potential. Finally, we provide evidence of the fact that proximity to a health care facility in the past may have affected individual habits regarding hygiene, preventive care and health awareness, which have been bequeathed over time and have influenced health outcomes of later generations.

This paper represents the first attempt to investigate how the exposure to historical medical missions may have determined advancements in current individual health outcomes. With this contribution, we add to the growing empirical literature that looks at historical institutions as important determinants of current outcomes and investigates path persistence in developing countries.[8] More specifically, we contribute to previous literature that recognizes the importance of the activities of religious organizations as possible fundamental sources of economic and social outcomes.[9] Previous works have identified Protestantism as an important determinant of

---

[7]See e.g. Chaudhury and Hammer (2004) and Banerjee et al. (2004).

[8]See e.g. Acemoglu et al. (2001), Glaeser and Shleifer (2002), Nunn (2009), and Alesina et al. (2013).

[9]See e.g. Barro and McCleary (2003, 2005).

advancements in human capital by focusing on its contribution to literacy and education.[10]

Recent works study how religious institutions and affiliation may help explaining differences in health outcomes. Almond and Mazumder (2011) analyze births to Arab parents in Michigan and find that prenatal exposure to Ramadan results in lower birth weight. While Bhalotra et al. (2010) show that Muslim children in India have a significantly higher probability of survival in infancy than Hindu children, Brainerd and Menon (2013) demonstrate that this advantage does not persist beyond infancy. Finally, Menon and McQueeney (2015) find that Christian infants in India have higher health outcomes compared to infants of other religious identities. We differ from this strand of literature in that we focus on the long-term effect of religious institutions.

Important policy implications may be drawn from our analysis. On one hand, understanding historical determinants of regional variation in health conditions may provide useful guidance in addressing issues related to health inequality in modern India. On the other hand, in light of the existence of long-run positive spillovers, the expansion of healthcare access becomes an even more critical priority, as it provides benefits for both current and future generations. Health can influence not only individuals' early survival and life expectancy, but also cognitive performance, schooling decisions, and productivity as adults. Understanding accumulation of health capital in developing countries means understanding one of the foremost determinants of economic development.[11]

The rest of the paper is organized as follows. Section 2 provides a brief historical description of the Protestant medical enterprise in colonial India. Section 3 describes our novel dataset, and discusses the use of anthropometric indicators as a measure of health. Section 4 presents our empirical strategy. Section 5 discusses the estimation results, while section 6 checks the robustness of our findings. Section 7 investigates potential transmission channels. Finally, section 8 concludes.

2 Medical Missions in Colonial India

The period between 1800-1914 represents an era of great missionary expansion led by the Protestants of Britain and the United States. The intensity of the missionary evangelization was such that that period is known as the *great century* of missions. India was the preferred destination of the Protestant missionary venture due to its British element. In the mid-nineteenth century more than a quarter of all the Protestant missionaries were stationed in India. On the eve of the

---

[10]See e.g. Weber (1930), Becker and Woessmann (2009), Gallego and Woodberry (2009), Bai and Kung (2014), Cagé and Rueda (2014), and Mantovanelli (2014).

[11]See Schultz (2009).

first World War 5,200 Protestant missionaries (of whom 2,500 British and 1,800 American) were posted in the Indian sub-continent. At that time, only China rivaled India in terms of Protestant presence.[12]

At first, missionary societies showed very little interest in medicine. Protestant organizations did not consider the establishment of hospitals and the provision of medical care as crucial for the scope of their missions. The missionary movement was convinced that the biblical command of evangelizing every nation was limited only to preaching and teaching the Gospel.[13]

Evangelizing strategies started to change from the 1860s onwards in correspondence with the growing awareness of the deficiencies of the orthodox missionary methods. Anxiety over the lack of tangible results provided a powerful trigger for a change in official mission policies and practices. The belief was that the missionary sphere could rightfully include a range of activities beyond those of preaching and teaching:

> *"I feel convinced that medical missions are amongst the means best calculated [...] to give a fresh impetus to the cause of the Gospel, to help lift the chariot out of the rut in which sometimes it seems for a moment set fast and its progress retarded."*
> (Rev. Thomas V. French, Bishop of Lahore, quoted by Bishop McDougal in a speech at the Oxford Missionary Conference of 1877)

Medical missions were finally recognized as valuable auxiliaries to the work of propagating Christianity,

> *"The work of the doctor is to open the door that the evangelist may enter in."*
> (Medical Missionary Journal [1874], 9, 58-59)

In contrast to official colonial medicine, Protestant mission medicine sought to place itself within non-European social and institutional milieus and reached out to embrace all classes of the *native* society, regardless of caste or gender.[14]

In light of the recognition of the importance of medicine as an effective weapon in the evangelization process, the last decades of the nineteenth century witnessed an unprecedented growth of the medical mission enterprise. The number of doctors and medical agents sent to India by the Protestant societies rose from 28 in 1870 to 335 in 1912. This was accompanied by a rapid increase in the number of mission hospitals and dispensaries. By 1912, approximately three million patients were treated in Protestant medical facilities annually. The quality of medical missionary medicine increased as well in the closing years of the nineteenth century. At the Bombay Missionary Conference of 1892-93 a resolution was passed stating that medical missionaries in India

---

[12]For more details refer to Richter (1908).

[13]The constitution of the Baptist Missionary Society, for example, stated that the propagation of Christianity consisted of "the preaching of the Gospel, the translation of the Holy Scriptures and the establishment of Schools."

[14]See Hardiman (2008).

should invariably possess a medical degree or diploma sufficient to qualify as a licensed medical practitioner in the West. Fitzgerald (2001) indicates that, despite initial fears, it was surgery that determined acceptance of Western medicine among the Indian people and their greater willingness to use the resources of medical missions. Missionary reports document that over time patients became increasingly willing to enter a mission hospital for periods of in-patient care and were generally applying for medical assistance at an earlier stage in their illness.[15] In addition to foreign staff, medical missions trained and employed Indians working as assistants (around 700 in 1912), creating another vital contribution to the local communities.

3 Data

3.1 Data Sources

One of the main contributions of this paper is the creation of a unique and fully geocoded dataset that combines historical, geographic, and contemporary information.

Contemporary Data. Current individual data are from the *2003 India World Health Survey* (WHS). The World Health Surveys (WHS) were launched by the World Health Organization to monitor and evaluate critical health outcomes and health systems through cross-country comparable household surveys. The WHS were implemented between 2002 and 2004 in 70 countries selected to represent all regions of the world. The 2003 India WHS covers individuals and households residing in seven states: Assam, Karnataka, Madhya Pradesh, Maharashtra, Rajasthan, Uttar Pradesh and West Bengal. While the survey is centered around health-specific issues, it also provides demographic information of the respondents and data about household composition and educational level. Crucially for the purpose of our research, the survey discloses the geographic coordinates of where the respondents live and some anthropometric indicators.[16] Figure 1 shows the locations of survey respondents.

To investigate possible transmission channels of the long-term effect of access to Protestant medical missions, we use data from the Village Directory of the 2001 Census of India and the 2007-2008 District Level Household & Facility Survey (DLHS-3). We obtain information about the location of current health facilities from the former, and data about health and preventive care practices from the latter. The 2001 Village Directory contains detailed information about population and infrastructure/amenities - including hospitals and several types of health facilities

---

[15]See Fitzgerald (2001).

[16]The decision of not using the Demographic Health Survey, known as National Family Health Survey for India, is data driven, as geographic coordinates are not available for this country.

- of all villages and towns in India. The DLHS-3, one of the largest demographic and health surveys ever implemented in India, allows us to assess the utilization of services provided by health care facilities and people's health practices.[17] The survey does not provide any village level identifier, but it contains information about the village population as per Census 2001. We follow Banerjee and Sachdeva (2014) and use the village population to match the DLHS-3 data set with the Census 2001 data. The matching provides information on census village codes and villages names for all of the DLHS-3 sample.[18] We successfully match and geocode 21,151 villages.[19] To the best of our knowledge, this paper is the first one exploiting DLHS-3 information to construct a fully geocoded version of the survey.[20]

Historical Data. Information about the location of Protestant missions in colonial India comes from the *Statistical Atlas of Christian Missions*, published in 1910 in correspondence with the World Missionary Conference held in Edinburgh (Scotland). We geocode the maps of India contained in the Atlas to identify the exact location of Protestant missions as of 1908. Figure A1 in the Appendix shows an example of these maps. To identify missionary medical infrastructures we rely on the *Centennial Survey of Foreign Missions* (Dennis (1902)) which provides a worldwide list of all the Protestant missions with medical facilities and a record of their basic characteristics. For example, we know the designation of the facility (hospital or dispensary), the foundation date, the name of the supporting Protestant society, and the number of patients and surgical cases. A reproduction of one page of the survey is provided in Figure A2 in the Appendix. We digitize and geocode this information for the Indian subcontinent. We use the same data source to identify Protestant missions with educational institutions. Our final sample includes 1,069 Protestant missions. Of these, 183 are equipped with either a hospital or dispensary. Figure 1 shows the locations of the Protestant medical missions.

Beginning in the middle of the nineteenth century, the Government of India started a great transportation infrastructure project aimed at developing a railroad network in the Indian subcontinent. Railroads likely played a significant role in determining the missionaries' location decisions. Areas with a well-developed railways network were more likely to be targeted by missionaries as they were more easily accessible. To take this into account we digitize and geocode historical maps from the *Constable's Hand Atlas of India*. Figure A3 in the Appendix

---

[17]It consists of one village level, one household level, and two individual level questionnaires (ever married women of age 15-49 and unmarried women of age 15-24). While it covers all India and the sample size is much larger than in the WHS, the DLHS-3 does not contain anthropometric indicators.

[18]The matching is done generating village identifiers using state, district, and village population. We drop duplicates from the census data and all DLHS villages that have the same population at the tehsil/taluk (sub-district administrative units) level.

[19]The geocoding is based on Google coordinates for a string containing the names of *state*, *district*, and *village*.

[20]Banerjee and Sachdeva (2014) match the DLHS-3 data set with the Census 2001 data to analyze the impact of a nationwide road construction program on the usage, provision, and awareness of preventive health care. Their dataset does not include geographic coordinates. We are grateful to Ashish Sachdeva for suggesting this approach.
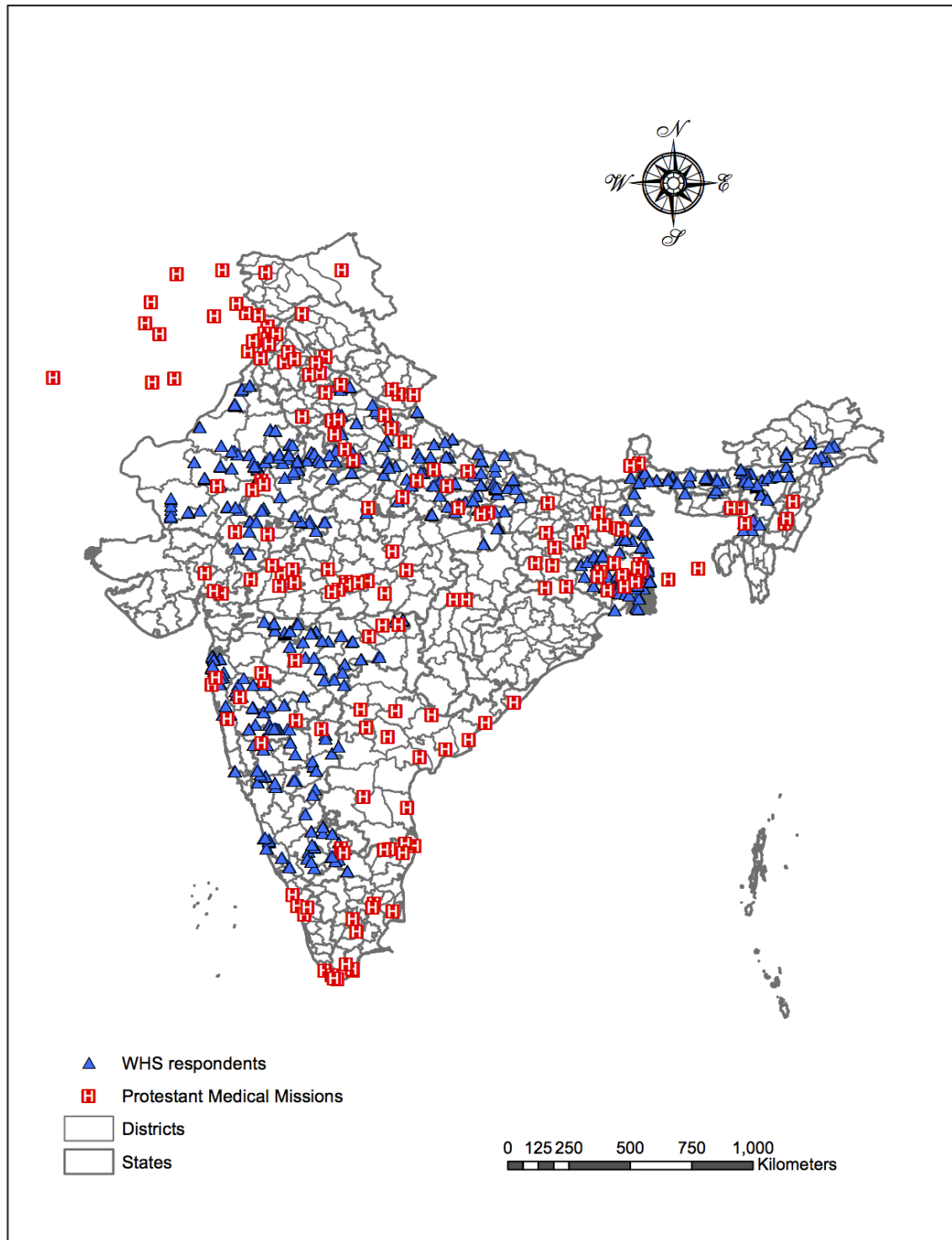
Figure 1: Protestant Medical Missions and WHS Respondents

reports a section of one of those maps showing the status of the Indian railways network as of 1891.

In our empirical analysis we also include information about the location of historical Catholic missions. We rely on maps contained in the *Atlas Hierarchicus*, an official document of the Vatican published in 1913 reporting the worldwide geographic distribution of Catholic missions as of 1911. Figure A4 in the Appendix shows one of these maps.

Finally, we obtain additional historical information from the History Database of the Global Environment - HYDE 3.1.[21] This database provides rasters of gridded total (estimated) historical population (inhabitants/gridcell) and gridded cropland (km^2/gridcell) from 10,000 B.C. to 2005

---

[21]See Klein Goldewijk et al. (2010, 2011).

A.D. at a 5×5 minutes resolution (about 85 km^2 around the equator). We extract the 1900 estimates for an area within a 5-kilometer radius from each WHS location.

