UNITED NATIONS
UNIVERSITY
UNU-WIDER

This paper is a draft submission to the

WIDER Development Conference

Human capital and growth

6-7 June 2016 Helsinki, Finland

Tertiary Education and Prosperity: Catholic Missionaries to Luminosity in India*

Amparo Castelló-Climent,† Latika Chaudhary‡ and Abhiroop Mukhopadhyay⋆

† *University of Valencia* ‡ *Naval Postgraduate School* ⋆ *Indian Statistical Institute*

March 2016

Abstract

This paper estimates the impact of tertiary education on economic prosperity across Indian districts. To address the endogeneity of tertiary education, we use the location of Catholic missionaries circa 1911 as an instrument. The choice of India is an important part of our identification strategy. Although Catholics have influenced elite education in India, they constitute a very small share of the population and their influence beyond education has been limited. Catholics were not at the forefront of tertiary education in colonial India, rather they established many high quality colleges after independence. This accounts for our strong first stage results. Controlling for a rich set of geographical and historical characteristics, our second stage finds a positive effect of tertiary education on development, as measured by light density at night. Our results are robust to different measures of development and are not driven by alternative channels through which missionaries could impact current income.

JEL classification: I25, N35,O15

Key words: Human Capital, Catholic Missionaries, Subregional Analysis

*We would like to thank participants at the Xth Annual Conference on Economic Growth and Development, XVIIth World Economic History Congress, 2015 Stanford Institute for Theoretical Economics Conference, ANAECO Workshop, 2016 Royal Economic Society Conference, and seminar participants at NYU-Stern, University of Arizona, UNSW(Sydney), Monash University (Melbourne), Delhi School of Economics and CMI (Bergen) for their positive feedback. We are grateful to Federico Mantovanelli for sharing the historical maps, and to Daron Acemoglu and Romain Wacziarg for comments. This work would not have been possible without the effort put in by Athisii Kayina. We acknowledge the financial support from the Spanish Ministry of Economy and Competitiveness through ECO2011-29283 project, and the Planning and Policy Research Unit at the Indian Statistical Institute (Delhi). The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

1 Introduction

Identifying the fundamental determinants of development has a long pedigree in economics. A large literature relying largely on cross-country variation has emphasized the role of institutions (Acemoglu et al., 2001, 2002, 2014), geography (Sachs 2003), openness to trade (Frankel and Romer 1999), and human capital (Lucas 1988). Among these factors, the empirical evidence linking education to income has produced perhaps the weakest findings at the macro level (Benhabib and Spiegel 1994). The lack of a robust relationship between education and income is at odds with the vast labor literature, which finds strong causal effects of each additional year of schooling on individual earnings (Card 2001). How do we reconcile the two sets of findings?

One explanation is perhaps that the macro literature has relied on incorrect measures of education. Most of the literature uses average years of schooling to capture education differences across countries (Benhabib and Spiegel 1994; Cohen and Soto 2007; de la Fuente and Doménech 2006). The large number of people with no education skew average years of schooling for poor countries. Moreover, primary education is often of poor quality in these countries (Chaudhury et al. 2006), further exacerbating measurement problems. This may explain why years of education correlate poorly with economic outcomes at the macro level (Pritchett 2001). Another mutually nonexclusive explanation is the focus on cross-country analysis. The vast differences in culture, institutions and access to technology make it difficult to identify the causal effect of education on income. Problems of omitted variables and reverse causality plague many of the empirical studies (Acemoglu, Gallego and Robinson, 2014).

In this paper, we study the impact of human capital on development using data on Indian districts in 2006. Our focus on districts, an administrative unit below states, allows for tighter comparisons because these sub-national units at least share common governance and national policies thereby reducing some concerns related to omitted variables correlated with education and development. This strategy also allows us to address Acemoglu et al.'s (2014) concerns that empirical studies treating institutions and human capital as exogenous are misspecified. Since institutions are mainly determined at the central and state level, the advantage of our approach is we can account for institutions without the need of an additional instrument.

However, district-level data pose one problem in the Indian context - current income levels are not well measured. We address this shortcoming by using night lights data as a proxy for income, in line with the recent literature (Henderson, Storeygard and Weil 2012;

Micholapoulos and Papaioannou 2013, 2014; Alesina, Michalopoulous and Papaioannou 2015). We rely on information collected by the National Geophysical Data Center (NGDC) on the location of night lights between 8pm and 10pm, as captured by satellites of the United States Air Force Defense Meteorological Satellite Program (DMSP). Observations are available for an area of one squared kilometer and can be aggregated to the district-level. To measure human capital, we focus on the share of the adult population with tertiary education as only higher levels of education appear to be correlated with economic growth in India (Castelló-Climent and Mukhopadhyay 2013).

We find a strong positive association between the share of the population with tertiary education and light density at night in OLS models that control for state fixed effects. This relationship is robust to a rich array of factors that may jointly influence tertiary education and luminosity such as current population, population shares of socially disadvantaged groups, geography, and historical variables that control for initial conditions. The potential endogeneity of tertiary education, however, poses an empirical challenge because tertiary education and the evolution of income generally go hand in hand. We address this concern by using the location of Catholic missionaries in the early 20^{th} century as an instrument for current tertiary education. Using the first edition of the *Atlas Hierarchicus*, we extract the precise location of Catholic missionaries in 1911 and overlay the historical maps on district borders as of 2001 using a geographic information system (GIS) program.

The historical location of Catholic missionaries has to satisfy two conditions to be a valid instrument. First, the location of Catholic missionaries in 1911 has to be correlated with current tertiary education. Supporting this condition, we document a large, positive and statistically significant impact of Catholic missionaries circa 1911 on the share of the 2001 adult population with tertiary education. The tertiary education share is 1.2 percentage points higher in districts with a Catholic mission. This is a substantial effect on the order of 20% given mean tertiary education (5.8%).

Catholic missionaries were not heavily involved in providing education in the colonial period. For example, there were only 9 Catholic colleges (5%) as of 1911. Rather the mechanism linking Catholic missionaries to current tertiary education appears to operate from the establishment of several Catholic colleges after Indian independence. The historical network of Catholic missionaries was a natural platform from which Catholic influence on education radiated out. In support, we find a large and positive correlation between missionary location in 1911 and the stock of Catholic colleges in the beginning of the 21^{st}century. Qualitative evidence suggests stronger funding from the Vatican, better co-ordination among the different Catholic groups and Indian independence in 1947 that

guaranteed protection to minority institutions in the new constitution were collectively responsible for the increase.[1]

Second, our historical location instrument has to satisfy the exclusion restriction, i.e., the location of Catholic missionaries circa 1911 has to be unrelated to any factor that may impact the subsequent development of districts other than through current tertiary education. If Catholic missionaries historically located in richer and more educated districts, then our instrument would be invalid. However, this does not seem to be the case. Using historical data for a sample of districts, we find no significant correlation between the location of Catholic missionaries in 1911 and the historical provision of education or measures of wealth such as income-tax revenues. The key historical determinants of Catholic missionary location are proximity to the coast, proximity to railways and districts with a large share of tribal groups. As we describe in section 4, these quantitative results match historical accounts of Catholic missionaries locating in Portuguese settlements along the coast and in districts with large tribal populations.

In our instrumental variables regression, we include a rich set of geographical and historical controls to address any potential selection concerns regarding the location of Catholic missionaries. Our geographic variables such as latitude, longitude, average elevation, average river length, and a coastal indicator control for any positive selection in the choice of location vis-à-vis geography. We also include the historical share of Brahmins, tribal groups and the urban population share in addition to a dummy if a railway passed through the district in 1909. These variables control for selection related to historical access, social structure and urbanization. Finally, we include an indicator for districts that were historically part of Princely India as opposed to under direct British rule (British India) because Iyer (2010) shows that districts under Princely rule have better economic outcomes after independence. Our assumption is that conditional on including these extensive controls derived from a reading and analysis of the history, the location of Catholic missionaries is not confounded with other characteristics that may independently impact current development other than through current tertiary education.

Using Catholic missionaries as an instrument for current tertiary education, the second stage results find a positive and statistically significant effect of tertiary education on

[1] Article 30(1) of the Constitution of India gives linguistic and religious minorities a fundamental right to establish and administer educational institutions of their choice. The end of British colonization also marked a watershed for the growth of Catholic educational institutions as the British government followed an active strategy of discouraging missionary influence on the provision of education in India. Moreover, an early emphasis on Indianizing the clergy generated a large pool of Indian priests and nuns giving the Catholic church a unique advantage in post-independence India (Frykenberg 2008).

current income proxied by light density at night. A one standard deviation increase in the share of the population with tertiary education increases log light density by 0.38, an economic effect of 9% given mean log light density of 4.24. Although we find significant and positive IV estimates on tertiary education, one may still concerned about Catholic missionaries impacting current income via non-education channels. This would also violate our exclusion restriction.