Geographical Data. We control for geographic characteristics at the WHS location level. We use the *Global Gridded Population Database* from CIESIN to compute the population in year 2000 in an area within a 5-kilometer radius from each WHS location.[22] We also calculate the average elevation within the 5-kilometer buffer using the *SRTM 90m Digital Elevation Data* compiled by the CGIAR Consortium for Spatial Information.[23] Both population and elevation data are available at a 30×30 seconds resolution (about 0.8 km^2 around the equator). To account for the importance of access to water sources, we also include a variable measuring the number of water streams within the 5-kilometer buffer around the WHS locations. Data are from the *Digital Chart of the World*.[24]

3.2 Measuring Health

Measurement of health has recently evolved to rely on anthropometric indicators of physical development. Previous literature has used height, weight-for-height or body mass index to measure individual health.[25] In this paper we use body mass index (BMI) as our main measure of current individual health status.[26] We also consider height and a self-reported description of current health status as alternative measures of health.

That the relationship between health status and BMI is nonlinear is a well known fact, both at the individual and at the aggregate level. Waaler (1984) clearly illustrates this non-linearity by showing a U-shaped relationship between relative risk of mortality and BMI using Norwegian data from the 1970s. Especially in low income countries, however, an increase in the caloric intake shifts the lower tail of the BMI distribution to the right. Fogel (1994, 2004) interprets this shift as an accumulation of the population's health human capital, which tends to be associated with both declines in mortality and increases in labor productivity.

Figure 2 shows the results of a non-parametric estimation of the relationship between health status and BMI. Panel A suggests the presence of a non-monotonic concave relationship between

---

[22]http://www.ciesin.org/.

[23]http://srtm.csi.cgiar.org/.

[24]http://worldmap.harvard.edu/data/geonode:Digital_Chart_of_the_World.

[25]See e.g. Waaler (1984), Fogel (1994), Schultz (2002), Schultz (2003), Weil (2005), Steckel (2008), Deaton (2008), and Schultz (2009).

[26]The BMI is defined as $\frac{weight(kg)}{height^2(m)} = \frac{weight(lb)}{height^2(in)} \times 703$. The WHO provides an international classification of adult underweight, overweight and obesity according to BMI and fixes 18.5 as the cut-off value between underweight and normal-range, 25 as the cut-off value between normal-range and overweight and 30 as the cut-off value between overweight and obese. A WHO expert consultation in 2004 (Shiwaku et al. (2004)) addresses the recent debate about interpretation of recommended body-mass index (BMI) cut-off points for determining overweight and obesity in Asian populations, and develops population-specific cut-off points for BMI. The current BMI cut-off values for Asian Indians are 18.5, 23, 25 and 30 for the thresholds between underweight and normal-range, between normal-range and overweight, between overweight and pre-obese and between pre-obese and obese, respectively.
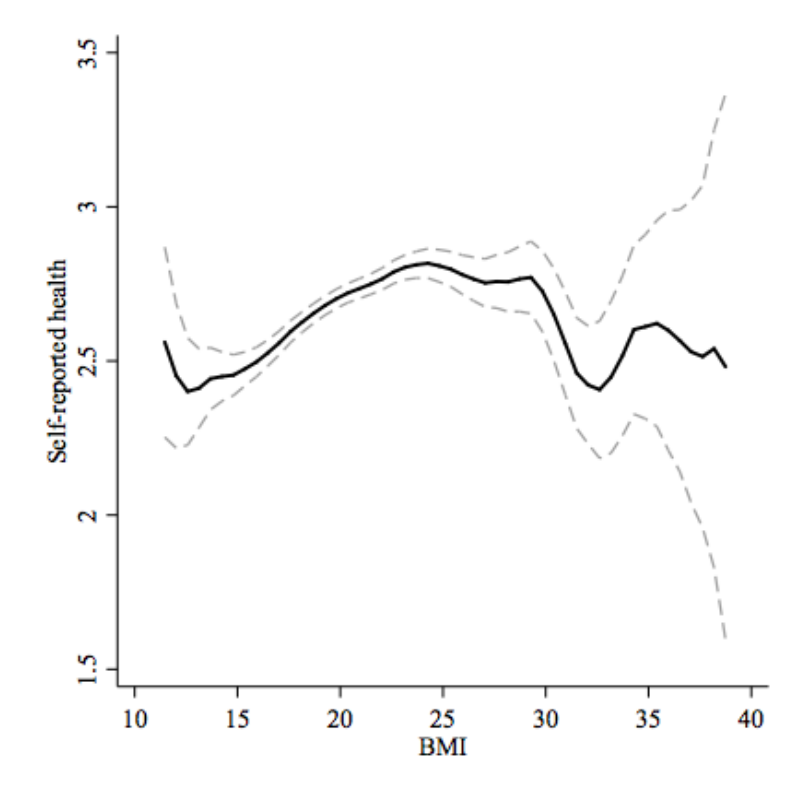
(A) Self-reported health status
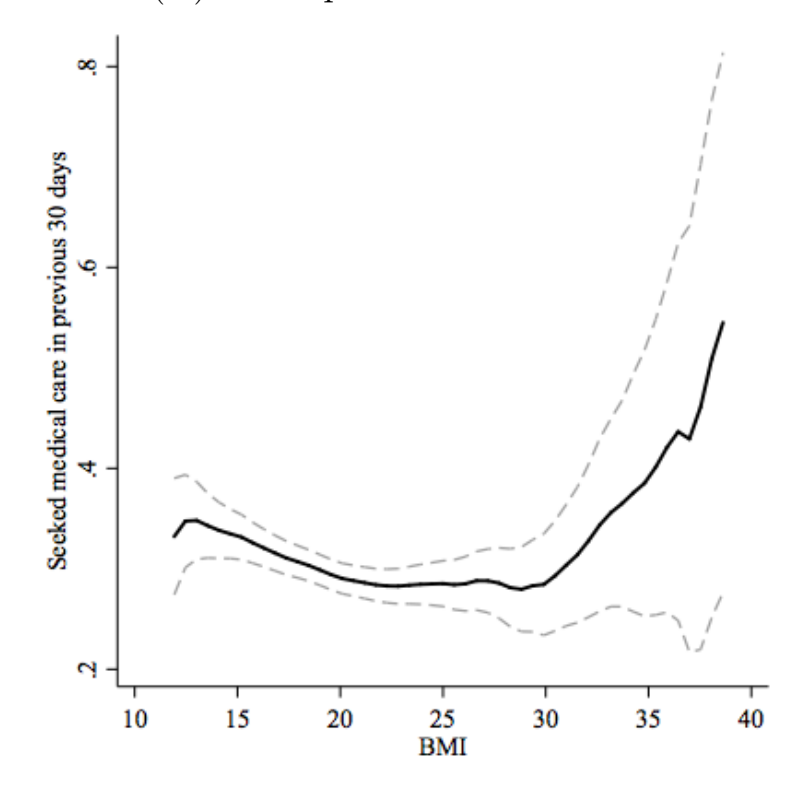
(B) Health care needed in the past month

Figure 2: Non-parametric regression of health status on BMI

respondents' self-reported health status and BMI, while panel B indicates a convex relationship between the probability of seeking health care in the previous month and BMI.

In our main empirical analysis we focus on individuals aged 20 to 60 with BMI between 15, the cutoff for starvation, and 30, the obesity threshold. We therefore concentrate only on the monotonic part of the BMI-health relationship. We exclude children, teenagers and elderly in order to avoid changes in individuals' BMI due to biological growth and aging. Figure 3 shows

the distribution of BMI in the main estimation sample.[27] Consistent with the previous literature, BMI is log normally distributed.[28]
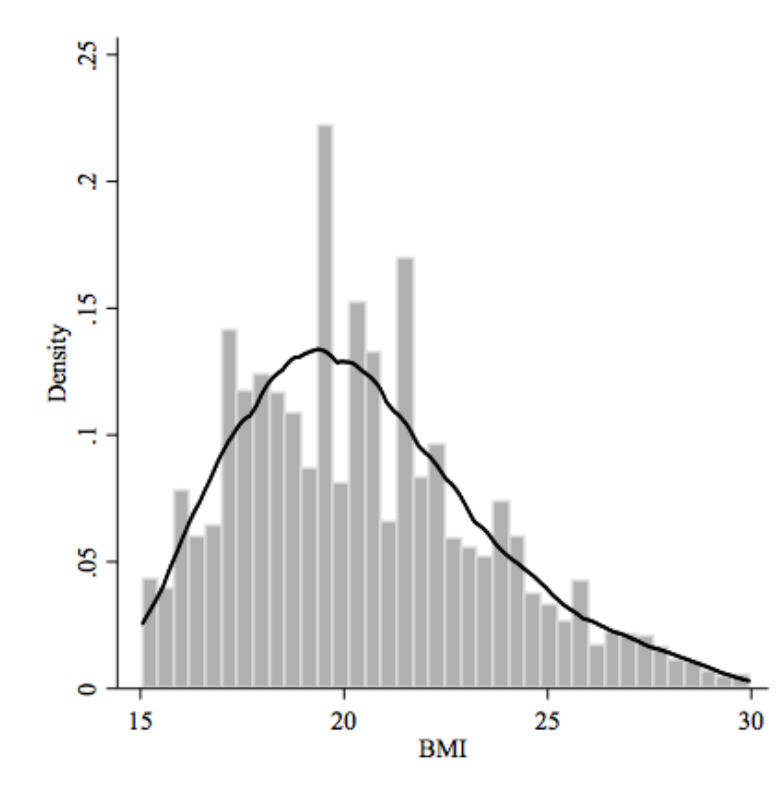

Figure 3: Distribution of BMI in the estimation sample

3.3 Descriptive Statistics

Table 1 contains some descriptive statistics for the variables used in our empirical analysis. Columns 1 to 3 present figures for the entire population in the survey, while columns 4 to 6 show figures for the baseline estimation sample (individuals of age 20 to 60, with BMI between 15 and 30). As there are no significant differences between the survey sample and our baseline estimation sample, we discuss here some descriptive statistics based on the latter.

The average BMI is 20.42, which falls in the normal range according to the WHO cutoffs, and the average height is 159 cm. Sixty-one percent of individuals rate their own health *good* or *very good*, while only 13 percent rate their health *bad* or *very bad*.[29] About half of the sample consists of females and individuals who completed primary school; the average age is 37 years and households have on average six members. The wealth quartiles are based on a wealth index computed as the first principal component that combines information about a set of household assets.[30]

The mean distance from a Protestant medical mission is 81 kilometers, and the mean distance

---

[27]Bandwidth: 0.75; Epanechnikov kernel function.

[28]See Burmaster and Crouch (1997) and Hjelmborg et al. (2008).

[29]Self-reported health status is an ordinal variable (from 0, "very bad", to 4, "very good") determined by the response to the survey question "How would you rate your health today?".

[30]Number of rooms in dwelling, refrigerator, scooter, radio, and bike.

from a generic Protestant mission - with or without a medical facility - is equal to 37 kilometers. On average, respondents live 65 kilometers away from a Catholic mission, while the distance from the location of Protestant missions with an educational institution is larger. Data on population in year 2000 are available at at the 30" × 30" grid cell. For year 1900, they are available at the 5' × 5' resolution, which explains why the average population in 1900 is larger than that in year 2000.

The location of a Protestant mission is likely not random. A number of geographic and historical factors may have driven the settlement decisions of missionaries in colonial India, such as access to clean water, population density, altitude, accessibility by railways, land crop suitability, and the proximity to a Catholic mission. Previous literature shows that Protestant missionaries choose to locate in geographically favorable and more accessible regions.[31]

In Table 2, we investigate how geographic and historical controls vary as we restrict the sample to areas closer to a Protestant mission or to a Protestant medical mission. Columns 1-2 (3-4) (5-6) report means and standard deviations of historical and geographic controls when the sample is restricted to observations below the 25th (50th) (75th) percentile of minimum distance from a mission (Panel A) or from a medical mission (Panel B).[32] Historical population in 5-kilometer radius increases as we restrict the sample to nearby areas, suggesting that missionaries settled in more densely populated areas. Moreover, areas close to missionary settlements have better access to colonial railways and water sources. The distance from a Catholic mission increases as farther away areas are included in the sample. On one hand, this suggests that the factors driving location decisions of Protestant missionaries are similar to those driving location decisions of Catholic missionaries. On the other hand, it indicates that the presence of a Catholic mission did not act as a deterrent to the settlement of Protestant missionaries.

---

[31] See e.g. Johnson (1967), Nunn (2010), and Cagé and Rueda (2014).

[32] The mission distance percentiles are 14 kilometers (25th), 29 kilometers (50th) and 49 kilometers (75th); the medical mission distance percentiles are 28 kilometers (25th), 75 kilometers (50th) and 121 kilometers (75th).

Table 1: Descriptive Statistics

	All sample			Estimation Sample		
	Obs. (1)	Mean (2)	St. Dev. (3)	Obs. (4)	Mean (5)	St. Dev. (6)
Body Mass Index (BMI)	9305	20.59	6.16	7046	20.42	3.02
Height (cm)	9305	158.73	11.35	7046	159.08	10.10
$\mathbb{1}$(Self-reported health = Very bad)	9566	0.02	0.14	7046	0.01	0.12
$\mathbb{1}$(Self-reported health = Bad)	9566	0.13	0.34	7046	0.12	0.32
$\mathbb{1}$(Self-reported health = Moderate)	9566	0.26	0.44	7046	0.26	0.44
$\mathbb{1}$(Self-reported health = Good)	9566	0.36	0.48	7046	0.38	0.49
$\mathbb{1}$(Self-reported health = Very good)	9566	0.22	0.41	7046	0.23	0.42
$\mathbb{1}$(Female)	9627	0.51	0.50	7046	0.52	0.50
$\mathbb{1}$(Married)	9624	0.77	0.42	7046	0.82	0.38
$\mathbb{1}$(Major ethnic group)	9625	0.81	0.39	7046	0.81	0.39
$\mathbb{1}$(Primary school completed)	9627	0.52	0.50	7046	0.53	0.50
$\mathbb{1}$(Wealth 1st quartile)	9469	0.29	0.45	7046	0.29	0.45
$\mathbb{1}$(Wealth 2nd quartile)	9469	0.22	0.41	7046	0.22	0.41
$\mathbb{1}$(Wealth 3rd quartile)	9469	0.26	0.44	7046	0.25	0.44
$\mathbb{1}$(Wealth 4th quartile)	9469	0.24	0.43	7046	0.24	0.42
Age	9626	38.85	15.22	7046	36.68	11.19
Household size	9619	5.99	2.42	7046	5.99	2.36
Population in 2000 (000s)	9181	5815.12	12851.37	7046	5745.19	12657.62
Altitude (meters)	9181	264.72	250.02	7046	262.04	247.97
No. rivers/water sources	9181	1.16	1.57	7046	1.17	1.59
No. colonial railways	9181	0.19	0.43	7046	0.18	0.43
Population in 1900 (000s)	9203	33523.47	83538.05	7046	33288.43	82041.03
Cropland in 1900 (km^2)	9203	38.48	33.15	7046	38.14	33.30
Latitude	9181	22.85	4.81	7046	22.99	4.77
Longitude	9181	80.01	6.43	7046	80.05	6.48
Distance from Protestant medical mission (km)	9181	81.37	59.58	7046	81.49	59.61
Distance from Protestant mission (km)	9203	36.51	34.06	7046	36.75	34.12
Distance from Catholic mission (km)	9203	64.13	52.86	7046	64.65	53.18
Distance from Protestant mission with university (km)	9203	225.09	177.98	7046	227.08	180.96
Distance from Protestant mission with boarding school (km)	9203	113.65	103.04	7046	114.98	104.98
Distance from Protestant mission with high school (km)	9203	156.36	173.67	7046	159.48	176.93
Distance from allopathic hospital from Census 2001 (km)	9203	20.64	25.42	7046	20.55	25.12

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. The wealth quartiles are based on a wealth index computed as the first principal component that combines information about a set of household assets: number of rooms in dwelling, and asset ownership, such as refrigerator, scooter, radio, and bike. Population, altitude, presence of rivers/water sources, access to colonial railways and cropland data are extracted in 5 km-radius buffers around the WHS locations. Population in year 2000 and elevation data are available at the 30" × 30" grid cell; original population in year 1900 and cropland data are available at the 5' × 5' resolution. Latitude and longitude are in decimal degrees.