The existing literature has suggested many such non-education channels. Studies have found that the historical presence of missionaries can influence current religious beliefs and values (Nunn 2010). Acemoglu et al. (2014) highlight the interrelationship between human capital and institutions arguing that measures of human capital can capture the effect of institutions if some measure of institutions is not directly included in the analysis. Bai and Kung (2015) argue that the presence of Europeans promoted the diffusion of knowledge and hence economic prosperity in China. In the case of India, Calvi and Mantovanelli (2014) find that proximity to a historical Protestant medical mission is correlated with current health outcomes. To ensure our results are not driven by these alternative mechanisms, we estimate specifications that control for measures of health, infrastructure, the current religious composition of the district, the share of migrants, and the historical presence of Europeans. The coefficient on tertiary education remains positive, statistically significant and stable across these specifications.

We conduct several tests that indicate the model does not suffer from a weak instruments problem and is not under-identified. Robustness checks also confirm that the coefficient on the share of tertiary education is not picking up the effect of other levels of schooling such as low primary education. Finally, we use alternative proxies for development such as consumption per-capita and find similar results.

Our paper contributes to four different literatures. First, a large literature has studied the relationship between human capital, development and economic growth. Early studies by Barro (1991) and Mankiw, Romer and Weil (1992) found a significant and positive effect of human capital on income and growth. However, subsequent studies by Benhabib and Spiegel (1994) and Pritchett (2001) found insignificant or negative coefficients in specifications where human capital was measured in growth rates. To reconcile the findings, Cohen and Soto (2007) and De la Fuente and Doménech (2006) argued that poor data quality could account for the weak effect of human capital on growth.[2] If the stock of human capital is measured with error, their first differences will bias the estimated coefficient

[2]Krueger and Lindahl (2001) have also pointed to measurement error and the lack of signal in education data as an explanation for the findings in Benhabib and Spiegel (1994) and Pritchett (2001).

towards zero. Other studies have also raised concerns of reverse causality because higher income and faster growth rates generate more resources to invest in education (Bils and Klenow 2000). Motivated by the conflicting evidence and by the possibility that richer places can afford more education, our paper is one of the few attempts to convincingly address such concerns. On the measurement side, we focus on a specific level of education, namely the share of the population with at least a college degree. To address endogeneity problems, we exploit the location of Catholic missions in the early 20^{th} century as an exogenous source of variation for current tertiary education in India.

Second, a small and growing literature has begun to study the roots of development from a sub-national perspective. For example, Acemoglu and Dell (2010) examine sub-national variation in a sample of countries in the Americas and show that differences across regions within the same country are even larger than differences in income across countries. They find that about half the between-country and between-municipality differences in labor income can be accounted for by differences in human capital. In a similar vein, Gennaioli et al. (2013) find that human capital is one of the most important determinants of regional GDP per-capita in a large sample of regions covering 110 countries. Although the paper partially addresses endogeneity concerns using panel data techniques, their results cannot be interpreted as causal. The major contribution of our paper is the identification of a causal effect of tertiary education on development, using historical data as an exogenous source of variation in current human capital.

Third, our paper contributes to a burgeoning literature on how history, in particular colonization, influences current outcomes. Acemoglu et al. (2001, 2002) argue that colonies with a more favorable disease environment encouraged more settlement of European colonizers and promoted institutions protecting private property rights. Sokoloff and Engerman (2000) also focus on historical institutional development, but argue that factor endowments shaped institutions. Areas predisposed to sugarcane production saw larger imports of slaves, establishment of slave plantations and more unequal institutions, whereas areas with higher land-labor ratios and small farms lead to egalitarian economic and political institutions. In contrast, Glaeser et al. (2004) suggest that European settlers brought with them their own human capital and not institutions per se. Easterly and Levine (2012) compute a new measure of the share of European population during the early stages of colonization. Their findings are in line with the Glaeser et al. (2004) view.[3] By focusing on the intervention of Catholic missionaries in former colonies, we iso-

[3]Other scholars stress the genetic distance relative to the world technological frontier (Spolaore and Wacziarg 2009) and the genetic diversity within populations (Ashraf and Galor 2013). See Spolaore and

late a specific historical channel of human capital transmission. Unlike Banerjee and Iyer (2005) and Iyer (2010) that study how differences in formal colonial institutions impact current outcomes in India, our focus is on Catholic missionaries and how their historical location impacts human capital today. In our case, history influences the present through the location of Catholic missionaries who were absent from the provision of education in the past but played a bigger role after Indian independence.

Fourth and finally, our study contributes to the growing literature on religion and human capital. Much of this literature focuses on the positive impact of Protestants on education. For example, Becker and Woessmann (2009) find that Protestants had a strong effect on literacy in 19^{th} century Prussia. Mantovanelli (2014) argues that Protestant missionaries can account for current differences in literacy across India. Nunn (2014) compares Protestant to Catholic missionary activity in Africa and finds that both had a long-term positive impact on education. But, the impact of Protestant missionaries is stronger for women while the impact of Catholic missionaries is stronger for men.[4] Studying Africa again, Gallego and Woodberry (2010) find that Protestant missionaries had a larger impact on long-term education than Catholics, but mainly in states where Catholic missionaries were protected from competition by Protestant missionaries.

Our paper offers a fresh perspective to the literature. First, we find that only historic Catholic missionaries are correlated with current tertiary education in India and not Protestant missionaries. The impact of Catholic missionaries on the supply of tertiary education increased over the second half of the 20^{th} century as Catholics established colleges across India. Catholic emphasis on tertiary education is perhaps unsurprising and matches accounts of Catholics, especially the Jesuits, leading the growth of tertiary education in other parts of the world (Codina 2000). Second, most studies that analyze the long-term consequences of Christianity have focused on Africa and South America. Unlike these countries, the Christian population is a minority in India, with only a 2 percent share identifying as Catholic. Whereas Christianity has played a critical role in shaping a part of education, especially the influence of Catholics on elite education, their influence beyond education has been limited. Thus, looking at India is a good identification strategy for analyzing the impact of education on development.

The structure of the paper is as follows. The next section describes the data. In section 3 we present the OLS results and discuss potential biases. Section 4 discusses

Wacziarg (2013) for an excellent survey of the literature.

[4]Becker and Woessmann (2008) examine village-level data from the Prussian Census of 1816 and identify a negative relationship between the prevalence of the Protestant religion and the educational gender gap, measured as average male education minus average female education.

the instrumental variables strategy. Section 5 presents the main IV results. We describe several robustness checks in section 6, and conclude in section 7.

2 Data

Our analysis is conducted at the district-level, an administrative unit in India analogous to a US county. Empirical analyses that use historical data (or panel data) for India usually work with 13-16 major Indian states (of 1991 census-year vintage). The common practice in all such papers is to drop small states (like Delhi) and the extreme north-eastern part of India.[5] Analogous to previous work, our analysis is restricted to 500 districts in 20 states of India (of 2001 census-year vintage) that cover more or less the same area as covered by other studies.[6]

While district-level data allow us to conduct sub-regional analyses, the main shortcoming in the Indian context is that statistical agencies do not report district-level GDP. To address this issue, we rely on night lights data. Recent work by Henderson et al. (2012) and Pinkovskiy (2013) suggest luminosity is a good proxy for income.[7] The data on night-light luminosity is recorded worldwide for every pixel by the Operational Linescan System (OLS) flown on the Defense Meteorological Satellite Program (DMSP) satellites. The data is available online from the US National Oceanic and Atmospheric Administration (NOAA). Following Henderson et al. (2012) and Michalopoulos and Papaioannou (2013), we use satellite images on light density at night as a proxy for economic development. We aggregate 2006 luminosity across all pixels within 2001 district boundaries. Then, we divide total luminosity by the area of the district to calculate light density at night.[8]

[5] In the case of north-eastern India, this is largely to account for the poor quality of current data and problems of mapping historic boundaries to current boundaries.

[6] The number of states is dependent on the data available for the question being asked. For example, Besley and Burgess (2000) use 16 major states of India, whereas Banerjee and Iyer (2005) use district-level data from 13 major states. The states we study are Andhra Pradesh, Assam, Bihar, Chattisgarh, Gujarat, Harayana, Himachal Pradesh, Jammu and Kashmir, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, Uttaranchal and West Bengal. The larger number of states as compared to the cited studies reflect the bifurcation of states between 1991 and 2001.

[7] Chen and Nordhaus (2010) note some problems with satellite image data but argue that luminosity is still useful for regional analysis especially when income data are poor.

[8] There are two kinds of luminosity data available from NOAA. The datasets that are available at an annual frequency measure luminosity from 0 to 63, with 0 measuring no light. However, these suffer from saturation problems: that is, they do not reflect luminosity differences beyond the top value of 63 (as would be the case with growth in cities). The other data set, classified as calibrated radi-

As is standard in the literature, we calculate the log of this measure. This measure varies from a minimum of -0.953 to a maximum of 6.407 with a mean of 4.24. *Figure* 1 illustrates the night lights map and district-level luminosity side-by-side. In the district map, lighter colors correspond to higher luminosity. There is tremendous heterogeneity in luminosity across Indian regions such as between the South (high luminosity) and East (low luminosity). Nevertheless, there is also heterogeneity between contiguous districts. Our analysis explores whether tertiary education can account for differences in luminosity across districts within the same state.