Table 2: Geographical and Historical Controls

	Sample Restricted (Below Percentile of Distance)					
	25%		50%		75%	
	Mean (1)	St. Dev (2)	Mean (3)	St. Dev (4)	Mean (5)	St. Dev (6)
Panel A: Distance from Protestant mission						
Population in 2000 (000s)	12.22	18.99	9.91	17.21	7.37	14.50
Altitude (meters)	269.46	317.21	252.50	275.35	258.54	266.86
No. rivers/water sources within 5 km	1.31	1.81	1.31	1.73	1.21	1.61
Latitude	21.10	4.73	21.94	4.91	22.29	4.91
Longitude	80.17	6.31	80.41	6.31	80.35	6.21
Population in 1900 (000s)	77.07	133.91	55.13	110.97	41.75	94.58
Cropland in 1900 (km^2)	39.84	34.23	41.20	33.22	41.73	33.50
No. colonial railways within 5 km	0.38	0.60	0.29	0.52	0.22	0.46
Distance from Catholic mission (km)	37.43	39.36	44.29	43.42	50.36	42.25
Distance from Protestant mission with university (km)	179.74	150.30	190.08	160.31	194.29	163.11
Distance from Protestant mission with boarding school (km)	79.73	83.96	90.57	88.82	94.29	89.08
Distance from Protestant mission with high school (km)	9203	156.36	173.67	7046	159.48	176.93
Panel B: Distance from Protestant medical mission						
Population in 2000 (000s)	16.83	21.40	10.31	16.98	7.42	14.54
Altitude (meters)	199.48	281.18	219.16	259.60	252.04	257.92
No. rivers/water sources within 5 km	1.39	1.88	1.35	1.70	1.24	1.66
Latitude	22.44	4.15	22.83	4.25	23.04	4.47
Longitude	79.99	6.07	80.39	5.87	80.24	6.13
Population in 1900 (000s)	98.60	145.72	56.70	112.08	41.49	94.81
Cropland in 1900 (km^2)	39.72	35.21	45.33	33.69	41.77	33.63
No. colonial railways within 5 km	0.45	0.62	0.27	0.52	0.22	0.47
Distance from Catholic mission (km)	30.30	30.70	45.14	37.37	55.09	40.34
Distance from Protestant mission with university (km)	110.55	117.00	139.15	121.84	187.69	151.23
Distance from Protestant mission with boarding school (km)	44.28	41.66	58.94	48.17	83.78	69.46
Distance from Protestant mission with high school (km)	62.44	78.74	87.60	89.20	123.43	136.51

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. The mission distance percentiles (based on all sample statistics) are 14 kilometers (25th), 29 kilometers (50th) and 49 kilometers (75th); the medical mission distance percentiles (based on all sample statistics) are 28 kilometers (25th), 75 kilometers (50th) and 121 kilometers (75th). Population, altitude, number of colonial railways and cropland data are extracted in 5 km buffers around the WHS locations. Original population in year 2000 and elevation data are available at the 30" × 30" grid cell; original population in year 1900 and cropland data are available at the 5' × 5' resolution. Latitude and longitude are in decimal degrees.

4 Empirical Strategy

Our baseline specification is as follows:

$$BMI_{ivd} = \alpha_d + \beta Hospital\ distance_{ivd} + X'_{ivd}\gamma_1 + W^{g\prime}_{vd}\gamma_2 + W^{h\prime}_{vd}\gamma_3 + \epsilon_{ivd} \tag{1}$$

where BMI_{ivd} is the body mass index of individual i, living in village v in district d.[33] *Hospital distance$_{ivd}$* is the logarithm of the minimum aerial distance between the Protestant medical mission and survey respondent i.[34] $\beta/100$ measures the change in BMI following a 1 percent change in the distance from a location of a Protestant medical mission.

We control for a large set of covariates influencing both current individuals' BMI and the location of a Protestant hospital at the end of the nineteenth century. X_{ivd} contains individual and household level characteristics such as gender, marital status, level of education, ethnic group, age, household size, and a measure of wealth. The vector W^g_{vd} includes village specific geographic controls, such as population density (2000 population living within a 5-kilometer radius), number of rivers or lakes and average altitude in a 5-kilometer radius, latitude and longitude. W^h_{vd} contains additional historical characteristics, such as estimated population and area covered by cropland in 1900, and access to colonial railways in 1892. We include district level fixed effects, α_d, to account for unobserved heterogeneity across districts.

In order to identify the causal effect of the proximity to a historical Protestant medical facility on current individuals' health outcomes, we address two types of selection bias. First, as discussed in section 3.3, historical and geographic characteristics may have driven the missionaries' decision of settling down in specific locations. Moreover, Protestant missions may have endogenously selected themselves into building a health facility and into providing health care.

We deal with the first potential selection issue in two ways. First, in order to correct for possible systematic differences between regions with and without Protestant missions, we restrict our sample of analysis to individuals living close to a Protestant mission. By limiting the analysis to more concentrated areas we attempt to minimize the risk of our findings being driven by (within district) unobserved differences. Second, we expand W^h_{vd} to include the minimum distance from a generic Protestant and Catholic missions. These variables, while potentially endogenous, should capture factors that also determine the location decisions of Protestant medical missionaries. If our results were purely driven by unobserved heterogeneity across locations, we would find a sig-

---

[33]The survey provides anthropometric information for one single individual per household, so the individual and household dimensions coincide in our dataset.

[34]We consider the logarithm of the distance, instead of the variable in levels, as we believe that percentage changes in distance from a Protestant medical mission matters more than level changes, i.e. we believe that an increase in distance from 10 km to 50 km does not have the same effect of an increase in distance from 110 km to 150 km.

nificant correlation between these variables and individuals' current health outcomes. Moreover, this approach allows us to check that our results are indeed driven by proximity to a Protestant medical mission and not by proximity to a generic mission.

We address the second potential source of selection bias by using an instrumental variable approach. The fact that Protestant missionaries built medical missions in historically and geographically favorable areas may bias the OLS estimates of β upwards. In contrast, potential downward bias of OLS estimates may arise due to measurement error, and to the fact that medical missions were built in areas where medical care was more needed.

Several associations of missionaries and evangelists started to emerge in the late eighteenth century with the purpose of *"administering funds, [...] sending out missionaries, initiating and conducting missionary operation, funding churches and institutions, and otherwise fulfilling the varied aims of mission effort"*.[35] To construct our instrument we follow Cagé and Rueda (2014) and exploit the heterogeneity in the missionary societies' propensity to undertake medical activities. For each of the 183 Protestant medical missions in our dataset, we observe the supporting missionary society. For each missionary society, we compute its share of medical missions *outside* India as:

$$Society\ hospital\ share = \frac{No.\ medical\ missions\ outside\ India}{No.\ missions\ outside\ India} \tag{2}$$

By construction, all medical missions supported by the same missionary society are assigned the same share.[36] *Society hospital share* can be interpreted as a measure of the propensity of the supporting missionary society to be involved in medical activities, and, therefore, as a measure of the likelihood to have a health facility in their missions. We then construct our instrument as the sum of these shares for all medical missions, $m(k)$, located $k \leq K$ kilometers away from individual i living in village v:[37]

$$Hospital\ share_{ivd}^{K} = \sum_{k \leq K} Society\ hospital\ share_{ivd}^{m(k)} \tag{3}$$

Figure A5 in the Appendix provides a graphical illustration of how we construct the instrument.[38]

---

[35]Dennis (1902).

[36]Table A1 in the Appendix contains information about the missionary societies we include in our main analysis. The table includes only societies affiliated with medical missions historically located less than 50 kilometers away from a WHS respondent. Data are from the *Centennial Survey of Foreign Missions*.

[37]Our findings do not change if we construct our instrumental variable as the sum of shares, defined as in equation (3), for all missions, both medical and non-medical. See section 6 for more details.

[38]The larger is the radius K, the weakly larger is the value of the instrument, by construction. We consider the sum of *Society hospital share* instead of the average, as we want to capture the fact that being close to missions affiliated with health oriented societies indeed is associated with a lower distance from a Protestant medical mission. Consider two individuals A and B, both located less than K kilometers away from a mission affiliated with the American Baptist Missionary Union, for which about 23 percent of missions outside of India are medical. If A and B only had one mission located less than K kilometers away, *Hospital share*$_{ivd}^{K}$ would equal 23 for both of them. Now suppose individual A, instead, was located less than K kilometers away from two missions, one supported by the American Baptist Missionary Union and the other by the Wesleyan Missionary Society, for which less than 3 percent of missions outside of India are medical. Considering the average of *Society hospital share* would assign a value of 12.2 to individual A, and therefore

15

The first stage of our IV approach is defined as:

$$Hospital\ distance_{ivd} = \alpha_d + \lambda Hospital\ share_{ivd}^K + X'_{ivd}\gamma_1 + W^{g\prime}_{vd}\gamma_2 + W^{h\prime}_{vd}\gamma_3 + \epsilon_{ivd} \qquad (4)$$

where $Hospital\ distance_{ivd}$ is the logarithm of the minimum distance between the Protestant medical mission and the location of survey respondent i and $Hospital\ share_{ivd}^K$ is our instrument as defined in (3).

Identification comes from the assumption that $Hospital\ share_{ivd}^K$ is uncorrelated with the error term in our outcome equation (1), while adequately correlated with $Hospital\ distance_{ivd}$. The larger *Society hospital share*, the more likely it is for a mission associated with this society to be equipped with a hospital. The larger $Hospital\ share_{ivd}^K$, the more likely it is for an individual to be closer to a medical mission. The exclusion restriction requires $Hospital\ share_{ivd}^K$ to affect current individuals health outcomes only through the proximity to a Protestant medical mission. More specifically, *Society hospital share* must be uncorrelated with any other long-term determinants of individuals' health. In section 6, we address concerns about potential violations of this assumption.

5 Results

5.1 Baseline Model

Table 3 presents the baseline OLS estimates. Column 1 shows that BMI and the distance from a Protestant medical mission are significantly and negatively related. The estimated coefficient is equal to -0.271, indicating that a 1 percent increase in the distance is associated with a decrease in individual's BMI by 0.00271. In other words, doubling the distance is associated with a decrease in BMI by 0.19.[39]

Estimates in columns 2 to 5 indicate that the negative correlation between proximity to a medical mission and health outcome today is robust to the inclusion of district fixed effects, individual characteristics and geographic and historical village level controls. While village level controls do not seem to play a role in determining individuals' health status - with altitude and access to colonial railways being the only exceptions - some of the individual level variables are significantly associated with BMI. As expected, better educated individuals have higher levels of BMI, and the same holds for wealthier individuals. Moreover, BMI increases with age, but at a

---

make him *less likely* than individual B to live closer to a medical mission. This would be incorrect.

[39]We compute the expected change in y associated with a $p\%$ increase in x as $\ln\left(\frac{100+p}{100}\right) \times \hat{\beta}$.

Table 3: OLS Estimates

	BMI				
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)
Distance from Protestant medical mission (log)	-0.271*** (0.0527)	-0.370*** (0.0975)	-0.247*** (0.0912)	-0.246*** (0.0921)	-0.246*** (0.0923)
$\mathbb{1}$(Female)			-0.103 (0.0779)	-0.108 (0.0779)	-0.108 (0.0843)
$\mathbb{1}$(Married)			0.104 (0.104)	0.109 (0.103)	0.109 (0.0953)
$\mathbb{1}$(Primary school completed)			0.287*** (0.0890)	0.281*** (0.0892)	0.281*** (0.0873)
$\mathbb{1}$(Major ethnic group)			-0.135 (0.105)	-0.109 (0.105)	-0.109 (0.105)
Age			0.156*** (0.0221)	0.156*** (0.0222)	0.156*** (0.0225)
Age^2			-0.00148*** (0.000280)	-0.00147*** (0.000281)	-0.00147*** (0.000284)
Household size			-0.149** (0.0697)	-0.151** (0.0697)	-0.151** (0.0650)
Household size^2			0.00784* (0.00448)	0.00807* (0.00448)	0.00807* (0.00420)
$\mathbb{1}$(Wealth 2nd quartile)			0.287*** (0.0971)	0.285*** (0.0970)	0.285*** (0.0967)
$\mathbb{1}$(Wealth 3rd quartile)			0.536*** (0.0965)	0.538*** (0.0960)	0.538*** (0.0941)
$\mathbb{1}$(Wealth 4th quartile)			1.076*** (0.112)	1.083*** (0.113)	1.083*** (0.107)
Population in 2000 (000s)				0.00725 (0.00549)	0.00725 (0.00667)
Altitude (meters)				0.000853* (0.000458)	0.000853* (0.000460)
No. rivers				-0.0261 (0.0368)	-0.0261 (0.0368)
Latitude				0.209 (0.219)	0.209 (0.219)
Longitude				-0.246 (0.153)	-0.246* (0.138)
Population in 1900 (000s)				0.00000992 (0.000897)	0.00000992 (0.00147)
Cropland in 1900 (km^2)				0.00126 (0.00181)	0.00126 (0.00191)
No. colonial railways				-0.307* (0.177)	-0.307* (0.164)
N	7046	7046	7046	7046	7046
Mean Dependent Variable	20.42	20.42	20.42	20.42	20.42
District FE	No	Yes	Yes	Yes	Yes
Individual controls	No	No	Yes	Yes	Yes
Geographical controls	No	No	No	Yes	Yes
Historical controls	No	No	No	Yes	Yes
Spatial standard errors	No	No	No	No	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level in columns 1 to 4. Standard errors corrected for spatial correlation in column 5.

decreasing rate. The magnitude and significance of the coefficient of interest are almost unchanged across all specifications. Once district fixed effects, individual, and village level controls are included, the OLS estimates indicate that a 50 percent reduction in the distance from a Protestant medical mission is associated with an increase in BMI by 0.17.[40] In columns 1 to 4, robust standard errors are clustered at the primary sampling unit level, which corresponds to villages in rural areas and to wards in urban areas. Column 5 reports standard errors corrected for spatial correlation (Conley (2008)).[41]

5.2 Alternative Measures of Health and Proximity

In this paper, BMI and the logarithm of the minimum distance are our preferred variables of interest. Before proceeding further with our empirical analysis, we demonstrate that our findings are robust to alternative measures of health and proximity.

Results are confirmed when using height or a self-reported description of current health status to measure individual health.[42] The higher is the distance from a Protestant medical mission, the lower is the individuals' height. In particular, doubling the distance is associated with a decrease in height by about 0.75 centimeters. In addition, we find that an increase in distance from a historical medical facility is associated with an increase in the probability that an individual describes her health as "moderate" or worse and with a decrease in the probability that an individual describes her health as "good" or "very good". Table 4 presents the estimated coefficients. Figure A6 in the Appendix displays the ordered probit marginal effects.

When considering the logarithm of the distance, instead of the variable in levels, we assume that percentage changes in distance from a Protestant medical mission matter more than level changes. We test against this assumption by estimating a linear model. We do not find evidence of a linear relationship between the minimum distance from a Protestant medical mission and current individuals' BMI when the entire sample is considered. However, the existence of a linear relationship cannot be rejected if only relatively small distances are considered.[43]

Our findings are confirmed when we consider non-parametric measures of proximity, such as the presence of a Protestant medical mission nearby. On average, individuals living less than 5 or 25 kilometers away from the location of a historical medical mission have higher BMI than

---

[40]See footnote 39.

[41]Findings are robust to a change of the specification to a log-log model, where the coefficient can be interpreted as the the percentage change in BMI following a 1 percent change in the distance from a location of a Protestant medical mission. Results for this alternate specification are included in table A2 in the Appendix.

[42]Height has been widely used in the literature and includes the long-run effect of fetal and childhood nutritional limitations and disease environment (Fogel (1994), Schultz (2002), Schultz (2003), Weil (2005), and Steckel (2008)).

[43]We restrict the sample to individuals living less than 50 kilometers from a historical medical facility. However, the existence of a linear relationship cannot be rejected if individuals living up to 68 kilometers away are considered.

Table 4: Alternative Measures of Health

	Height (cm)	Self-reported Health Status	
	OLS	OLS	Ordered Probit
	(1)	(2)	(3)
Distance from Protestant medical mission (log)	-0.754**	-0.0627**	-0.129*
	(0.339)	(0.0318)	(0.0678)
N	7046	6995	6995
Mean Dependent Variable	159.08	2.69	2.69
District FE	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level. Column 3 contains the estimated coefficient of the latent model. The estimated coefficient in column 4 suggests that the latent variable increases with the proximity to a Protestant medical mission.

individuals who do not (by about 0.78 and 0.48, respectively). We find that the presence of a medical mission within 50 kilometer plays a role only when interacted with the logarithm of the minimum distance. For individuals living less than 50 kilometers away, doubling the distance would decrease their BMI by 0.25. For those living more than 50 kilometers away, doubling the distance would decrease their BMI by 0.19. Table 5 presents the estimation results.[44]

5.3 Geographical and Historical Selection

As discussed in section 3.3, the non-random location of missionary stations may introduce bias in the OLS estimates presented in table 3.[45]

In columns 1 to 3 of table 6, we restrict the estimation sample to areas within 49 kilometers (75th percentile), 29 kilometers (50th percentile), and 14 kilometers (25th percentile) from a Protestant mission. Such restrictions represent an attempt to correct for possible differences between regions with and without Protestant missions. Our coefficient of interest, β remains statistically significant as the sample size decreases, suggesting that our results are not driven by endogenous missionary location. The estimated coefficients on the logarithm of the distance from a Protestant medical mission are -0.331, -0.272, and -0.403, respectively. For individuals living within the 25th percentile of the minimum distance from a Protestant mission, a 50 percent

[44]To take into account the fact that the minimum distance may not capture actual proximity due to geographic and historical characteristics, we interact it with the presence of colonial railways, access to rivers and altitude, alternatively. The interaction term is not significant in any of the alternative specifications. Results are available upon request.

[45]Endogenous migration decisions of individuals may also introduce a bias in baseline estimates. While we do not observe the migration history of survey respondents and their ancestors, we can control for the mass migration occurred after the partition of India of 1947, and the secession of Bangladesh from Pakistan in 1971. Our findings are confirmed when we estimate our specification (1) in a subsample that excludes those Indian states that were more affected by these episodes of mass migration (Assam, West Bengal, and Rajasthan). Results are available upon request.

Table 5: Alternative Measures of Proximity

	BMI					
	Baseline sample	Distance ≤ 50 km	Baseline Sample			
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)
Distance from Protestant medical mission (km)	-0.00291 (0.00208)	-0.0433*** (0.00859)				
$\mathbb{1}$(Medical mission within 5 km)			0.775*** (0.224)			
$\mathbb{1}$(Medical mission within 25 km)				0.479** (0.217)		
$\mathbb{1}$(Medical mission within 50 km)					0.0578 (0.161)	
Distance from Protestant medical mission (log)						-0.276*** (0.0912)
$\mathbb{1}$(Medical mission within 50 km) × Distance from Protestant medical mission (log)						-0.0797* (0.0466)
N	7046	2666	7046	7046	7046	7046
Mean Dependent Variable	20.42	20.42	20.42	20.42	20.42	20.42
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level.

increase in the distance decreases BMI by 0.16. If only individuals living within the median distance are included, doubling the distance from a medical mission decreases the BMI by about 0.19, while if only the upper quartile is excluded from the sample, we find that a 100 percent increase in the distance decreases BMI by 0.23.

We then augment our set of controls to include measures of proximity to a Protestant mission - with or without a medical facility - and a Catholic mission. This alternative specification allows us to address two issues. First, it provides an additional way of testing whether unobserved heterogeneity across locations is the main driver of our findings. Common factors may have determined the location decisions of Protestant and Catholic missionaries, independent of the missionary activities. If these factors were driving our results, we would find a significant correlation between individual health today and these additional variables. Second, it allows us to assess whether our previous results are driven by vicinity to a medical mission or by proximity to other missionary activities. Column 5 to 6 show this is not the case. Our results indicate that that proximity to a generic Protestant or Catholic mission and current individual BMI are not correlated. In contrast, our coefficient of interest β remains negative and statistically significant

at the 5 percent level.

Table 6: Geographical and Historical Selection

	BMI				
	Distance $\leq$ percentile			Baseline	
	75%	50%	25%	Sample	
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)
Distance from Protestant medical mission (log)	-0.331*** (0.120)	-0.272* (0.141)	-0.403*** (0.128)	-0.246** (0.102)	-0.216** (0.107)
Distance from Protestant mission (log)				-0.000107 (0.103)	0.0347 (0.111)
Distance from Catholic mission (log)					-0.131 (0.117)
N	5230	3476	1825	7046	7046
Mean Dependent Variable	20.54	20.70	20.91	20.42	20.42
District FE	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level. Mission distance percentiles based on all sample statistics. 14 kilometers: 25th; 29 kilometers: 50th; 49 kilometers: 75th.

5.4 Instrumental Variable Estimation

We adopt an instrumental variable approach to deal with measurement error and the possible endogenous selection of Protestant missionaries into building a health facility. Table 7 shows the estimation results.[46] Columns 1 to 3 report the estimated coefficients for the first stage regression, without and with individual, geographic, and historical controls. The distance from a Protestant medical mission is negatively correlated with the instrument $Hospital\ share_{ivd}^{50}$. In all specifications, the coefficient of interest in equation (4), λ, is significant at the 1 percent level. Columns 4 to 6 show the 2SLS estimation results. The first stage F statistics is largely above 10, suggesting that $Hospital\ share_{ivd}^{50}$ is a non-weak instrument for our endogenous regressor $Hospital\ distance_{ivd}$.[47]

The relationship between distance from a medical mission and individuals' BMI presented in section 5.1 is robust to the instrumental variable approach: proximity to a Protestant medical missions increases BMI. Even though the level of significance of β decreases to 5 percent in column 5, the magnitude of the estimated coefficient more than doubles if compared to the OLS estimates

---

[46]We build our instrument using a 50 kilometer radius. In section 6 we show that our results are robust to changes in the radius.

[47]We report the Kleibergen-Paap F statistics, which is robust to non-iid errors.

in table 3.[48] When individual and village level controls are included, a 50 percent reduction in the distance from a Protestant medical facility increases current BMI by 0.43.

Table 7: IV Estimates

	Distance from Protestant medical mission (log)			BMI		
	OLS (1)	OLS (2)	OLS (3)	2SLS (4)	2SLS (5)	2SLS (6)
Hospital share (50km radius)	-1.138*** (0.211)	-1.012*** (0.238)	-0.774*** (0.215)			
Distance from Protestant medical mission (log)				-0.769*** (0.249)	-0.622** (0.300)	-0.734* (0.401)
Kleibergen-Paap rk Wald F statistic	-	-	-	29.02	18.03	13.01
N	7046	7046	7046	7046	7046	7046
Mean dependent variable	20.42	20.42	20.42	20.42	20.42	20.42
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	No	Yes	Yes	No	Yes	Yes
Geographical controls	No	Yes	Yes	No	Yes	Yes
Historical controls	No	Yes	Yes	No	Yes	Yes
Distance from Protestant mission (log)	No	No	Yes	No	No	Yes
Distance from Catholic mission (log)	No	No	Yes	No	No	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level.

6 Robustness Checks and Additional Results

We perform a series of robustness checks to test the sensitivity of our findings. In this section, all specifications are estimated with OLS, and, unless otherwise noted, contain the minimum distance from a Protestant mission and from a Catholic mission as additional controls.

6.1 Exclusion Restriction

The exclusion restriction for the validity of our IV approach requires *Hospital share*$_{ivd}^{K}$ to affect current individuals' health outcomes only through the proximity to a Protestant medical mission. This assumption may be problematic as societies with and without medical missionaries might differ along dimensions other than their likelihood to have a health facility. These differences may affect development and contemporary health outcomes independent of the health channel, which would represent a violation of the exclusion restriction.

---

[48]This significant increase in the magnitude of the effect is likely due to measurement error in our measure of distance and the consequent attenuation bias of the OLS estimates. The geocoding of Protestant medical missions is based on Google coordinates for a string containing the names of *state*, *district*, and *village/city*, which may suffer from measurement issues. While we acknowledge that the increase in the magnitude of the effect may be driven by violations of the exclusion restriction, in section 6 we show that this is unlikely.

We address this concern in three ways. First, we show that, once controlling for the number of missionary stations, societies involved and not involved in medical activities are not different from each other. Second, we follow Lewbel (2012) to supplement the model with external instruments and confirm the validity of our instrument by performing a test of overidentifying restrictions. Finally, we relax the exclusion restriction as in Conley et al. (2012) and show that the direct effect of the instrument on individual BMI would have to be quite large to render our IV results insignificant.

Societies With and Without Medical Missionaries. To test whether missionary societies involved in medical activities are systematically different from those that are not, we exploit information about 295 missionary society active worldwide in 1902.[49] Table A5 in the Appendix compares income, diffusion, and investments of missionary societies outside of India depending on whether they had medical missionaries or physicians. On average, societies involved in medical activities were founded earlier and had a larger number of stations worldwide. However, when looking at activities, size and income of each missionary station affiliated with medical and non-medical societies, no significant difference can be detected.

Testing Overidentifying Restrictions. If the number of instruments is strictly larger than the number of endogenous regressors, the validity of instruments can be checked by performing a test of overidentifying restrictions. While this is not a test for exogenenity of the instruments, but rather that the additional restrictions imposed by having additional instruments are valid, a rejection of the null hypothesis would cast doubt on the validity of the instruments. In order to achieve overidentification, we follow Lewbel (2012) and supplement the model with heteroskedasticity based external instruments.[50] We estimate the model in equation (1) by 2SLS using $Hospital\ share^{50}_{ivd}$ and the generated variables as instruments for $Hospital\ distance_{ivd}$.[51] Table A6 reports estimates obtained using only generated instruments and using both generated and excluded instruments, together with the Hansen J statistic, its degrees of freedom, and p-value. In both cases, the null hypothesis cannot be rejected.

Plausibly Exogenous Instrument. Following Conley et al. (2012), we relax the exclusion restriction and assume that the *potential* direct effect of the instrument on individual BMI is uniformly distributed in an interval $[0, \delta]$, with $\delta > 0$.[52] By varying δ, we identify the threshold

---

[49] Data are from the *Centennial Survey of Foreign Missions*. For simplicity, we only focus on "societies directly engaged in conducting foreign missions" (*class 1 societies* according to Dennis (1902)).

[50] The Pagan-Hall test strongly rejects the null hypothesis of homoskedastic disturbances ($\chi^2(170) = 279.602$), so that Lewbel's method can be applied. We implement it in Stata using the `ivreg2h` command (Baum and Schaffer (2012)).

[51] For simplicity, we center all variables to remove district fixed effects. This allows us to generate 19 additional instruments.

[52] Standard IV estimation requires that $\gamma = 0$ in the following equation $y = X\beta + Z\gamma + \epsilon$. Conley et al. (2012) relax the restriction

at which the second-stage coefficient on the distance from a medical missions becomes insignificant at the 10% level. Figure A7 shows how large the exclusion restriction violation would need to be in order to invalidate the reduced form results: as long the direct effect of *Hospital share*$_{ivd}^{50}$ on BMI is smaller than 0.24, our second stage is still significant at the 10% level.[53] To gauge magnitudes, we compare this to the overall reduced-form effect of our instrument on individual BMI, which is 0.63.[54] Therefore, to render our IV results insignificant, the direct effect of *Hospital share*$_{ivd}^{50}$ on BMI should be about as large as 40 percent of the overall effect.

6.2   Estimation sample

In section 3.2, we discuss the use of BMI as a measure of health status and our decision to restrict the estimation sample to individuals of age 20 to 60 and with BMI between 15 and 30. Given the non-monotonic relationship between individuals' BMI and health status, we expect our results to weaken when we include individuals in the far right tail of the BMI distribution (i.e., those who would fall into the obese category, as defined by the WHO). Moreover, as teenagers and elderly may experience changes in BMI due to biological growth and aging that are not directly related to their health status, we expect our results to be affected by the inclusion of individuals 15-20 and 60-65 in the sample. Our results are robust, but weaker, if we extend the estimation sample to individuals of age 15 to 65 and with BMI 18.5 to 30, 15 to 35, and 15 to 45.[55] The sample size decreases drastically once we include only people with BMI above the pre-obesity threshold. As hypothesized, however, the relationship between proximity to a Protestant medical mission and current individuals' BMI disappears. Table A3 in the Appendix shows the results for these different estimations samples.

6.3   Instrument

We test whether our findings are robust to the way we construct the instrumental variable. First, we find our results from the IV estimation to be robust to variations in the radius used to construct the instrument. Second, we show that our findings do not change if we construct our instrumental variable as the sum of societies' medical shares, defined as in equation (3), for all missions, both medical and non-medical. To this end, we use data from the *World Atlas Of Christian Missions (1911)* which includes information about medical and non-medical missions

---

that $\gamma = 0$, and replace it with the assumption that gamma is close to, but not necessarily equal to, zero.