In the regressions we focus on the population aged 25 years and above ensuring the completion of tertiary education is not censored by age. Using the 2001 census of India, we construct current district-level demographic and education variables.[9] The main independent variable of interest is the share of population 25 years and above who have completed tertiary education.[10] Although 5.8% of the Indian population over 25 has completed any tertiary education, the district-level percentages range from 1.4% to 21.3%. *Figure* 2 shows the spatial distribution of tertiary education across Indian districts. Analogous to the figure on luminosity, a lighter color represents a higher share of adults with completed tertiary education. While it is clear that South India has higher tertiary completion rates, again there is significant heterogeneity within states.

Tertiary education flourished in British India despite a low and stagnant level of literacy (just over 10% in 1931). Enrollment rates in arts and professional colleges increased six fold between 1891 and 1941 from 0.05% to 0.35% (Chaudhary 2015). In comparison, primary school enrollment in 1941 was only twice as much as in 1891 with one-third of school-age children attending any primary school in 1941. Most of the increase in tertiary education was driven by private demand because administrative positions in colonial government offices often required a college degree. Unlike the recent increase, this early development in tertiary education occurred in the liberal arts and not in technical degrees. After Indian independence in 1947, the policy focus switched to increasing and improving the number of

ance data allows measurement of luminosity till 255 and is less susceptible to the saturation criticism. However, it is sporadically available and hence less used in research that focuses on growth of luminosity. Since this paper uses cross sectional data, we are able to use the calibrated radiance data set. We use GIS tools to extract luminosity from the raster files provided by DMSP. Data are available on http://ngdc.noaa.gov/eog/dmsp/download_radcal.html

[9]The 2011 census of India was recently released. However, we take the current explanatory variables from the census in 2001 as the radiance night light data is not available beyond 2010.

[10]Our definition include individuals with degrees or diplomas in general education and professional education. There is also a provision to get a diploma after class X (grade 10), but these are usually at the level of higher secondary school. We do not include these individuals in our definition of tertiary education.

people with technical degrees. Unfortunately, the census enumeration of education at the district-level has evolved over time making it difficult to follow the share of the population with completed tertiary education. The best we can do is follow a consistent definition of graduate degree holders and above, which indicate the share of graduate degree holders increased from 0.5% of the population in 1971 to 3.1% in 2001.

The unconditional correlation between district log light density and share of 2001 tertiary education is 0.46. *Figure* 3 shows the scatter plot between the two variables. While the pictures suggest a large and positive correlation between luminosity and tertiary education, the correlation could be driven by confounding variables correlated with the two. To address this concern, we control for a broad array of factors that include current controls, geographic characteristics, and historical variables.

The list of current controls include the district population aged 25 years and above as well as the share of historically disadvantaged groups 25 years and older, referred to as Scheduled Castes and Scheduled Tribes under the Indian constitution. These data are taken from the 2001 Census. The list of current variables is necessarily parsimonious, a point we discuss below.

To capture geographic characteristics, we construct an indicator variable for districts with any coastal boundary, the latitude and longitude of the centroid of the district, the average length of rivers that pass through the district, the average elevation of the district, and the minimum distance from the centroid of the district to one of the million plus population cities of India.[11] The geographical variables are constructed using GIS tools. We include these variables to account for any direct impact of geography on development (Sachs 2003) and also to correct for any systematic bias that geography may cause in measuring night lights.

We rely on variables from the historical 1931 census of India to capture differences across districts in initial states that may impact current tertiary education and income. We focus on several historical variables: the urban population share in 1931, the tribal population share in 1931, the population share of Brahmins in 1931, the districts that were historically a part of Princely India, and the presence of railways in 1909.[12] Brahmins

[11]The million plus cities as of the 2001 census are Ahmedabad, Bangalore, Chennai, Delhi, Hyderabad, Jaipur, Mumbai, Pune and Surat. The files to extract the average elevation and the average river length were obtained from http://www.diva-gis.org/

[12]We use the methodology of Bharadwaj et al. (2008) and Kumar and Somanathan (2009) to match the 1931 district boundaries to 2001 district boundaries. On account of changes in these boundaries, we impute the same historical proportion to all districts in 2001 that are contained in consistent geographic units between 1931 to 2001. In few cases, these units cover a large number of districts because of many

typically occupy the top position in the Indian caste system. Although they traditionally worked as priests and teachers, Brahmins were disproportionately represented among landowners, lawyers and other elite occupations in the colonial era. Thus, the Brahmin population share may independently influence both subsequent tertiary education and development. We also create an indicator variable for districts that were historically a part of Princely India, i.e., under the direct control of hereditary rulers in the colonial period as opposed to under direct British rule (i.e., British India). The native rulers faced different incentives that may impact the subsequent development of education and income. Finally, we use the railway map in the Administration Report on Indian Railways for 1909 to construct an indicator for the presence of railways, as more accessible places are likely to have more educated people and can influence development per se. We report summary statistics for the main variables in *Table* 1.

3 Tertiary Education and Luminosity: OLS Estimates

We begin by estimating an OLS model using the share of the population over 25 who have completed tertiary education as our key independent variable (*tertiary*) and the log of light density at night (y) as our measure of development. The empirical model we estimate is:

$$y_{ds} = \alpha_s + \theta tertiary_{ds} + \rho' C_{ds} + \pi' G_{ds} + \delta' H_{ds} + \varepsilon_{ds} \quad (1)$$

where d denotes a district and s stands for state. We eliminate the impact of omitted variables that vary at the state level by allowing a state specific intercept term α_s, namely dummy variables for each state. The within-state comparison removes the effects of state-level policies (both current and past) that covary with *tertiary* as well as state-level omitted variables. The use of within-state variation, in contrast to inter-state variation, also eliminates cultural differences towards education and development. For instance, the differences in human capital between North and South India are often ascribed to differences in culture. We account for other observed differences by including the vector C of current variables, namely the population over 25 and the population share of Scheduled Castes and Tribes. We also take into account a vast empirical literature that has documented a strong correlation between geographical characteristics and current development.[13] We model y as a function of time-invariant characteristics with the vector

district boundary changes over time.

[13]Whether geographical factors have a direct impact on contemporaneous development or indirectly through its persistent historical effects is still under debate. See Nunn (2014) and Spolaore and Wacziarg

of geographical controls, G.

The list of current controls is parsimonious by intention. We exclude most current variables because they are likely endogenous to factors influencing contemporary luminosity. Instead, we take into account different initial conditions across districts through a set of historical controls, H, described in the previous section. These variables account for the evolution of other contemporaneous variables, for example, current urbanization, that are not included in our specification due to potential endogeneity.[14] We estimate the model using robust standard errors.[15]

Table 2 reports the results for the base-line OLS specification. A comparison across columns shows the impact of controlling for other covariates on the coefficient for tertiary education. Unsurprisingly, the addition of contemporaneous and geographic controls reduces the magnitude of the coefficient on tertiary education. The marginal effect of increasing the share of tertiary education by one standard deviation (0.03) in the model with no controls (*column* (1)) increases the log light density by 0.43 (mean value is 4.24). The marginal effect drops to 0.30 once we include all controls (*column* (4)). This is an economically significant effect on the order of 7% against the mean of log light density. Results also indicate that geographical characteristics, the share of tribal groups, and the presence of a railway line in 1909 are significantly related to current income levels.

One concern with using only the share of tertiary education is that the omitted category is the population share that is illiterate, has not completed primary education or has completed only up to secondary schooling. The share of tertiary education may differ across districts because of differences in primary schooling or illiteracy, apart from any differences in tertiary education. Hence, the coefficient on tertiary education may be picking up the effect of other levels of schooling. To address this concern, we directly control for the population share with some schooling but no higher than secondary education (*column* (5)). For ease of exposition, we refer to this variable as share of secondary schooling in the tables and text. The coefficient on tertiary education remains positive and statistically significant at the 1 percent level. As expected, the coefficient is smaller, but it is nonetheless economically significant. This suggests the findings on tertiary education are not driven by omitting other levels of schooling.

(2013) for excellent surveys on the literature.

[14] In Section 6 we analyze the robustness of the results to additional contemporaneous controls.

[15] We do not cluster standard errors at the state level because the number of states (17) are too few to generate accurate clustered standard errors (Angrist and Pischke 2009). As a robustness check we estimated p-values from wild bootstrap suggested by Cameron, Gellbach and Miller (2008) as a solution to the problem of few clusters. Our results on tertiary education are still significant.

Our results thus far find a strong correlation between the population share with tertiary education and log light density at night. This finding is robust to a broad array of controls, and is not driven by lower levels of schooling. However, this relationship cannot be interpreted as causal. The main challenge to ascribing a causal interpretation is that the share of individuals with tertiary education is likely to be endogenous. Reverse causality is a serious concern if individuals with tertiary education move to districts with higher income or higher light density. Although migration is low between Indian states (Munshi and Rozensweig, 2009), people are more mobile within states. Moreover, less is known about the migration of tertiary-educated labour within India. In order to obtain consistent estimates, we therefore need to address the issue of endogeneity. To this end, we turn to history and the role of Catholic missionaries.

4 Catholic Missionaries to Luminosity

In this section we first describe the history of Catholic missionaries and then discuss why the location of missionaries is a good instrument for 2001 tertiary education.

4.1 Catholic Missionaries

According to popular accounts, the apostle St. Thomas travelled to South India in the 1^{st} century A.D. (CSMC 1923). While it is unclear if the visit impacted the local population, an ancient group of Indian Christians (the St. Thomas Christians) with roots predating the arrival of Europeans believe they spiritually descended from the apostle St. Thomas. This oldest group of Catholics, the St. Thomas Christians, are largely concentrated in Kerala.[16]

Barring an occasional mention of Catholic priests, church history in India is silent till the arrival of Vasco de Gama and the Portuguese in Calicut in 1498. Along with their trading interests, the Portuguese had a strong desire to spread Catholicism in their overseas colonies (Richter 1908). To this end, Portuguese rulers enjoyed special ecclesiastical privileges, the *Padroado Real*. Granted by the Pope in 1452 and 1455, these charters gave the Portuguese Crown "exclusive authority to fill clerical positions within overseas domains" (Frykenberg 2008, p. 127). Rome believed that allowing the Portuguese Crown

[16]St. Thomas Christians, also known as Syrian Christians, had many historical disputes with the Catholic Church for example, the language of liturgy, Syrian or Latin, and promoting native clergy. Rome addressed these concerns in the late 19^{th} century. Since then the two churches of St. Thomas Christians, the Syro-Malabar Church and Syro-Malankara Church, have become an important part of the Catholic Church in India. See Frykenberg (2008) for details.

to appoint bishops and collect church taxes in exchange for establishing churches and missionaries was a low-risk high-return strategy, though subsequent Popes came to regret granting such extensive privileges to foreign monarchs.

Under *Padroado Real*, the first Catholic missions were set up in India in the 16^{th} century. The Franciscans and Dominicans were dominant early on but were taken over by the Jesuits after the arrival of Francis Xavier, co-founder of the Society of Jesus, in 1542. Missionary efforts were concentrated on the western coast and Goa become the center of Portuguese Catholic Church hierarchy. Missionaries also settled early in Portuguese strongholds such as Daman, Diu, Vasai (suburban area north of Mumbai), and Mumbai along the coast. Corroborating these accounts 78% of the Catholic missionaries we observe in 1911 are located in coastal districts.

While the early missionaries followed Portuguese conquest, missionaries also settled in the interior away from Portuguese strongholds. Our reading suggests the individual preferences of missionaries played a role. For example, an enterprising Jesuit named Robert de Nobili moved to the cultural city of Madurai, pretended to be an upper caste Hindu and established the Madurai Mission to recruit high-caste Brahmins into the Catholic fold. There are accounts of Jesuit missionaries following Nobili's methods, as well as non-Jesuit missionaries working to convert lower castes and tribes (CSMC 1923). Akbar, the Mughal Emperor of India in this period, allowed Catholics to set up missions in Gujarat but the missionaries chose the location. Historical centers of trade and production were not always the natural choice.

As Portuguese rule declined over the 17^{th} century, Catholic missions fell into disarray. While they had enjoyed Portuguese political patronage, neither the native Indian rulers nor the English East India Company were sympathetic to the Catholic cause. The suppression of the Jesuits in the 18^{th} century compounded the problem because they were the most active Catholic missionaries in the field. Finally, ecclesiastical disputes between the *Propaganda Fide*, a society backed by Rome, and the *Padroado Real* backed by Portugal made it difficult for all Catholic missionaries. These differences were resolved by Pope Leo XIII in 1885 with a charter that established the Indian Catholic Church. Most jurisdictions were placed directly under Rome, or the *Propaganda Fide* barring two, the Archdiocese of Goa and the Diocese of Mylapore, near modern-day Chennai.

After the establishment of the Hierarchy, the Catholic Church embarked on a extensive program of education in 20^{th} century India.[17] The main goals were to develop a dedicated

[17]The Hierarchy here refers to the officials, commonly bishops, who command authority within the Catholic Church.

native Indian clergy and high quality schools and colleges. Again the Jesuits lead the charge after their return in the 19^{th} century. India today has 3,851 Jesuit priests, more than any other country, despite the fact that less than 2% of the population is Catholic (Frykenberg 2008). These trained priests have contributed to the rise of many high-quality Jesuit schools and colleges that are among the best in the country. The historical network of Catholic missionaries was a natural platform from which Catholic influence on education spread. Corroborating the qualitative accounts, we find the large expansion in Catholic colleges occurred only after Indian Independence. In section 6, we document the growth in Catholic colleges over time, and show that missionary location in 1911 has a strong positive correlation with the stock of Catholic colleges in 2001.

4.2 Identification

Our empirical exercise uses the location of Catholic missionaries in the early 20^{th} century as an instrument for 2001 tertiary education.[18] We obtained the location of Catholic missionaries from a map published in the first edition of *Atlas Hierarchicus*, which marks the name of every place in India where there was a Catholic mission or missionary in 1911 (*Figure* 4).[19] This historical map is super-imposed on the 2001 district map of India to yield the location of Catholic missionaries in terms of 2001 district boundaries. *Figure* 5 displays Catholic locations after this exercise. As noted above, a majority of the missionaries are located along the coast in former Portuguese colonies. But, a sizable number are present inland with more in peninsula India than in the North or the East. We use the location of Catholic missionaries to construct an indicator variable taking the value 1 if, in 1911, there was a Catholic missionary in the area covered by district d in 2001. We believe a simple indicator is more exogenous compared to an intensive measure of the number of Catholic institutions or say number of Catholic institutions per square kilometer, which is likely to be correlated with the historical presence of Christians in a district.[20]

[18] Although we do not know the date when Catholic missionaries first arrived in a district, we believe it was either decades before or immediately after the establishment of the Catholic hierarchy in 1885. Not many Catholic missions were established in the early to mid-19^{th} century and Catholic missionaries were mostly dormant in this period. The older Portuguese missions were set up in the 16^{th} and 17^{th} centuries.

[19] The maps in the *Atlas Hierarchicus*, by Karl Streit (1913), are based on the latest edition of Sohr-Berghaus' Handatlas (1902-1906). The purpose of the Atlas was to gain insight on the situation, hierarchy and delineation of the territorial division of the Catholic Church in the whole world.

[20] Our IV results are robust to using such intensive measures and to controlling for the historical share of Christians. These results are available upon request.

The presence of Catholic missionaries has to satisfy two conditions to be a valid instrument. First, the instrument has to be correlated with tertiary education in 2001. In the next section we provide compelling evidence that districts with a historical presence of Catholic missionaries are significantly correlated with tertiary education in 2001. Further, conventional weak instrument tests, such as the F-statistic and the Cragg-Donald Wald statistic, indicate the instrument is strong. Second, the location of Catholic missionaries as of 1911 has to be uncorrelated with the error term, i.e., unobservable factors that may influence current light density. The location of Catholic missionaries can violate the exclusion restriction in two ways. First, Catholic missionaries may have located in richer or more educated areas. Second, the exclusion restriction can be violated if the location of Catholic missionaries impacts contemporary luminosity through channels other than contemporary tertiary education. In this section we analyze the endogenous location of Catholic missionaries. In Section 6 we examine other potential channels through which Catholic missionaries could affect current development.

Catholic missionaries could have located in richer or more educated districts that could independently impact current development. The main concern here is whether Catholics positively selected districts. To assess the potential exogeneity of Catholic missionaries, we collected information on a sub-sample of districts for the period before our map of Catholic locations was published in 1911. This smaller sample of districts covers the British Indian provinces of Bengal, Bihar and Orissa, Bombay and Madras where we have decent data on education and measures of income. In $Table\ 3A$ we regress the presence of a Catholic missionary in a district in 1911 on geography and other historical variables from 1901. We show the effects of geography on the entire sample of 500 districts ($Column$ (1)), and then for the sub-sample of districts where we have historical information (($Columns$ (2)-(3)). The impact of geography is similar across the two samples, which suggests we can draw cautious conclusions about the full-sample based on this selected sample.

Catholic missionaries were more likely to be located in coastal districts, in districts with a railway presence and in districts with a larger share of tribals. While coastal districts and railways indicate positive selection, tribal districts are indicative of negative selection. Reassuringly, we find the location of Catholic missionaries is uncorrelated with the share of Brahmins and income tax revenues per capita, a proxy for income. We also find that the number of schools are not correlated with the presence of catholic missionaries. In fact, the selection is negative; the coefficient on the number of colleges is significant but negative. This is consistent with the historical record that Catholic missionaries were less involved in education relative to Protestant missionaries. For example, there were only 9

Catholic colleges (5%) in 1911 compared to 40 Protestant (33%).[21] The rest were either public or under private Indian management. Public schools were more widespread at the primary and secondary level, but again Catholics were largely absent. These correlations are suggestive that the location of Catholic missionaries was not systematically correlated with historical measures of income or education. Moreover, the fact that missionaries located in coastal areas and in districts with a higher share of tribal groups is an indication that Catholic missionaries did not follow a uniform strategy.