[53]Figure A7 is produced using the `plausexog` code by Clarke (2014) and the *local to zero* approach (Conley et al. (2012)).

[54]This is obtained by estimating the following model by OLS: $BMI_{ivd} = \alpha_d + \beta Hospital\ share_{ivd}^{50} + X'_{ivd}\gamma_1 + W^{g'}_{vd}\gamma_2 + W^{h'}_{vd}\gamma_3 + \epsilon_{ivd}$.

[55]See footnote 26 for detailed description of the WHO cutoffs.

active in 1911 and their affiliated missionary societies.[56] Table A4 in the Appendix presents the estimation results.

6.4 Proximity to Protestant Schools

Parallel to and in most cases preceding the development of medical activities, Protestant missionaries invested in several educational, cultural and philanthropic activities. Previous literature stressed the role of Protestantism in increasing human capital by focusing on its contribution to the advancement of literacy and education.[57]

To assess whether the long-term effect is transmitted via an educational, rather than medical, channel, we control for the minimum distance to different types of educational institutions. Figure A8 in the Appendix shows the spatial distribution of Protestant boarding schools, high schools and universities operating in colonial India in 1902. While there is some overlapping between the locations of medical facilities and schools, it is far from perfect. We exploit this variation to disentangle the effect of proximity to educational from that related to proximity to medical facilities. To this end, we augment our empirical specification with the logarithm of the minimum distance between individual i and a Protestant educational institution. As shown in table A7, we still find a negative and statistically significant relationship between hospital distance and current BMI. Proximity to Protestant educational institutions does not seem to play a role, both when excluding (column 4) and including (column 5) the distance from a Protestant medical mission.

6.5 Medical Mission Activities

We investigate whether the effect of proximity to a Protestant medical mission varies with the size of the medical facility and with the type of medical procedures performed.

First, we interact the logarithm of the minimum distance with an indicator variable equal to one if the closest medical mission had a total number of patients, treatments and surgeries above the median.[58] Second, we test the hypothesis that it was surgery that determined the acceptance of Western medicine among the Indian people by interacting the distance with an indicator variable equal to one if surgeries were performed at the closest medical mission.[59]

---

[56]We classify a mission to be a medical mission if the presence of a doctor is recorded. Non-medical missions were often affiliated with societies with relatively low *Society hospital share*. The correlation coefficient between the two versions of the instrument is 0.89 and statistically significant at the 1 percent level.

[57]See e.g. Cagé and Rueda (2014) and Mantovanelli (2014).

[58]We collect additional information about activities and characteristics of Protestant medical missions from the *Centennial Survey of Foreign Mission*. In particular, for each medical mission we retrieve data about the total number of patients, treatments and surgeries in 1902 and we use their sum as a proxy for the size of the medical mission. When more than one hospital or dispensary are present in the same village (city), the total number of patients, treatment and surgeries is computed at the village (city) level.

[59]Fitzgerald (2001) suggests it was surgery that determined acceptance of Western medicine among the Indian people and their greater willingness to use the resources of medical missions.

Table A8 shows the OLS estimates for these alternative specifications. The estimates indicate that there is no independent effect on individuals' health coming from the closest medical mission being sizable. Proximity to a Protestant medical mission, however, has a larger impact on individuals who live close to a mission with an above median number of treatments and patients. We show that individuals for which surgeries were performed at the closest medical mission have, on average, a higher BMI. However, when an interaction term is included, the independent effect of proximity to surgery disappears. The effect of proximity to a medical mission is larger when surgery was available. These differences are however not statistically significant.

7 Long-Term Transmission Channels

In this section, we look at the mechanisms underlying the long-term effect of Protestant medical missions on individual health. We identify three potential channels: infrastructure, health potential (health status of previous generations, income and nutrition), and health culture (hygiene and health awareness).[60]

7.1 Persistence of Infrastructure

In the previous sections, we omit current access to health care from our baseline specification to avoid identification issues. We include it now as we want to investigate the independent effects of long-term and short-term access to health care. We measure short-term access to health care today with the minimum distance (logarithm) from a village with at least one hospital.[61] We first assess how proximity to a Protestant medical mission and access to current health facilities correlate. We then estimate the following model:

$$
\begin{aligned}
BMI_{ivd} = {} & \alpha_d + \beta_1 Hospital\ distance^{LR}_{\ ivd} + \beta_2 Hospital\ distance^{SR}_{\ ivd} \\
& + X'_{ivd}\gamma_1 + W^{g\prime}_{\ vd}\gamma_2 + W^{h\prime}_{\ vd}\gamma_3 + \epsilon_{ivd}
\end{aligned}
\tag{5}
$$

Our coefficients of interest are β_1 and β_2. If the only relevant channel was the persistence of health infrastructure over time, we would expect the coefficient on $Hospital\ distance^{LR}_{\ ivd}$ to become smaller in magnitude and, potentially, insignificant once including $Hospital\ distance^{SR}_{\ ivd}$ in the model.

---

[60]In section A.2 in the Appendix, we illustrate these transmission channels in the context of a health production function.

[61]Using information from the Village Directory of the 2001 Census of India, we successfully geocoded 4,617 villages in Assam, Karnataka, Madhya Pradesh, Maharashtra, Rajasthan, Uttar Pradesh and West Bengal with at least one allopathic hospital and computed the aerial minimum distance between WHS respondents and the allopathic hospitals. For the sake of simplicity, we exclude Ayurvedic, Unani and homeopathic hospitals, dispensaries, maternity and child welfare centers, maternity homes, primary health centers, family welfare centers, T. B. clinics, nursing homes, private and subsidized medical practitioners and community health workers.

Table 8 presents the estimation results. The long-run effect of access to health care facilities on individuals' health is not driven by persistence of infrastructure. Our measure of short-term access to medical facilities is significantly and positively correlated with the distance from a Protestant medical mission, but it does not seem to independently affect individuals' health.[62] Even when controlling for current access to health care, the distance from a Protestant medical facility remains significantly and negatively associated with BMI. The magnitude of the coefficient on $Hospital\ distance^{LR}{}_{ivd}$ is practically unchanged compared to the results in table 6.

Table 8: Transmission Channels: Persistence of Infrastructures

	Distance from all. hospital (log)	BMI	
	OLS (1)	OLS (2)	OLS (3)
Distance from Protestant medical mission (log)	0.244*** (0.0518)		-0.210* (0.108)
Distance from allopathic hospital (log)		-0.0767 (0.0701)	-0.0532 (0.0703)
N	7046	7046	7046
Mean Dependent Variable	2.54	20.42	20.42
District FE	No	Yes	Yes
Individual controls	No	Yes	Yes
Geographical controls	No	Yes	Yes
Historical controls	No	Yes	Yes
Distance from Protestant mission (log)	No	No	Yes
Distance from Catholic mission (log)	No	No	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level.

7.2 Health Potential and Nutrition

The cross-sectional nature of the WHS does not allow us to match survey respondents with health outcomes of previous cohorts. Height as an adult, however, includes the long-run effect of fetal and childhood nutrition and exposure to diseases, together with the genetic height potential. Via the intergenerational transmission of health, access to health care and the health status of individuals in one generation could affect the height of individuals in the next generation. The results presented in table 4 indicate that proximity to a Protestant medical mission positively affects individual height, providing preliminary evidence of the fact that the health potential may indeed be an important transmission channel.

---

[62]This findings are in line with Banerjee et al. (2004) and Das and Hammer (2007), which show that health care provision in India is highly dysfunctional, due for example to the high levels of absenteeism, and not necessarily impacting individuals' health outcomes in a positive way.

We now test whether these results are driven by access to better nutrition. To this end, we augment our baseline specification to include measures of quality and quantity of available food (expenditure on food and number of servings of fruit and vegetables consumed per day).[63] We also control for total household expenditure, as a proxy for income. Table 9 presents the estimation results. As expected, we find a positive and significant relationship between household expenditure on food and individual health. The same holds for the total household expenditure. We still find that the proximity to a Protestant medical mission has an independent effect on individual health outcomes. Moreover, the magnitude of the estimated coefficients is not affected, suggesting that the long-term effect of access to health care on current individual height and BMI cannot be simply explained by better nutrition and higher income.

Table 9: Transmission Channels: Health Potential

	Height (cm)		BMI	
	OLS (1)	OLS (2)	OLS (3)	OLS (4)
Distance from Protestant medical mission (log)	-0.653** (0.330)	-0.628* (0.328)	-0.224** (0.113)	-0.215* (0.113)
Fruit & vegetables (servings per day)	-0.127 (0.127)	-0.135 (0.121)	0.0121 (0.0431)	0.00937 (0.0421)
Household expenditure on food (log)	0.833*** (0.263)	0.524* (0.299)	0.349*** (0.0820)	0.240** (0.0990)
Fruit & vegetables (servings per day) × Household expenditure on food (log)	0.0138 (0.0172)	0.0147 (0.0165)	-0.00183 (0.00590)	-0.00150 (0.00577)
Household total expenditure (log)		0.508* (0.267)		0.180** (0.0875)
N	5729	5729	5729	5729
Mean Dependent Variable	158.83	158.83	20.47	20.47
District FE	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes
Distance from Protestant mission (log)	Yes	Yes	Yes	Yes
Distance from Catholic mission (log)	Yes	Yes	Yes	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level.

7.3 Health Culture and Practices

The quality of medical missionary medicine increased in the final years of the nineteenth century, especially after the Bombay Missionary Conference of 1892-93. At that time, European ideas of

---

[63]Expenditure on food includes the value of any food produced and consumed by the household, and excludes alcohol, tobacco and restaurant meals. Moreover, respondents are asked how many servings of fruit and vegetables they consume per day. Unfortunately, no information about other aliments is included in the survey.

modern medicine spread worldwide through the activities of medical missionaries.[64] We use data from the DLHS-3 to investigate whether proximity to a health care facility in the past may have affected individual habits regarding hygiene, self and preventive care and health awareness.[65] The empirical specification is as follows:

$$Health\ culture_{ivd} = \alpha_d + \beta Hospital\ distance_{vd} + X'_{ivd}\gamma_1 + W^{g\prime}_{vd}\gamma_2 + W^{h\prime}_{vd}\gamma_3 + \epsilon_{ivd} \qquad (6)$$

where $Hospital\ distance_{vd}$ is now the minimum distance between a DLHS-3 village v and a Protestant medical mission (logarithm). We focus on outcomes related to hygiene and health practices, and perinatal maternal and child care. In particular, $Health\ culture_{ivd}$ is, alternatively, an indicator variable equal to one if the respondent treats water to make it safer to drink, if she practices open defecation, or if she turns to traditional healers when she gets sick. To measure disease awareness, we combine the respondent's answers to seven questions regarding knowledge about the danger signs of pneumonia and create an index using a principal component analysis. Finally, to investigate practices related to maternal and child health, we use indicator variables for receiving or purchasing iron folic acid tablets or bottles during the last pregnancy, for receiving a full antenatal check-up (ANC) during last pregnancy, for having a safe delivery (i.e., assisted by skilled personnel, and for taking the newborn for a check-up after birth).

Table 10 presents the estimation results.[66] Our measures of hygiene, preventive care and health knowledge are significantly associated with the distance from a Protestant medical mission. In particular, living far from a village historically equipped with a Protestant medical mission decreases the probability of treating water and it increases the probability of practicing open defecation. Interestingly, we do not find any effect on the use of traditional medicine; the fraction of people reporting to turn to traditional healers and home treatments when sick is however quite small. Moreover, distance from a Protestant medical mission is associated with a lower knowledge of the early signs of pneumonia. The closer to a Protestant medical mission, the better the practices related to maternal and child health. We find that proximity to a historical medical facility increases women's probability of taking iron supplements while pregnant, of receiving a full antenatal check-up, and of having a safe delivery. Finally, it increases children's probability of receiving a medical check-up within 24 hours since birth.[67]

---

[64]See Fitzgerald (2001). New scientific discoveries and theories set the foundations of contemporary biomedicine and gained gradual acceptance in Europe and the United States. The pivotal role of the *germ theory of disease* in modern medicine is often described as "the most important single concept for the history of medicine" (Worboys (2000)).

[65]See section 3 for more details. Table A9 in the Appendix presents the descriptive statistics of the variables included in the analysis.

[66]In columns 1 to 3 we use information from the Household survey, while in columns 4 to 8 we consider responses from the Ever Married Women questionnaire. Reported are the marginal effects at the average value of the independent variables. As the number of observations is much larger than the number of fixed effects, the incidental parameter problem associated with estimating a probit model with fixed effects is not a concern (Greene et al. (2002), Fernández-Val (2009)).

[67]All results are robust to estimating linear probability models and to including the presence of a health facility today in DLHS-3

8   Conclusion

In this paper, we investigate the long-term effect of access to health care on current individual health outcomes by focusing on the medical missionary enterprise in India during the second half of the nineteenth century. We combine historical information about the location and activities of Protestant medical missions in colonial India with contemporary individual level data. We show that proximity to the location of a Protestant medical mission positively affects current individuals' body mass index (BMI) and other individual health indicators, such as height and self-reported health status. We verify that our findings are not driven by other non-medical missionary activities, and show that the vicinity to Protestant missions without health facilities does not affect current health outcomes. The investigation of potential transmission channels shows that persistence of infrastructure and current access to health care do not play a role, while improvements in health potential and changes in health culture and preventive care are the main drivers of the long-term effect. In particular, we find that proximity to a historical medical facility improves health habits related to hygiene, health awareness, and perinatal maternal and child health.

This paper represents a first attempt to analyze the persistence of health in developing countries and to investigate how exposure to Protestant medical missions may have determined advancements in current individual health. In light of the existence of long-run effects, the expansion of healthcare access becomes an even more critical priority, as it beneficially affects both current and future generations. In a related project we investigate whether such patterns exist in the African continent. Further work should focus on explicitly connecting the long-term effect of access to health care, current health status, productivity, and economic outcomes.

---

village v as additional geographic control. Results are available upon request.

Table 10: Transmission Channels: Health Culture and Practices

	Hygiene & Health Practices				Maternal & Child Health			
	Pr(Treat water)	Pr(Open defecate)	Pr(Traditional medicine)	Knowledge pneumonia	Pr(Iron suppl. while pregnant)	Pr(Full ANC)	Pr(Safe delivery)	Pr(Child checked within 24hrs)
	Probit (1)	Probit (2)	Probit (3)	Tobit (4)	Probit (5)	Probit (6)	Probit (7)	Probit (8)
Distance from Protestant medical mission (log)	-0.0362***	0.0264*	-0.0321	-0.00992***	-0.0170*	-0.0311*	-0.0296*	-0.0376**
	(0.0134)	(0.0157)	(0.0333)	(0.00383)	(0.00987)	(0.0171)	(0.0174)	(0.0186)
N	273109	275386	219033	198225	198168	74017	72728	71864
Mean Dependent Variable	0.30	0.66	0.04	0.22	0.81	0.13	0.42	0.39
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distance from Protestant mission (log)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distance from Catholic mission (log)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

DHLS-3 data. Columns 1 to 3 uses information from the Household survey. Columns 4 to 8 uses responses from the Ever Married Women questionnaire. Respondents are ever married women 15-49. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village level. Marginal effects in columns 1 to 3 and 5 to 8.