Most studies on Christian missionaries focus on Protestants. In keeping with that literature, we also constructed an indicator variable for districts with a Protestant mission as of 1908 using information in the Statistical Atlas of Christian Missions (1910). We observe Protestant missionaries in 58% of the districts. Their location is closely tied to the strength of the East India Company and British Crown (Richter 1908). Hence, we worry more about the potential exogeneity in their choice of location. Interestingly, Protestants were actively involved in education in the colonial era, but lost their dominance after Independence. Despite their alleged stance of religious neutrality, the British Crown was more favorable to Protestant missionaries compared to Catholics. Hence, the loss of informal state patronage probably hurt the Protestants in post-1947 India.

Similar to *Table* 3*A*, we ran regressions on the location of Protestant missionaries and historical characteristics reported in *Table* 3*B*. In the case of Protestants, the coastal indicator has even larger predictive power. Protestants also set up missions in more ethnically diverse districts and those at a higher elevation. But the correlation between Protestant missionaries and the provision of education is again insignificant.

In the main regressions we would like to control for all historical variables pre-dating Catholic missionary location. Apart from the geographical variables that also control for coastal location and railways, which measure accessibility, there is no data to control for historical characteristics before the 20^{th} century. Hence, we control for historical characteristics in the first year the data are available, namely 1931. The first colonial census was conducted in 1872, but these early censuses were unreliable. More systematic enumeration began with the 1891 census, but information on the Princely States that account for one-third of the colonial Indian population was reported for aggregate regions, not individual states. Some of the information on the Princely States was also incorrectly enumerated in the early censuses.[22] We use the 1931 census because it has the most

[21] We constructed these averages using information in the Progress of Education in India (Great Britain, 1914).

[22] For example, literacy is incorrectly enumerated in the 1901 census for the Central India Agency States. We collected data on the historical variables for the British Indian districts in 1901 that are correctly

detailed and accurate information on the Princely States and districts of British India. Migration and urbanization was low and largely unchanged in the colonial era, so these historical variables are decent, though not ideal, proxies.

Our set of historical variables proxy for missionaries positively selecting districts with more Brahmins, urbanized areas and those that are better connected to railways, or negatively selecting districts with more tribal groups. Whether the district was a part of Princely India is also important because Christian missionaries were more common in British India.[23] If the location of Catholic missionaries are correlated with initial conditions that we control for in the main regressions, there is less of a concern because we are already accounting for that observable historical characteristic.

5 Tertiary Education and Luminosity: IV Estimates

5.1 Main Results

We present the IV results in *Table* 4. Moving across the columns we add more controls, with *column* (4) being the complete specification. The first stage results, displayed in *Panel A*, show the indicator for Catholic missionaries is positive and significant across specifications, with the magnitude going down as we add more controls. In *column* (4), which includes our full set of controls, the marginal effect of Catholic missionaries on contemporary tertiary education is 0.015. Thus, the population share of individuals with tertiary education is 25% higher in districts with a Catholic missionary given the mean tertiary education share of 0.058. This rather large point estimate is highly significant. Reassuring, we do not suffer from a weak instrument problem; the *F stat* is 20 and the Cragg-Donald Wald statistic is 28.[24] Other interesting results from the first stage show that districts with centroids closer to big cities and those with a higher urbanization rate in 1931 have more tertiary education today. As expected, a large presence of disadvantaged

enumerated in that census. Our IV results are robust and in fact larger when we use these 1901 controls for the districts of British India.

[23]This effect is especially strong for Protestant missionaries but nonetheless also negative and significant for Catholic missionaries. Including an indicator for a historical Princely State is problematic because these areas were not randomly selected making the indicator potentially endogenous. Our OLS and IV results on tertiary education are the same whether we include a Princely State dummy. We choose to show the results with this variable because they were more Catholic missionaries in British India.

[24]The null hypothesis is that the instrument is sufficiently weak that the 2SLS estimator is at least b percent as biased as the OLS estimator. The critical values provided by Stock and Yogo (2005) for values of b =10%, 15%, 20% and 25% bias are 16.38, 8.96, 6.66 and 5.53 respectively. We reject the null hypothesis as the Cragg-Donald Wald statistic exceeds the corresponding critical values.

communities is negatively correlated with tertiary education.[25]

Second stage results are displayed in *Panel B*. The IV coefficient on tertiary education is positive and significant at the 1 percent level with a magnitude of 12.97 (*column* (4)). A standard deviation increase in tertiary education raises log light density by 0.39, an economic effect of 9% given mean log light density. One standard deviation of tertiary education is equivalent to a 3 percentage point higher tertiary completion rate, as compared to the omitted share. The omitted share, by construction, is a combination of those who are illiterate, those with some schooling, those who have passed high school but not gone ahead to complete tertiary education.

As noted in our OLS results, one may argue the share of tertiary education is picking up the impact of other levels of schooling. In *column* (5) we address this issue by including the share of population with some schooling but no higher than secondary schooling. The IV coefficient on tertiary education is positive and significant, while the coefficient on schooling is not statistically significant. Moreover, the coefficient on tertiary education is similar in magnitude to the estimate in *column* (4), which confirms the omission of lower levels of education is not driving our results on tertiary education.

5.2 Endogeneity of Secondary Schooling

Although we control for the share of secondary schooling in *Table* 4, this raises a problem because this share is not exogenous. Moreover, the positive correlation between secondary and tertiary education may render the coefficient on secondary education insignificant in *column* (5). Hence, we need an instrument for secondary schooling. The location of Protestants missionaries offers one possibility given the burgeoning literature showing a positive impact of Protestant missions or missionaries on literacy. The principle of "Sola Scriptura" underpins the relationship between the Protestant religion and literacy because the Bible is the supreme authority in matters of doctrine and practice and one has to be literate to read the Bible. According to Gallego and Woodberry (2010) and Nunn (2014) this lead Protestant missionaries to promote education around the world. Studies have linked Protestants to literacy in Germany (Becker and Woessmann 2008) and India (Mantovanelli 2014). Scholars have also used variation in Protestant missionaries as an instrument for education (Acemoglu et al. 2014).

Following this literature we use the location of Protestant missionaries as of 1908 as an instrument for the share of secondary schooling in 2001. *Table* 5 presents the results using the two instruments: Catholic and Protestant missionaries to instrument for the share of

[25] In the interests of parsimony, we do not report these results, but they are available upon request.

the population with tertiary education and the population with secondary education.[26] The first stage for tertiary education shows a large and positive coefficient on Catholic missionaries, and a positive but insignificant coefficient on Protestant missionaries. This corroborates our understanding of the post-1947 education landscape when many Catholic colleges came on the scene and accounts for the persistent impact of Catholic missionaries on contemporary tertiary education. Although Protestants focused more on basic literacy, it is still somewhat surprising that we find no effect of Protestant missionaries on tertiary education. More broadly though these results suggest our Catholic instrument is not picking up the impact of Christian missionaries more generally.

Similar to the literature, we find Protestant missionaries had a positive impact on the share of the adult population with some schooling but no higher than secondary schooling. Districts with Protestant missionaries in 1908 are positively correlated with the secondary schooling share in 2001. However, we interpret these results with caution because of the low values of the Cragg-Donald and Kleibergen-Paap rk Wald tests. The location of Protestant missionaries seems to be a weak instrument for schooling.[27] In fact, the Kleiberger-Paap p-value in *column* (1) indicates the model is under-identified. As we have two endogenous variables, the model can be under-identified if only one of the instruments is valid. In our case only Catholic missionaries appears to be a valid instrument for tertiary education.

The second stage results are similar to the previous findings, again highlighting that our main IV results in *Table* 4 are not picking up the effect of other levels of education. The coefficient on tertiary education is less significant than earlier models because the combination of Protestant and Catholic missionaries are weak instruments jointly for tertiary and secondary education.

6 Sensitivity Analysis

6.1 Is location of Catholic missionaries a plausible instrument?

Our identification relies on the assumption that Catholic missionaries affect current luminosity through their influence on tertiary education. We report two pieces of evidence in

[26] If we use Protestant and Catholic missionaries as joint instruments for just tertiary education, we find similar results as in *Table* 4. The Hansen over-identification test in the parsimonious and full specification displays p-values equal to 0.144 and 0.973 respectively, giving further evidence of the exogeneity of the instruments.

[27] Acemoglu at el. (2014) use Protestant missionaries as an instrument for average levels of schooling and they also report a weak first stage in some models.

support of this assumption. In *Table* 6, we first show that the presence of Catholic missionaries in a district in 1911 is uncorrelated with the provision of education and education outcomes in the colonial period.[28] Catholic missionaries are positively correlated with the number of Catholic colleges only after Indian independence. In *columns* (1) and (2), we use the total literacy rate and English literacy rate in 1931 as dependent variables. In this context, English literacy is a better measure of tertiary education because most colleges used English as the language of instruction. The coefficient on Catholic missionaries is small and insignificant. In *columns* (3) to (6), we study the correlation between Catholic missionaries and the number of degree granting Catholic colleges as of 1911, 1951, 1971 and 2001 respectively.[29] There is a small correlation in 1951, which grows over time. By 2001 there is a significant and positive correlation between the historical presence of Catholic missionaries and Catholic colleges. Taken together these findings suggest Catholic missionaries were uncorrelated with tertiary education historically. Their impact on current tertiary education operates through an increase in the supply of colleges beginning in the 1950s.

Although the historical presence of Catholic missionaries had no impact on historical tertiary education, Catholics missionaries may have influenced current income via alternative channels. We therefore conduct rigorous checks on the validity of our exclusion restriction. Apart from education, Christian missionaries undertook other social activities like building hospitals and promoting better sanitation. Given the positive relationship between health, education and development, Catholic missionaries could influence current income by improving the health of the population.[30] In *Table* 7, *column* (1) we include infant mortality, our measure of health, as an additional control. The coefficient on tertiary education remains large and statistically significant. We also test for a direct relationship between infant mortality and the location of Catholic missionaries. First stage results estimating the impact of Catholic missionaries on alternative channels are displayed in

[28] Unfortunately, there are no other measures of economic outcomes at the district-level for colonial India.

[29] We obtained the information on Catholic colleges from the Catholic Directory of India 2010 that enumerates all Catholic education institutions with their address and year of opening. We used this information to construct the number of degree granting colleges in each district-year pair.

[30] Recent studies have analyzed the relationship between religion and health in India. Calvi and Mantovanelli (2014) find that proximity to a historical Protestant medical mission has a positive long-run effect on current health. They show that it is the proximity to a Protestant mission equipped with a medical facility that matters for current health and not the proximity to a generic Protestant mission. Using a broad set of instruments for Christian identity today, Menon (2015) finds that Christian infant girls score higher in terms of height for age than lower caste Hindu girls. In their first stage, however, the number of Catholic missions in 1910 is not statistically significant in any specification.

Appendix Table 1. Column (1) indicates the presence of missionaries in 1911 is unrelated to contemporaneous measures of health, suggesting the health mechanism is not a plausible channel by which Catholic missionaries influence current income.

The IV estimates on tertiary education could also be driven by general religiosity, which encourages attitudes of thrift, work ethic and honesty. For example, Nunn (2010) finds that descendants of populations that experienced greater missionary contact in colonial Africa are more likely to self-identify as Christians today. To assess this possibility, we control for the direct effect of religion by including the current population share of Christians as a control in *column* (2). While the current Christian share is negatively related to light density, the coefficient on tertiary education is unchanged.[31]

In *column* (3) we control for different measures of infrastructure because Catholic missionaries may have encouraged the construction of roads and irrigation facilities. Moreover, it may be the case that more government schools were established in districts with historical Catholic missionaries. If infrastructure is correlated with education and fosters development, the omission of infrastructure could bias the estimated coefficient. We include the proportion of villages with paved roads, the share of land that is irrigated, and government schools. Even after including these controls, tertiary education still has a positive and highly significant coefficient.

In *column* (4) we control for the share of migrants to the total population over 25 in the district. Although migration between states is not too widespread in India, we worry people may migrate in larger numbers to districts with a historical Catholic presence in pursuit of higher education or because these districts are more welcoming to outsiders. This would then bias our IV coefficient because the presence of Catholic missionaries may be impacting light density via higher migration rates. It is reassuring to see the coefficient on tertiary education is robust to controlling for migration.

In *column* (5) we control for the historical presence of Europeans in a district. A recent paper by Bai and Kung (2015) finds a significant impact of Protestant missionaries on income in China, another country where few individuals converted to Christianity. Although their paper suggests formal schooling as a link from Protestant missionaries

[31]The literature finds mixed results of religiosity on economic growth depending on how religion is measured. Using survey data on religiosity for a broad panel of countries, Barro and McCleary (2003) and McCleary and Barro (2006) find a positive effect of religious beliefs on growth– related to beliefs in hell, heaven and after-life–, while the influence of attendance to religious services is negative. In our context, the negative sign is perhaps because a larger proportion of disadvantaged groups such as the former lower castes and tribes were more likely to convert to Christianity in India. Our IV results on tertiary education are unchanged if we include the share of Christians in 1931 as a control.

to income, one could imagine general European presence leading to a diffusion of western knowledge that in turn could promote development. This would suggest our IV results are picking up the beneficial impact of western or European presence in a district. However, the coefficient on tertiary education is essentially unchanged when we control for the number of Europeans in a district in 1931.[32]

Finally, in *column* (6) we control for the historical population density of a district in 1931. While we always include the 1931 urbanization rate among our historical controls, one could argue that population density better captures the attraction of Catholic missionaries to districts with more people to convert. As seen this does not change the results on tertiary education.

One final concern is there may still be alternative mechanisms through which Catholic missionaries could affect current income that we are not accounting for. For example, one could argue that the presence of missionaries could affect the attitudes and culture not only of converters but also of people living in the area. As changes in attitudes and culture could affect income directly (Tabellini 2010), missionaries might have impacted current income through a culture and attitude channel. However, the absence of a survey on values at the district level makes it difficult to test this mechanism. As a robustness exercise, we ran the following reduced form regression to asses the validity of our exclusion restriction:

$$y_{ds} = \alpha_s + \lambda Catholic\ Mission_{ds} + \rho' C_{ds} + \pi_r' G_{ds} + \delta_r' H_{ds} + \varepsilon_{ds} \tag{2}$$

If the effect of Catholic missionaries on current development is mainly driven by its influence on tertiary education, we would expect that once we control for the level of tertiary education, the effect of Catholic missionaries disappears. The estimated coefficient of λ in *Table* 8 in *Panel A* ranges from 0.447 with a standard error of 0.073 in the parsimonious specification to 0.147 (standard error = 0.063) when all controls are accounted for. In all cases, the estimated coefficient of Catholic missionaries is highly significant. When we control for the population share with tertiary education, the results in *Panel B* show that the coefficient on Catholic missionaries is not statistically significant in any specification.

6.2 Sensitivity Checks

In this last sub-section, we first use different proxies for development in lieu of night lights. The literature analyzing district-level development in India has been scarce due to a lack

[32]Our results are unchanged if we standardize the Europeans by the total population of the district in 1931, or if we focus only on British citizens rather than Europeans.

of good quality data on local GDP per capita. Other scholars have used agricultural investment, agricultural productivity and the stock of health and education infrastructure as proxies for economic prosperity (e.g. Banerjee and Iyer 2005; Iyer 2010). Our use of night lights as a district-level outcome is part of our contribution to this literature. But, in *Table* 9 we present IV estimates for other measures of development.

In line with Henderson et al. (2012) and Michalopoulos and Papaionaou (2013, 2014), we have used light density as a proxy for development. To assess the sensitivity of the results, we computed alternative measures of lights. Column (1) shows the IV estimates on tertiary education for *lights per-capita* controlling for current, geographical, and historical variables as before. The coefficient on tertiary education is positive and statistically significant.[33]

The advantage of using night lights is they offer a good proxy for development at the district-level, when national accounts are unavailable. The disadvantage, however, is that it is difficult to interpret the estimated coefficient in line with the existing literature. Household consumption is a good proxy for disposable income and offers another proxy for development that is perhaps easier to interpret. So as an illustrative exercise we estimate the effect of tertiary education on per capita consumption using data on consumption expenditures from the National Sample Survey Organization. These data come from quinquennial surveys of a randomly selected sample of households conducted by the Ministry of Statistics and Programme Implementation.[34] The results, displayed in *column* (2), show a positive and highly significant coefficient on tertiary education. In quantitative terms, a one percentage point increase in the population share of individuals with completed tertiary education increases consumption per capita by 3.55 percent.[35]

Finally, in *Table 10,* we check the robustness of our results on light density to the presence of outliers. As our results could be driven by atypical observations or extreme

[33]These results on lights per capita should be interpreted with caution because there is a discrepancy in the timing of the lights and population data. High resolution radiance data is only available in 2006, whereas the population is taken from the decennial census in 2001. Extrapolating population for 2006 would require to fit a model using 2011 census data. We would need to make several assumptions for that exercise, some of which involve splitting districts in 2011. Land is, of course, consistent throughout in that regard.

[34]Data is taken from from the NSS 61st Round (Schedule 1.0 on Household consumption) 2004-2005.

[35]As with luminosity, we are not able to match the year for population and consumption. The first year post 2001 when a reliable district level figure can be calculated is 2004-05. This number may also be on the lower side as consumption, measured from household surveys, is much lower than consumption in national accounts (Deaton, 2005). Pinkovskiy and Sala-i-Martin (2014) also suggest that night lights come closer to measures of income per capita per national accounts compared to survey estimates.

values, we first drop districts with light density above the 99^{th} percentile and below the 1^{st} percentile in *column* (1). In *column* (2) we exclude districts with million plus population cities. In *column* (3) we exclude districts that contain the largest city in the state and in *column* (4) we drop districts that include the state capital. Across these specifications, the coefficient on tertiary education remains positive and statistically significant.