Bibliography

ACEMOGLU, D., S. JOHNSON, AND J. A. ROBINSON (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, 1369–1401. [2]

ALESINA, A., P. GIULIANO, AND N. NUNN (2013): "On the Origins of Gender Roles: Women and the Plough," *The Quarterly Journal of Economics*, 128, 469–530. [2]

ALMOND, D. AND B. MAZUMDER (2011): "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance During Pregnancy," *American Economic Journal: Applied Economics*, 3, 56–85. [3]

BAI, Y. AND J. K.-S. KUNG (2014): "Diffusing Knowledge While Spreading God's Message: Protestantism and Economic Prosperity in China, 1840–1920," *Journal of the European Economic Association*. [3]

BANERJEE, A., A. DEATON, AND E. DUFLO (2004): "Health, Health Care, and Economic Development: Wealth, Health, and Health Services in Rural Rajasthan," *The American economic review*, 94, 326. [2], [27]

BANERJEE, R. AND A. SACHDEVA (2014): "Pathways to Preventive Health, Evidence from India's Rural Road Program," *Unpublished Manuscript*. [6]

BARRO, R. J. AND R. MCCLEARY (2003): "Religion and Economic Growth," NBER Working Papers 9682, National Bureau of Economic Research, Inc. [2]

BARRO, R. J. AND R. M. MCCLEARY (2005): "Which Countries Have State Religions?" *The Quarterly Journal of Economics*, 120, 1331–1370. [2]

BAUM, C. F. AND M. E. SCHAFFER (2012): "IVREG2H: Stata Module to Perform Instrumental Variables Estimation Using Heteroskedasticity-based Instruments," Statistical Software Components, Boston College Department of Economics. [23]

BECKER, S. O. AND L. WOESSMANN (2009): "Was Weber Wrong? A Human Capital Theory of Protestant Economic History," *The Quarterly Journal of Economics*, 124, 531–596. [3]

BHALOTRA, S. AND S. B. RAWLINGS (2011): "Intergenerational Persistence in Health in Developing Countries: The Penalty of Gender Inequality?" *Journal of Public Economics*, 95, 286 – 299. [50]

BHALOTRA, S., C. VALENTE, AND A. VAN SOEST (2010): "The Puzzle of Muslim Advantage in Child Survival in India," *Journal of Health Economics*, 29, 191–204. [3]

BRAINERD, E. AND N. MENON (2013): "Religion and Health in Early Childhood: Evidence from the Indian Subcontinent," Working Papers 65, Brandeis University, Department of Economics and International Business School. [3]

BURMASTER, D. E. AND E. A. CROUCH (1997): "Lognormal Distributions for Body Weight as a Function of Age for Males and Females in the United States, 1976–1980," *Risk Analysis*, 17, 499–505. [10]

CAGÉ, J. AND V. RUEDA (2014): "The Long-term Effects of the Printing Press in Sub-Saharan Africa," *Available at SSRN 2292660.* [2], [3], [11], [15], [25]

CHAUDHURY, N. AND J. S. HAMMER (2004): "Ghost Doctors: Absenteeism in Rural Bangladeshi Health Facilities," *The World Bank Economic Review*, 18, 423–441. [2]

CLARKE, D. (2014): "PLAUSEXOG: Stata Module to Implement Conley et al.'s Plausibly Exogenous Bounds," Statistical Software Components, Boston College Department of Economics. [24]

CONEUS, K. AND C. K. SPIESS (2012): "The Intergenerational Transmission of Health in Early Childhood - Evidence from the German Socio-Economic Panel Study," *Economics and Human Biology*, 10, 89 – 97. [50]

CONLEY, T. G. (2008): "Spatial Econometrics," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf and L. E. Blume, Basingstoke: Palgrave Macmillan. [18]

CONLEY, T. G., C. B. HANSEN, AND P. E. ROSSI (2012): "Plausibly Exogenous," *Review of Economics and Statistics*, 94, 260–272. [23], [24]

DAS, J. AND J. HAMMER (2007): "Money for Nothing: The Dire Straits of Medical Practice in Delhi, India," *Journal of Development Economics*, 83, 1–36. [27]

DEATON, A. (2008): "Height, Health, and Inequality: The Distribution of Adult Heights in India," *American Economic Review*, 98, 468–74. [8]

DENNIS, J. S. (1902): *Centennial Survey of Foreign Mission: A Statistical Supplement to "Christian Missions and Social Progress", Being a Conspectus of the Achievements and Results of Evangelical Missions in all Lands at the close of the Nineteenth Century*, New York: Revell. [6], [15], [23]

FERNÁNDEZ-VAL, I. (2009): "Fixed Effects Estimation of Structural Parameters and Marginal Effects in Panel Probit Models," *Journal of Econometrics*, 150, 71–85. [29]

FITZGERALD, R. (2001): "'Clinical Christianity': The Emergence of Medical Work as a Missionary Strategy in Colonial India," in *Health, Medicine and the Empire: New Perspectives on Colonial India*, ed. by B. Pati and M. Harrison, London: Sangam Books. [1], [5], [25], [29]

FOGEL, R. W. (1994): "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy," *American Economic Review*, 84, 369–95. [8], [18]

——— (2004): *The Escape from Hunger and Premature Death, 1700-2100*, no. 9780521808781 in Cambridge Books, Cambridge University Press. [8]

GALLEGO, F. A. AND R. WOODBERRY (2009): "Christian Missionaries and Education in Former African Colonies: How Competition Mattered," Working Papers ClioLab 2, EH Clio Lab. Instituto de Economía. Pontificia Universidad Católica de Chile. [3]

GLAESER, E. L. AND A. SHLEIFER (2002): "Legal Origins," *The Quarterly Journal of Economics*, 117, 1193–1229. [2]

GREENE, W., C. HAN, AND P. SCHMIDT (2002): "The Bias of the Fixed Effects Estimator in Nonlinear Models," *Unpublished Manuscript, Stern School of Business, NYU.* [29]

HARDIMAN, D. (2008): *Missionaries and Their Medicine: A Christian Modernity for Tribal India*, Manchester University Press. [4]

HJELMBORG, J. V., C. FAGNANI, K. SILVENTOINEN, M. MCGUE, M. KORKEILA, K. CHRISTENSEN, A. RISSANEN, AND J. KAPRIO (2008): "Genetic Influences on Growth Traits of BMI: A Longitudinal Study of Adult Twins," *Obesity*, 16, 847–852. [10]

JOHNSON, H. B. (1967): "The Location of Christian Missions in Africa," *Geographical Review*, 57, pp. 168–202. [11]

KLEIN GOLDEWIJK, K., A. BEUSEN, AND P. JANSSEN (2010): "Long-term Dynamic Modeling of Global Population and Built-up Area in a Spatially Explicit Way: HYDE 3.1," *The Holocene.* [7]

KLEIN GOLDEWIJK, K., A. BEUSEN, G. VAN DRECHT, AND M. DE VOS (2011): "The HYDE 3.1 Spatially Explicit Database of Human-induced Global Land-use Change Over the Past 12,000 Years," *Global Ecology and Biogeography*, 20, 73–86. [7]

Lewbel, A. (2012): "Using Heteroskedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models," *Journal of Business & Economic Statistics*, 30, 67–80. [23], [46]

Mantovanelli, F. G. (2014): "The Protestant Legacy: Missions and Literacy in India," *Unpublished Manuscript.* [3], [25]

Mathew, E. (1999): "Growth of Literacy in Kerala: State Intervention, Missionary Initiatives and Social Movements," *Economic and Political Weekly*, 2811–2820.

Menon, N. and K. McQueeney (2015): "Christianity and Infant Health in India," *Unpublished Manuscript.* [3]

Nunn, N. (2009): "The Importance of History for Economic Development," *Annual Review of Economics*, 1, 65–92. [2]

——— (2010): "Religious Conversion in Colonial Africa," *American Economic Review*, 100, 147–52. [11]

Pal, P. and J. Ghosh (2007): "Inequality in India: A Survey of Recent Trends," Working Papers 45, United Nations, Department of Economics and Social Affairs. [1]

Richter, J. (1908): *A History of Missions in India*, F. H. Revell. [1], [4]

Schultz, T. P. (2002): "Wage Gains Associated with Height as a Form of Health Human Capital," *American Economic Review*, 92, 349–353. [1], [8], [18]

——— (2003): "Wage Rentals for Reproducible Human Capital: Evidence from Ghana and the Ivory Coast," *Economics & Human Biology*, 1, 331–366. [1], [8], [18]

——— (2009): "Population and Health Policies," Working Papers 974, Economic Growth Center, Yale University. [3], [8]

Shiwaku, K., E. Anuurad, B. Enkhmaa, K. Kitajima, and Y. Yamane (2004): "Appropriate Body-Mass Index For Asian Populations and Its Implications For Policy and Intervention Strategies," *Lancet.* [8]

Srinivisan, R. (2010): "Health Care in India-Vision 2020," *New Delhi, India: Government of India, Planning Commission of India.* [1]

Steckel, R. H. (2008): "Heights and Human Welfare: Recent Developments and New Directions," NBER Working Papers 14536, National Bureau of Economic Research, Inc. [8], [18]

WAALER, H. T. (1984): "Height. Weight and Mortality The Norwegian Experience," *Acta Medica Scandinavica*, 215, 1–56. [8]

WEBER, M. (1930): *The Protestant Ethic and the Spirit of Capitalism, 1700-2100*, London: G. Allen and Unwin. [3]

WEIL, D. N. (2005): "Accounting for the Effect of Health on Economic Growth," Tech. rep. [8], [18]

WORBOYS, M. (2000): *Spreading Germs: Disease Theories and Medical Practice in Britain, 1865-1900*, Cambridge University Press. [29]
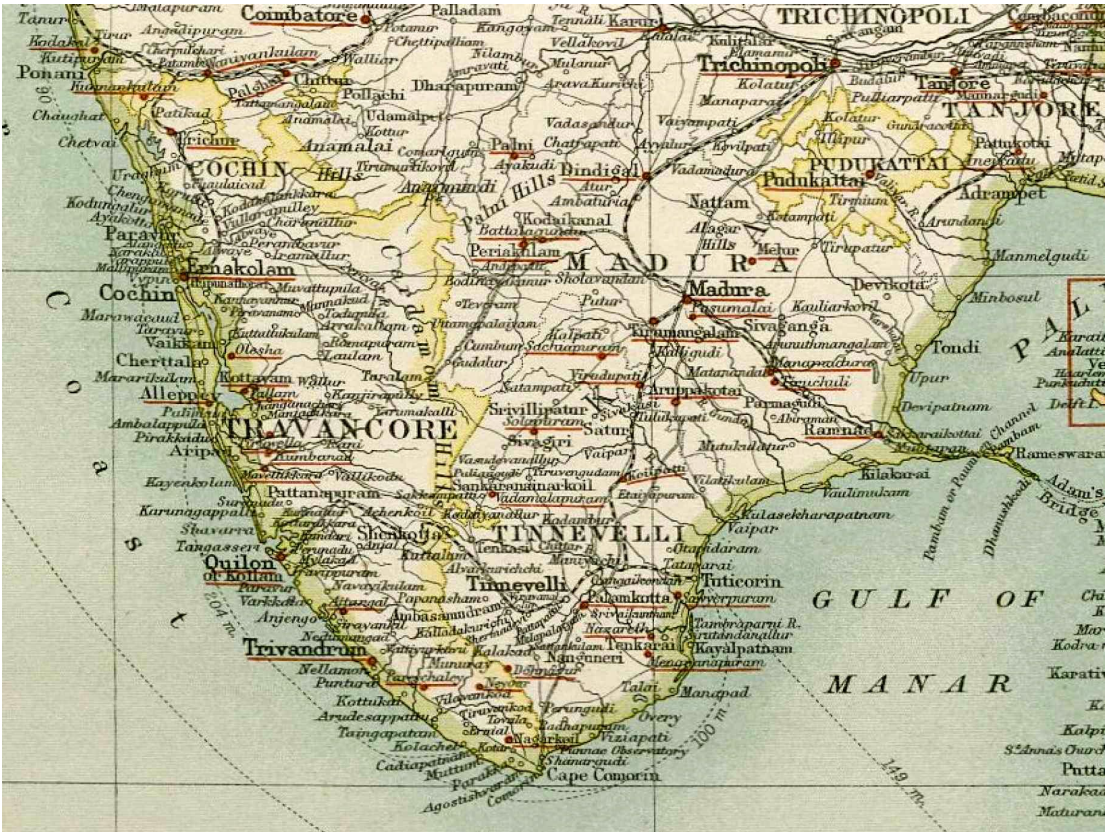
A Appendix

A.1 Additional Figures and Tables

Figure A1: Protestant missions locations as of 1910 (Statistical Atlas of Christian Missions)

Figure A2: Page from the Centennial Survey Survey of Foreign Missions
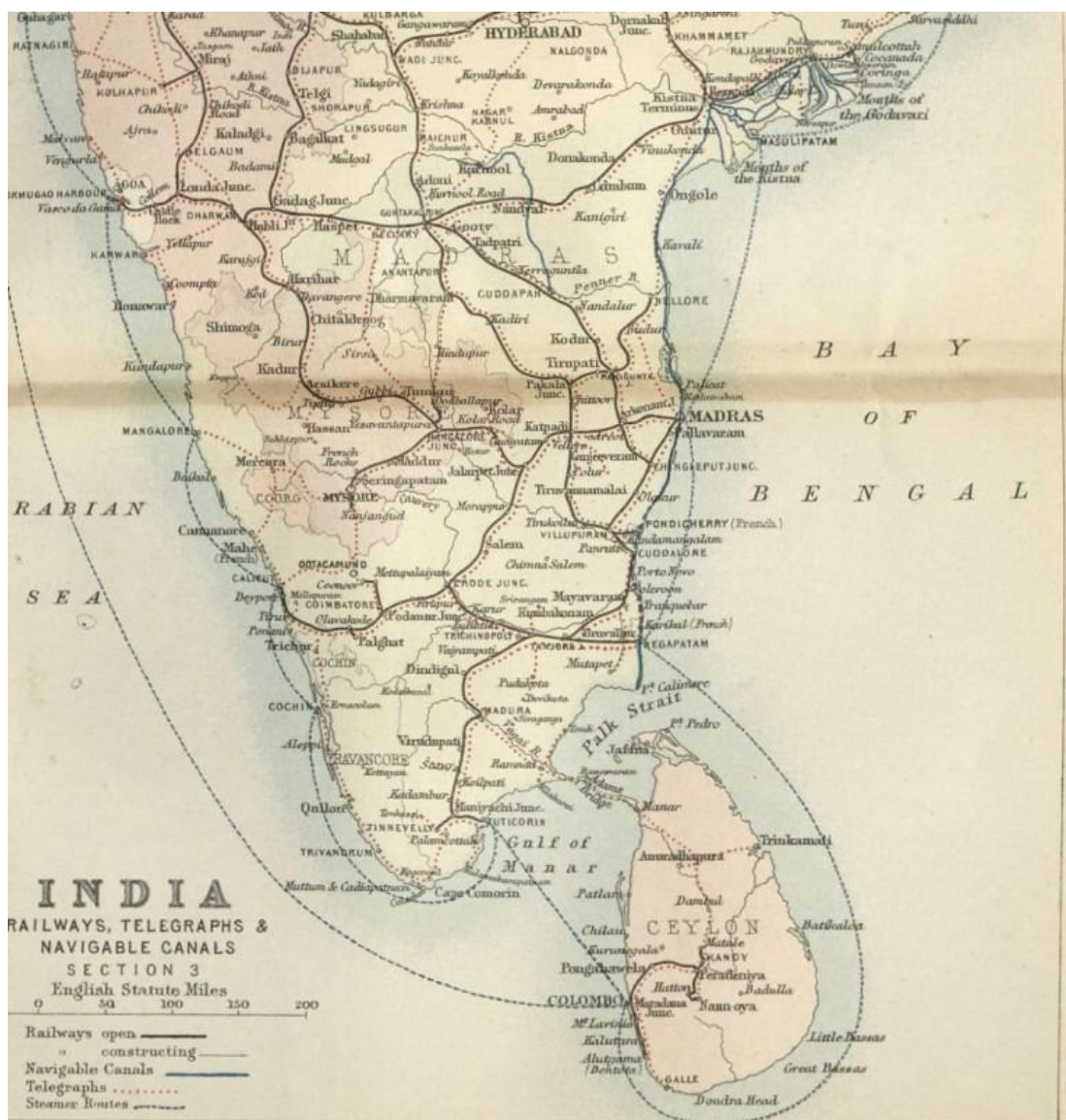
Figure A3: Status of Indian railways network as of 1891 (Constable's Hand Atlas of India)