As mentioned earlier, St. Thomas Christians are an ancient Christian community that pre-date the arrival of Europeans. They are an important part of the Indian Catholic community and have set up many Catholic schools and colleges in Kerala where they account for majority of the Catholic population. Given their ancient lineage, one may be concerned the IV results are driven by Kerala and St. Thomas Christians. Hence, in *Column* (5), we drop the state of Kerala. Again, the results on tertiary education are essentially unchanged. In results unreported here we also checked whether one particular state may be responsible for the results by dropping one state at a time. We also used the more intensive form of missionaries per square km as the instrument. The coefficient on share tertiary was essentially unchanged across these additional robustness checks.

7 Conclusion

This paper investigates if higher levels of education lead to higher levels of development. The literature has studied this question extensively but has failed to draw strong conclusions because the evidence so far is mixed. Many studies suffer from identification problems such as reverse causality and omitted variables. For example, as richer and faster growing economies have more resources to invest in education, the direction of causality can go from development to education (Bils and Klenow, 2000).[36] Moreover, education can also be correlated with omitted variables that influence development. The solution to the problem is not trivial as we need to find an exogenous source of variation in education.

Although there have been attempts to tackle endogeneity in governance quality (Acemoglu et al. 2001), trade (Frankel and Romer 1999) and foreign aid (Burnside and Dollar 2000), few papers have seriously addressed the endoneneity of human capital. We use history to address this problem. In line with the empirical literature that looks at history to identify the causal impact of institutions on development (see Acemoglu et al, 2000, 2001), we address the endogeneity of education by exploiting the location of Catholic

[36]Bils and Klenow (2000) point out that the empirical literature has documented correlations but has not identified the direction of causation. Using calibration techniques, they argue the positive correlation between education and growth, found in Barro (1991), can be explained by a channel that goes from expected growth to schooling.

missionaries in colonial India.

Christian missionaries in general, and Catholic missionaries in particular are associated with high quality education around the world. However, the novelty of our identification strategy is to focus on India, a country where Christians form a small minority and have relatively limited influence over other institutions correlated with development. We also use district-level data with state fixed effects to control for institutions and cultural characteristics. In our analysis, we focus on tertiary education and propose a new instrument for the share of population with tertiary education. This level of education is particularly important in the Indian context because it is the main educational variable correlated with economic growth at the state level (Castelló-Climent and Mukhopadhyay 2013).[37] The focus on one country and a particular level of education is therefore part of our strategy to better identify the causal effect of education on development.

We find that exposure to Catholic missionaries at the beginning of the 20^{th} century has had a long-term impact on the current composition of education. Our identification strategy relies on the fact that Catholic missionaries were not actively involved in the provision of education in the colonial period. The number of colleges and schools are uncorrelated with the presence of Catholic missionaries in 1911. Moreover, Catholic missionaries in our context did not locate in richer or more educated places. Nevertheless, the historical network was a natural platform from which Catholics expanded the number of high quality colleges after independence. Using the location of Catholic missionaries in 1911 as an exogenous source of variation in levels of higher education, we find a strong and positive effect of the population share with tertiary education on contemporaneous levels of development, as measured by night light density. The magnitude of the effect is sizable; a standard deviation increase in tertiary education raises log light density by 0.39, an economic effect of 9% given mean log light density.

A broad array of sensitivity analyses indicate that it is implausible that Catholic missionaries have influenced current development through channels other than tertiary education. We also show the results hold with a rich set of geographical, historical and current controls. The effects of tertiary education on development are also not driven by lower levels of education. Finally, the results hold for alternative measures of development, and are not driven by outliers.

[37] Most of the literature has used average years of schooling to asses their effect on development and growth. However, average years of schooling is a broad measure of education and cannot disentangle the differential effects of each level of schooling (Vandenbussche et al. 2006). Other things being equal, an increment in average years of schooling in two countries may have different effects on development if the increment is driven by increases in primary, secondary or tertiary education (see Aghion et al. 2009).

Our main goal in this paper has been to identify the causal effect of tertiary education on development. However, we are unable to comment more broadly on the causal effect of only completing lower levels of education because in the Indian context these lower levels do not seem to play an important role on development. A broader examination across other countries could shed more light on the differential effect of lower and higher levels of education on development. In fact, the finding that Catholic missionaries mainly influence tertiary education, whereas Protestant missionaries mainly influence secondary and lower levels of schooling is an interesting result. If those findings are corroborated in other settings, the results offer new avenues for future research relating human capital to growth and development.

References

[1] Acemoglu, D. and M. Dell (2010). "Productivity Differences Between and Within Countries," *American Economic Journal: Macroeconomics*, Vol. 2(1), pp. 169-88.

[2] Acemoglu, D., F. A. Gallego y J. A. Robinson (2014). "Institutions, Human Capital and Development," *Annual Review of Economics*, Vol. 6, pp. 875-912.

[3] Acemoglu, D., S. Johnson and J. A. Robinson (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, 1369-1401.

[4] Acemoglu, D., S. Johnson and J. A. Robinson (2002). "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 118, 1231-1294.

[5] Aghion, Ph., L. Boustan, C. Hoxby and J. Vandenbussche (2009). "The Causal Impact of Education on Economic Growth: Evidence from the United States," Brookings Papers on Economic Activity, Spring 2009, Conference Draft.

[6] Alesina, A., S. Michalopoulos and E. Papaioannou (2015). "Ethnic Inequality," *Journal of Political Economy*, forthcoming.

[7] Angrist, J. D. and J-S. Pischke (2009). "Mostly Harmless Econometrics," Princeton Univesity Press.

[8] Ashraf, Q. and O. Galor (2013). "The 'Out-of-Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development." *American Economic Review*, Vol. 103 (1), pp. 1-46.

[9] Bai, Y. and J-K. Kung (2015). "Diffusing Knowledge While Spreading God's Message: Protestantism and Economic Prosperity in China, 1840-1920," *Journal of the European Economic Assocciation*, 13(4), pp. 669-98.

[10] Banerjee, A. and L. Iyer (2005) "History, Institutions and Economic Performance: the Legacy of Colonial Land Tenure Systems in India" *American Economic Review*, Vol. 95(4), pp. 1190-213.

[11] Barro, R. (1991). "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics*, 106 (2), 407–443.

[12] Barro, R. and R.M. McCleary (2003). "Religion and Economic Growth across Countries," *American Sociological Review*, Vol. 68(5), pp.760-81.

[13] Becker, S. O. and L. Woessmann (2008). "Luther and the Girls: Religious Denomination and the Female Education Gap in Nineteenth-century Prussia," *Scandinavian Journal of Economics*, Vol. 110(4), pp. 777-805.

[14] Becker, S. O. and L. Woessmann (2009). "Was Weber Wrong? A Human Capital Theory of Protestant Economic History," *Quarterly Journal of Economics*, Vol. 124(2), pp. 531-596.

[15] Benhabib, J. and M. Spiegel (1994). "The Role of Human Capital in Economic Development." *Journal of Monetary Economics*, 34 (2), pp. 143-174.

[16] Besley, T. and R. Burgess (2000). "Land Reform, poverty Reduction, and Growth: evidence from India," *Quarterly Journal of Economics*, 115(2), pp. 389-430.

[17] Bharadwaj, P., A. Khwaja and A. Mian (2008). "The Big March: Migratory Flows after the Partition of India." *Economic and Political Weekly* Vol. 43, No. 35, 39-49.

[18] Bils, M and P. Klenow (2000). "Does Schooling Cause Growth." *American Economic Review*, 90(5), pp.1160-1183.

[19] Burnside, C. and D. Dollar (2000). "Aid, Policies, and Growth." *American Economic Review*, Vol. 90 (4), pp. 847–68.

[20] Calvi, R. and F. Mantovanelli (2014). "Long-Term Effects of Access to Health Care: Medical missionaries in Colonial India" Mimeo Boston College.

[21] Cameron, A.C., J. Gelbach and D. Miller (2008). "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, Vol. 90, pp. 414-27.

[22] Card, D. (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, Vol. 69(5), pp. 1127-60.

[23] Castelló-Climent, A. and A. Mukhopadhyay (2013). "Mass Education or a Minority Well Educated Elite in the Process of Growth: the Case of India," *Journal of Development Economics*, Vol. 105, pp. 303-320.

[24] Catholic Students' Mission Crusade (CSMC) (1923). *India and its missionaries.* New York: Macmillan.

[25] Cohen, D. and M. Soto (2007). "Growth and Human Capital: Good Data, Good Results," *Journal of Economic Growth*, Vol. 12, pp. 51–76.

[26] Codina, G. S.J. (2000). "A Century of Jesuit Education, 1900-2000," in Jesuits: Yearbook of the Society of Jesus.