Figure A4: Catholic missions locations as of 1911 (Atlas Hierarchicus)
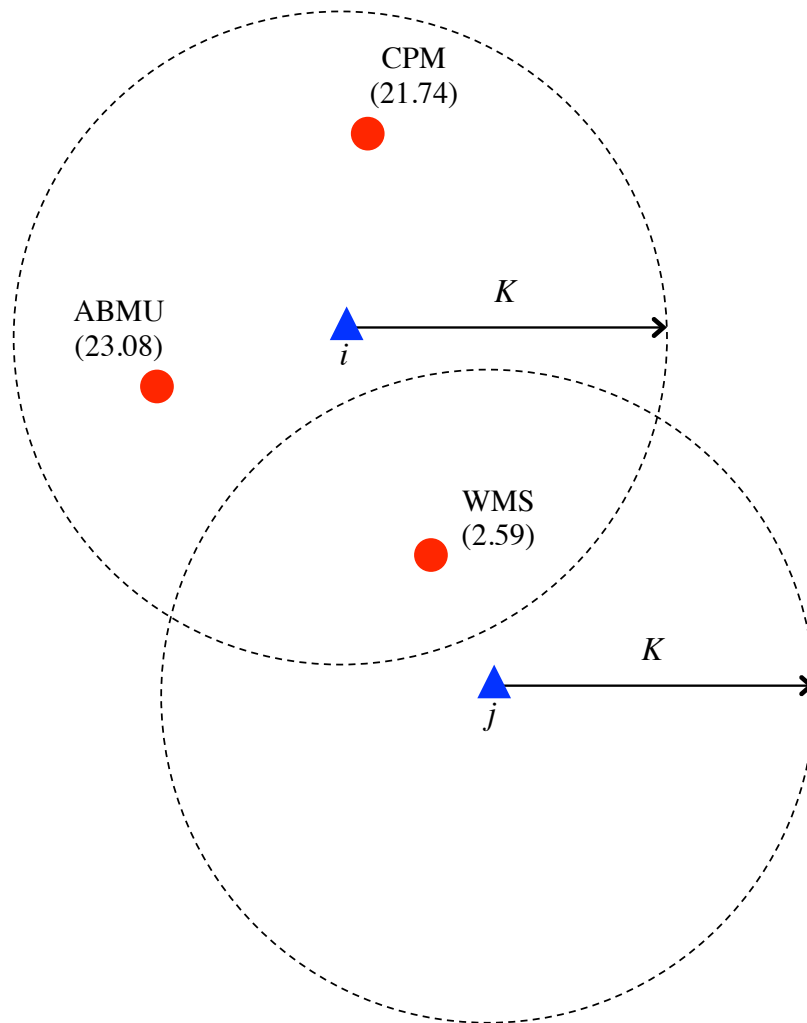
Figure A5: Instrumental Variable: An Illustrative Example

Note: Consider the case of two individuals (i and j) and three Protestant medical missions associated with three different missionary societies: Canadian Presbyterian Mission (CPM, with *Society Hospital Share = 21.74%*), American Baptist Missionary Union (ABMU, 23.08%), and Wesleyan Missionary Society (WMS, 2.59%). K is the buffer radius (in kilometers). For individual i, *Hospital share$_i^K$* $= 21.74 + 23.08 + 2.59 = 47.41$. For individual j, *Hospital share$_j^K$* $= 2.59$.
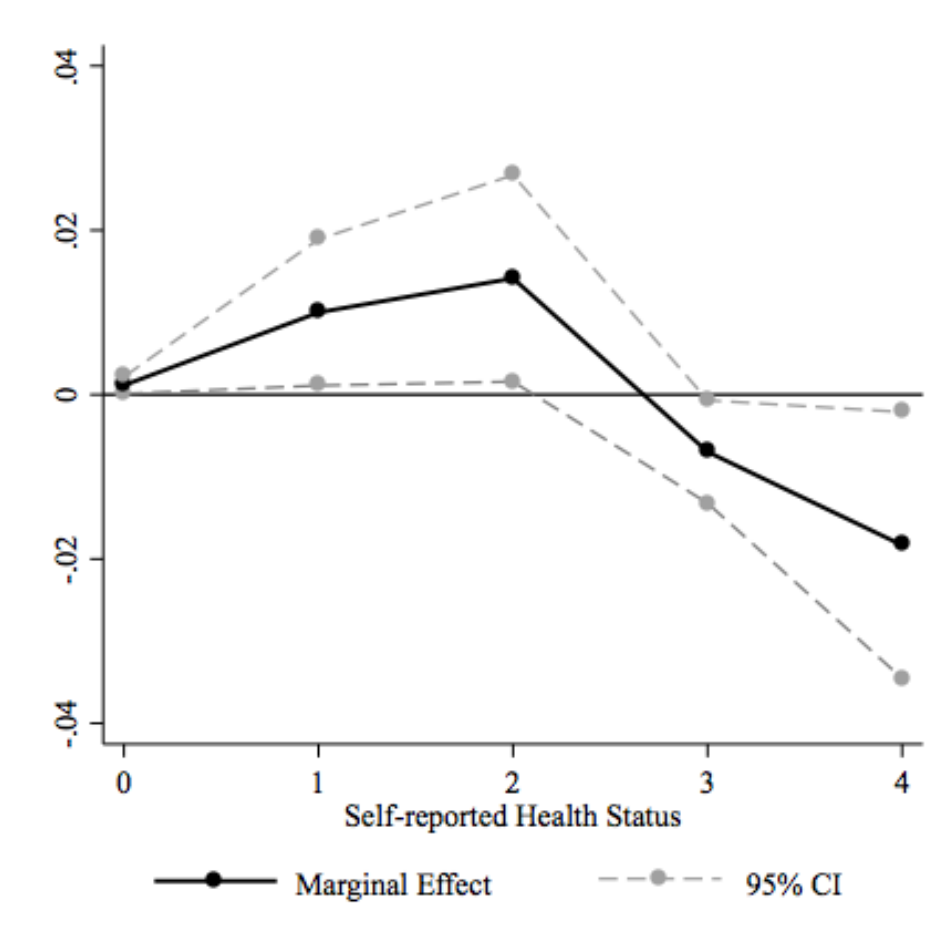
Figure A6: Self-reported Health: Ordered Probit Marginal Effects

Local To Zero (LTZ) method. Conley et al. (2012)
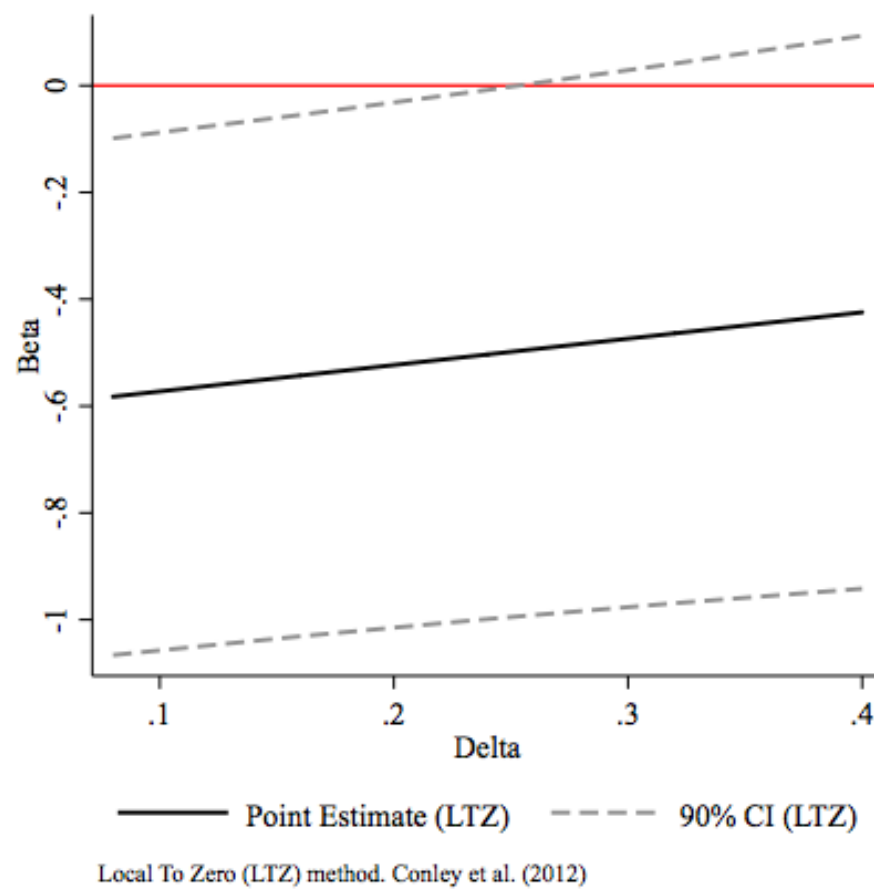
Figure A7: Estimated β by Direct Effect of Instrument
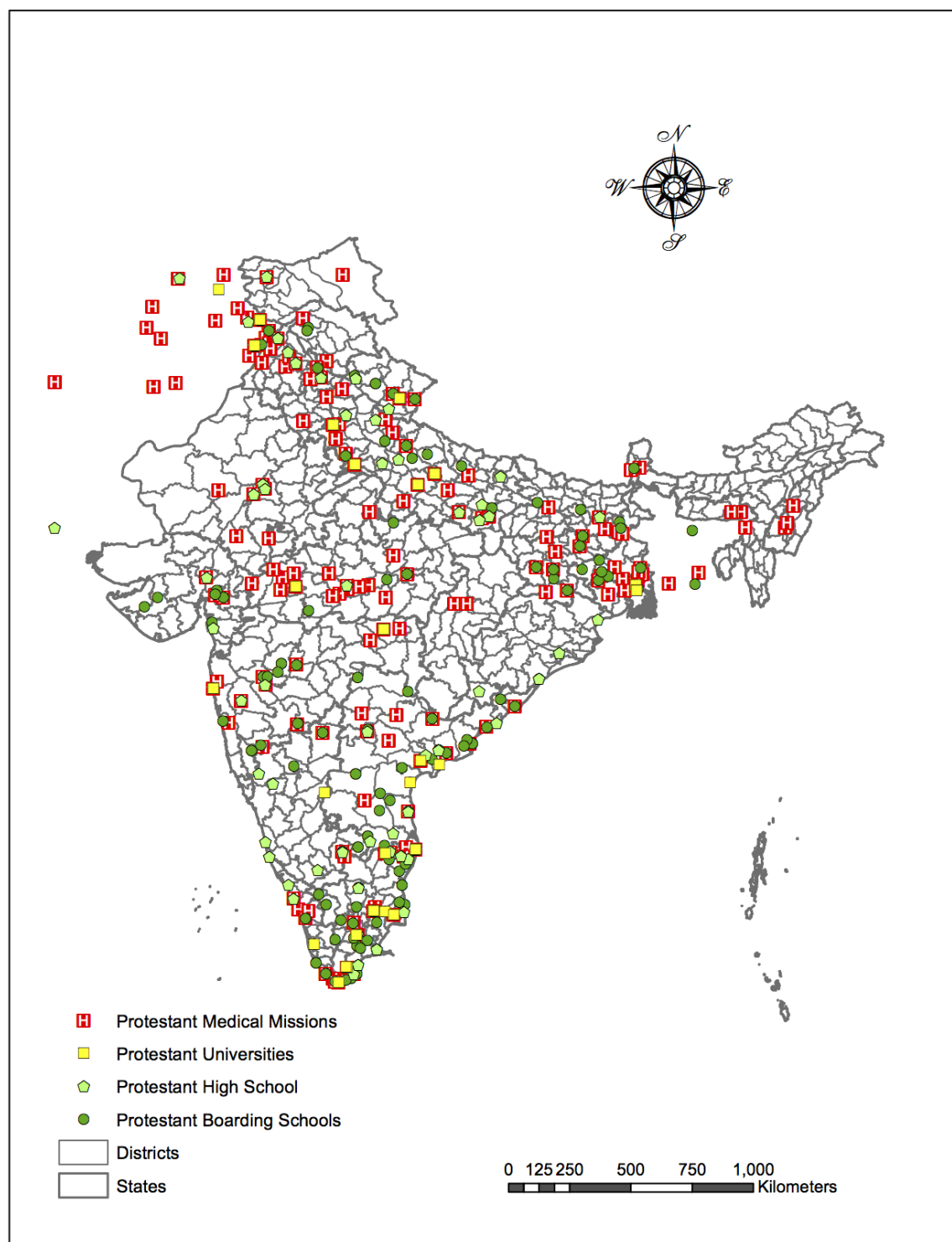
Figure A8: Protestant Medical Missions and Schools

Table A1: Missionary Societies and Hospital Shares

Society	Origin	*Society hospital share*
Canadian Presbyterian Mission	Canada	21.74%
Church of England Zenana Missionary Society	England	3.08%
Cambridge Mission to Delhi	England	0.00%
Church Missionary Society	England	8.87%
English Baptist Missionary Society	England	0.87%
English Presbyterian Church Mission	England	69.23%
Friends' Foreign Missions Association	England	22.22%
London Missionary Society	England	19.39%
Society for the Propagation of Gospel	England	1.60%
Society of St John the Evangelist	England	0.00%
Wesleyan Missionary Society	England	2.59%
Zenana, Bible and Medical Missions	England	0.00%
Basel Evangelical Missionary Society	Germany	7.14%
Independent	Independent	58.82%
Poona and Indian Village Mission	India	0.00%
Ranagat Medical Mission	India	0.00%
Bethel Santhal Mission	India	0.00%
Presbyterian Church of Ireland Missionary Society	Ireland	21.05%
Church of Scotland Mission	Scotland	21.05%
Edinburgh Medical Missionary Society	Scotland	66.67%
Free Church of Scotland	Scotland	26.83%
United Presbyterian Church of Scotland	Scotland	7.83%
American Board for Commissioners of Foreign Missions	United States	29.59%
American Baptist Missionary Union	United States	23.08%
Free Baptist Foreign Missionary Society	United States	0.00%
Methodist Episcopal Missionary Society	United States	18.66%
Presbyterian Board of Foreign Missions North	United States	42.34%
Woman's Union Missionary Society	United States	14.29%
Welsh Calvinist Methodist Foreign Missions Society	Wales	0.00%

Society hospital share is the ratio between each society's number of medical missions and total missions outside India. Only societies included in the estimation sample. Data from the Centennial Survey of Foreign Mission (1902).

Table A2: OLS Estimates (Log-Log Model)

	BMI (log)				
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)
Distance from Protestant medical mission (log)	-0.0119***	-0.0166***	-0.0110***	-0.0110***	-0.0110***
	(0.00240)	(0.00447)	(0.00416)	(0.00413)	(0.00416)
$\mathbb{1}$(Female)			-0.00653*	-0.00673*	-0.00673*
			(0.00358)	(0.00359)	(0.00387)
$\mathbb{1}$(Married)			0.00448	0.00467	0.00467
			(0.00472)	(0.00472)	(0.00436)
$\mathbb{1}$(Primary school completed)			0.0132***	0.0130***	0.0130***
			(0.00409)	(0.00410)	(0.00403)
$\mathbb{1}$(Major ethnic group)			-0.00561	-0.00446	-0.00446
			(0.00482)	(0.00482)	(0.00484)
Age			0.00723***	0.00720***	0.00720***
			(0.00102)	(0.00102)	(0.00103)
Age^2			-0.0000689***	-0.0000685***	-0.0000685***
			(0.0000129)	(0.0000129)	(0.0000130)
Household size			-0.00697**	-0.00703**	-0.00703**
			(0.00314)	(0.00314)	(0.00293)
Household size^2			0.000371*	0.000380*	0.000380**
			(0.000203)	(0.000203)	(0.000190)
$\mathbb{1}$(Wealth 2nd quartile)			0.0134***	0.0133***	0.0133***
			(0.00443)	(0.00443)	(0.00440)
$\mathbb{1}$(Wealth 3rd quartile)			0.0249***	0.0250***	0.0250***
			(0.00439)	(0.00437)	(0.00430)
$\mathbb{1}$(Wealth 4th quartile)			0.0497***	0.0499***	0.0499***
			(0.00507)	(0.00510)	(0.00483)
Population in 2000 (000s)				0.000358	0.000358
				(0.000249)	(0.000307)
Altitude (meters)				0.0000359*	0.0000359*
				(0.0000203)	(0.0000204)
No. rivers/water sources				-0.00136	-0.00136
				(0.00168)	(0.00169)
Latitude				0.00890	0.00890
				(0.0100)	(0.0101)
Longitude				-0.0110	-0.0110*
				(0.00700)	(0.00639)
Population in 1900 (000s)				-0.00000119	-0.00000119
				(0.0000397)	(0.0000639)
Cropland in 1900 (km^2)				0.0000555	0.0000555
				(0.0000825)	(0.0000862)
No. colonial railways				-0.0142*	-0.0142*
				(0.00810)	(0.00749)
N	7046	7046	7046	7046	7046
Mean Dependent Variable	3.05	3.05	3.05	3.05	3.05
District FE	No	Yes	Yes	Yes	Yes
Individual controls	No	No	Yes	Yes	Yes
Geographical controls	No	No	No	Yes	Yes
Historical controls	No	No	No	Yes	Yes
Spatial standard errors	No	No	No	No	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level in columns 1 to 4. Standard errors corrected for spatial correlation in column 5.

Table A3: Robustness Check: Estimation Sample

	Age	BMI			
	15-65	18.5-30	15-35	15-45	25-45
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)
Distance from Protestant medical mission (log)	-0.173* (0.105)	-0.229** (0.108)	-0.248* (0.129)	-0.239* (0.134)	0.325 (0.288)
N	7883	4966	7155	7174	699
Mean dependent variable	20.36	20.80	20.60	20.65	27.95
District FE	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes
Distance from Protestant mission (log)	Yes	Yes	Yes	Yes	Yes
Distance from Catholic mission (log)	Yes	Yes	Yes	Yes	Yes

In column 1, only individuals with BMI 15 to 30 are included in the estimation sample. In column 2 to 5, only individuals of age 20 to 60 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level.

Table A4: Robustness Check: Instrument

	Distance from Protestant medical mission (log)			BMI		
				10km (med. only)	30km (med. only)	50km (all miss.)
	OLS (1)	OLS (2)	OLS (3)	2SLS (4)	2SLS (5)	2SLS (6)
Hospital share 10 km (medical only)	-2.103*** (0.527)					
Hospital share 30 km (medical only)		-1.319*** (0.343)				
Hospital share 50 km (all missions)			-0.143*** (0.0430)			
Distance from Protestant medical mission (log)				-0.445* (0.265)	-0.584** (0.275)	-0.771* (0.406)
Kleibergen-Paap rk Wald F statistic	-	-	-	15.93	14.77	10.99
N	7046	7046	7046	7046	7046	7046
Mean dependent variable	4.02	4.02	4.02	20.42	20.42	20.42
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level.

Table A5: Robustness Check: Societies With and Without Medical Missionaries

	Non-Medical (1)	Medical (2)	Difference (3)
Date of Organization	1876	1857	19***
			(5.44)
No. Missionary Stations	44.50	220.34	-175.84***
			(-4.45)
Income from Home Sources per Station (US$)	1195.42	1346.60	-151.18
			(-0.64)
Income from the Foreign Field per Station (US$)	224.18	181.44	42.74
			(0.43)
No. Missionaries per Station	1.65	1.58	0.07
			(0.28)
No. Native Christians per Station	412.14	117.19	294.95
			(0.85)
No. Churches per Station	0.52	1.02	-0.50
			(-1.34)
No. Sunday Schools per Station	0.52	1.49	-0.97
			(-1.67)
N	203	92	295

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table compares societies operating worldwide in 1902 depending on whether they had medical missionaries. Societies compared are those also present in India. The number of missionary stations includes principals and all other substations.

Table A6: Robustness Check: Test of Overidentifying Restrictions

	Generated Instr. Only (1)	Generated and Excluded Instr. (2)
Distance from Protestant medical mission (log)	-0.4838*	-0.5759***
	(0.268)	(0.214)
N	7046	7046
Hansen J Statistics	16.3	16.4
Degrees of Freedom	18	19
P-value	0.573	0.632

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All variables are centered to remove district fixed effects. 19 additional instruments are generated using the residuals from the first stage regression (Lewbel (2012)).

Table A7: Additional Results: School Proximity

	BMI				
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)
Distance from Protestant medical mission (log)	-0.227* (0.120)	-0.193* (0.113)	-0.204* (0.107)		-0.216* (0.122)
Distance from Protestant boarding school (log)	0.0249 (0.110)			0.0280 (0.123)	0.0914 (0.141)
Distance from Protestant high school (log)		-0.0748 (0.110)		-0.123 (0.119)	-0.0883 (0.124)
Distance from Protestant university (log)			-0.0547 (0.128)	-0.0441 (0.173)	-0.0603 (0.167)
Distance from Protestant mission (log)	0.0353 (0.111)	0.0405 (0.112)	0.0322 (0.110)	-0.00421 (0.113)	0.0408 (0.114)
Distance from Catholic mission (log)	-0.134 (0.117)	-0.117 (0.115)	-0.124 (0.117)	-0.161 (0.114)	-0.120 (0.116)
N	7046	7046	7046	7046	7046
Mean Dependent Variable	20.42	20.42	20.42	20.42	20.42
District FE	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes
Distance from Protestant mission (log)	Yes	Yes	Yes	Yes	Yes
Distance from Catholic mission (log)	Yes	Yes	Yes	Yes	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level.

Table A8: Additional Results: Medical Mission Activities

	BMI			
	OLS (1)	OLS (2)	OLS (3)	OLS (4)
$\mathbb{1}$(No. treatments and patients $\geq$ median)	0.265 (0.178)	0.900 (0.564)		
Distance from Protestant medical mission (log)		-0.194* (0.108)		
Distance from Protestant medical mission (log) $\times$ $\mathbb{1}$(No. treatments and patients $\geq$ median)		-0.147 (0.135)		
$\mathbb{1}$(Surgery at closest Protestant medical mission)			0.377** (0.179)	0.645 (0.582)
Distance from Protestant medical mission (log)				-0.251** (0.112)
Distance from Protestant medical mission (log) $\times$ $\mathbb{1}$(Surgery at closest Protestant medical mission)				-0.046 (0.143)
N	7046	7046	7046	7046
Mean Dependent Variable	20.42	20.42	20.42	20.42
District FE	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes
Distance from Protestant mission (log)	Yes	Yes	Yes	Yes
Distance from Catholic mission (log)	Yes	Yes	Yes	Yes

Only individuals of age 20 to 60 and BMI 15 to 30 are included in the estimation sample. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the village (ward) level.

	Obs. (1)	Mean (2)	St. Dev. (3)
Panel A: Household Survey			
$\mathbb{1}$(Water Treatment)	485219	0.30	0.46
$\mathbb{1}$(Open Defecation)	485533	0.66	0.47
$\mathbb{1}$(Traditional Medicine)	458569	0.04	0.19
Distance from Protestant medical mission	465092	89.26	106.09
Distance from Protestant mission	485540	49.93	100.41
Distance from Catholic mission	485540	79.09	105.37
Household size	485540	5.36	2.62
Age of head of household	485537	46.43	14.39
$\mathbb{1}$(Hindu)	485539	0.77	0.42
$\mathbb{1}$(Primary school completed by head of hh)	287838	0.80	0.40
$\mathbb{1}$(Wealth 1st quartile)	485537	0.26	0.44
$\mathbb{1}$(Wealth 2nd quartile)	485537	0.27	0.44
$\mathbb{1}$(Wealth 3rd quartile)	485537	0.23	0.42
$\mathbb{1}$(Wealth 4th quartile)	485537	0.24	0.43
Village population (Census 2001)	485540	2299.22	4272.09
Population in 1900	485540	131.74	228.70
No. colonial railways	485540	0.35	0.86
Cropland in 1900 (km^2)	485540	35.91	32.91
Altitude (meters)	485540	385.67	533.22
No. rivers/water sources	485540	1.90	2.00
Latitude	465092	24.24	5.16
Longitude	465092	81.59	6.41
Panel B: Ever Married Women Survey			
Pneumonia Knowledge Index	441168	0.22	0.24
$\mathbb{1}$(Iron supplements during last pregnancy)	441168	0.81	0.39
$\mathbb{1}$(Full ANC during last pregnancy)	157503	0.13	0.34
$\mathbb{1}$(Safe delivery)	157488	0.42	0.49
$\mathbb{1}$(Child received check-up within 24 hrs from birth)	154556	0.40	0.49
Distance from Protestant medical mission	422332	87.85	100.11
Distance from Protestant mission	441168	49.95	94.53
Distance from Catholic mission	441168	78.05	99.11
Household size	441168	6.36	3.12
Woman's age	441168	31.45	8.62
$\mathbb{1}$(Hindu)	441166	0.78	0.41
$\mathbb{1}$(Primary school completed)	441123	0.38	0.49
$\mathbb{1}$(Wealth 1st quartile)	441168	0.28	0.45
$\mathbb{1}$(Wealth 2nd quartile)	441168	0.27	0.44
$\mathbb{1}$(Wealth 3rd quartile)	441168	0.21	0.40
$\mathbb{1}$(Wealth 4th quartile)	441168	0.24	0.43
Village population (Census 2001)	441168	2307.91	4212.86
Population in 1900	441168	137.47	233.89
No. colonial railways	441168	0.37	0.88
Cropland in 1900 (km^2)	441168	37.39	32.88
Altitude (meters)	441168	357.25	487.81
No. rivers/water sources	441168	1.87	1.97
Latitude	422332	24.26	5.05
Longitude	422332	81.29	6.20

DLHS-3 data. The wealth quartiles are based on a wealth index computed as the first principal component that combines information about a set of household assets: number of rooms in dwelling, and asset ownership, such as refrigerator, scooter, radio, and bike. Population, altitude, number of colonial railways and cropland data are extracted in 5 km-radius buffers around the DLHS village locations. Elevation data are available at the 30" × 30" grid cell; original population in year 1900 and cropland data are available at the 5' × 5' resolution. Latitude and longitude are in decimal degrees.

A.2 Long-Term Transmission and Health Production Function

To illustrate the three transmission channels, it may be useful to consider individual health as being determined by a set of inputs as follows:

$$H = \mathcal{H}(P, I, D) \tag{A1}$$

where P is health potential, I are health inputs and D are exogenous environmental conditions. These may depend on local observable characteristics, i.e. $D = \mathcal{D}(W)$. We define A_{SR} as current access to health care and A_{LR} as long-term access to health care, i.e. access to health care of previous generations. Since access to health care and health status of one generation can leave their mark on the health potential of the following cohorts (Bhalotra and Rawlings (2011), Coneus and Spiess (2012)), we consider individual health potential to be determined by A_{LR} and by socio-economic individual characteristics, i.e. $P = \mathcal{P}(A_{LR}, X)$. Moreover, health inputs may depend on current access to health care, health culture C, such as hygiene and health awareness and health promoting practices, and socio-economic individual characteristics, i.e. $I = \mathcal{I}(A_{SR}, C, X)$. In particular, health culture may be shaped by current and past access to health care: $C = \mathcal{C}(A_{SR}, A_{LR})$. On one side, proximity to health facilities today may directly affect health culture and stimulate the diffusion of healthy practices. On the other side, health and hygiene awareness may have been bequeathed by the family or the social environment and therefore depend on the access to health care of previous generations. An individual's health production function is therefore a function of long-term access to health care, access to health care today, socio-economic and demographic individual characteristics and local environmental conditions. Assuming a simple Cobb-Douglas production function and conditioning on observable characteristics at the individual and at the village level, individual health is produced as follows:

$$H = \mathcal{H}(A_{LR}, A_{SR}) = \kappa (A_{LR})^{\beta_1} (A_{SR})^{\beta_2} \tag{A2}$$

which we estimate empirically in section 7.1. We assume κ to be additively separable in district level and individual idiosyncratic components, i.e. $\kappa_{ivd} = \alpha_d + \epsilon_{ivd}$.