[27] Chaudhary, L. (2015). "Caste, Colonialism and Schooling: Education in British India." in Latika Chaudhary, Bishnupriya Gupta, Tirthankar Roy and Anand Swamy eds., A New Economic History of India. London: Routledge.

[28] Chaudhury, N., J. Hammer, M. Kremer, K. Muralidharan, and F. H. Rogers (2006). "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives*, Vol. 20(1), pp. 91-116.

[29] Chen, X. and W. D. Nordhaus (2010). "The Value of Luminosity Data as a Proxy for Economic Statistics." NBER Working Papers N^o 16.317.

[30] Deaton, A. (2005). "Measuring Poverty in a Growing World (or Measuring Growth in a Poor World)." *The Review of Economics and Statistics*, Vol. 87(1), pp.1-19.

[31] De la Fuente, A. and R. Doménech (2006). "Human Capital in Growth Regressions: How Much Difference Does Data Quality Make?," *Journal of the European Economic Association*, 4(1), pp. 1-36.

[32] Easterly, W. and R. Levine (2012). "The European Origins of Economic Development," NBER Working Paper N^o 18162.

[33] Frankel, J. A. and D. Romer (1999). "Does Trade Cause Growth?" *American Economic Review*, Vol. 89(3), pp. 379-399.

[34] Frykenberg, R. E. (2008). "Christianity in India: From Beginnings to the Present," Oxford University Press.

[35] Gallego, F. A. and R. D. Woodberry (2010). "Christian Missionaries and Education in Former African Colonies: How Competition Mattered," *Journal of African Economies*, Vol. 19(3), pp. 294-329.

[36] Gennaioli, N., R. La Porta, F. Lopez-de-Silanes, A. Shleifer (2013). "Human Capital and Regional Development," *Quarterly Journal of Economics,* Vol. 128(1), pp. 105-64.

[37] Glaeser, E. L., R. La Porta, F. Lopezde-Silanes, and A. Shleifer (2004). "Do Institutions Cause Growth?" *Journal of Economic Growth*, Vol. 9 (3), pp. 271–303.

[38] Great Britain. (1914). "East India (education). Progress of education in India, 1907-1912. Sixth quinquennial Review. " His Majesty's Stationery Office.

[39] Henderson, J. V., A. Storeygard and D. N. Weil (2012). "Measuring Economic Growth from Outer Space." *American Economic Review*, Vol. 102(2), pp. 994-1028.

[40] Iyer, L. (2010). "Direct versus Indirect Colonial Rule in India: Long-term Consequences," *Review of Economics and Statistics*, Vol. 92(4), pp. 693-713.

[41] Krueger, A. and M. Lindahl (2001). "Education for Growth: Why and For Whom?" *Journal of Economic Literature*, vol. XXXIX, 1101-36.

[42] Kumar, H. and R. Somanathan (2009). "Mapping Indian Districts Across Census Years, 1971-2001," *Economic and Political Weekly* , Vol. 36(9), pp. 783-92.

[43] Lucas, R. (1988). "On the Mechanics of Economic Development," *Journal of Monetary Economics*, Vol. 22, pp. 3-42.

[44] Mankiw, N. G., D. Romer and D. Weil (1992). "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics,* 107, pp. 407-37.

[45] Mantovanelli, F. G. (2014). "The Protestant Legacy: missionaries and Literacy in India," Mimeo Boston College.

[46] McCleary, R.M. and Barro, R. (2006). "Religion and Economy," *Journal of Economic Perspectives*, Vol. 20(2), pp.49-72.

[47] Menon, N. (2015). "Christianity and Infant Health in India," IZA DP N^o 9177.

[48] Michalopoulos, S. and E. Papaioannou (2013) "Pre-colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, Vol. 81(1), pp. 113-52.

[49] Michalopoulos, S. and E. Papaioannou (2014) "National Institutions and Subnational Development in Africa," *Quarterly Journal of Economics*, pp. 151-213.

[50] Munshi, K. and M. Rosenzweig (2009) "Why is Mobility in India so Low? Social Insurance, Inequality, and Growth," *NBER Working Papers* 14850, National Bureau of Economic Research, Inc.

[51] Nunn, N. (2010). "Religious Conversion in Colonial Africa," *American Economic Review Papers and Proceedings*, Vol. 100 (2), pp. 147-52.

[52] Nunn, N. (2014). "Gender and Missionary Influence in Colonial Africa," in Emmanuel Akyeampong, Robert Bates, Nathan Nunn and James A. Robinson (eds.), Africa's Development in Historical Perspective, Chapter 16, pp. 489-512. Cambridge University Press, 2014.

[53] Nunn, N. (2014). "Historical Development," *Handbook of Economic Growth*, Chapter Seven, Vol. 2A, pp. 347-402.

[54] Pinkovskiy, M. L. (2013). "Economic Discontinuities at Borders: Evidence from Satellite Data on Lights at Night" Mimeo, Federal Reserve Bank of New York.

[55] Pinkovskiy, M. and X. Sala-i-Martin (2015). "Lights, Camera,...Income! Estimating Poverty Using National Accounts, Survey Means, and Lights," Federal Reserve Bank of New York Staff Reports N^o. 669.

[56] Pritchett, L. (2001). "Where Has All the Education Gone," *World Bank Economic Review*, Vol. 15(3), pp. 367-91.

[57] Richter, J. (1908). *A history of missionaries in India.* Edinburgh: Oliphant.

[58] Sachs, J. (2003), "Institutions don't rule: Direct Effects of Geography on Per Capita Income" *NBER Working Papers* 9490, National Bureau of Economic Research, Inc.

[59] Sokoloff, K. L., and S. L. Engerman (2000). "Institutions, Factor Endowments, and Paths of Development in the New World." *Journal of Economic Perspectives*, Vol. 14 (3), pp. 217–32.

[60] Spolaore, E. and R. Wacziarg (2009). "The Diffusion of Development." *Quarterly Journal of Economics*, Vol. 124(2), pp. 469-529.

[61] Spolaore, E. and R. Wacziarg (2013). "How Deep Are the Roots of Economic Development?," *Journal of Economic Literature*, Vol. 51(2), pp. 325-69.

[62] Stock J.H. and M. Yogo (2005). "Testing for Weak Instruments in Linear IV Regression," In Andrews DWK Identification and Inference for Econometric Models. New York: Cambridge University Press, pp. 80-108.

[63] Tabellini, G. (2010). "Culture and Institutions: Economic Development in the Regions of Europe." *Journal of the European Economic Association*, Vol. 8(4), pp. 677-716.

[64] Vandenbussche, J., P. Aghion and C. Meghir (2006). "Growth, Distance to Frontier and Composition of Human Capital." *Journal of Economic Growth*,11, pp. 97-127.

[65] World Missionary Conference (1910). *Statistical atlas of Christian missions: containing a directory of missionary societies, a classified summary of statistics, an index of mission stations, and a series of specially prepared maps of mission fields.* Edinburgh: The Conference.

Table 1: Summary Statistics

Variables	Obs	Mean	Std. Dev.	Min	Max
Current					
Log Light Density	500	4.240	1.092	-0.953	6.407
Log Consumption per-capita	496	8.925	0.308	7.991	9.915
Share Tertiary (25+)	500	0.058	0.03	0.014	0.213
Share Secondary (25+)	500	0.339	0.097	0.115	0.664
Share SC	500	0.162	0.079	0	0.501
Share ST	500	0.11	0.177	0	0.938
Pop 25+	500	874,426	613,867	41,358	4,670,683
Geography					
Coastal	500	0.098	0.298	0	1
Longitude	500	79.793	5.073	69.778	95.627
Latitude	500	23.108	5.907	8.308	34.534
Av. River Length	500	12.156	3.737	2.932	30.342
Min Dist Big City	500	336.522	181.811	3.563	947.762
Av. Elevation	500	403.836	619.58	3.967	4941.724
Historical					
Catholic Missionary	500	0.3	0.459	0	1
Protestant Mission	500	0.582	0.494	0	1
Share Urban, 1931	500	0.107	0.08	0	.50
Share Brahman, 1931	500	0.056	0.043	0	.27
Share Tribal, 1931	500	0.034	0.088	0	.69
Princely State	500	0.34	0.474	0	1
Railway Line, 1909	500	0.782	0.413	0	1
Historical, 1901 - Controls in Tables 3A and 3B					
Share Urban, 1901	156	0.086	0.074	0.000	0.354
Share Brahman, 1901	156	0.038	0.026	0.004	0.159
Share Lower Castes, 1901	156	0.213	0.117	0.006	0.592
Share Tribes, 1901	156	0.047	0.111	0.000	0.548
Ethnic Frac, 1901	156	0.741	0.139	0.317	0.898
Income-Tax per Capita, 1901	156	0.051	0.044	0.000	0.364
Colleges, 1901	156	2.327	3.380	0.000	17.000
Schools, 1901	156	823.385	627.194	0.000	4553.000

Note: In the regressions, Pop 25+ is reported per 1,000,0000 and Min Dist Big City and Av. Height are reported per 100. The summary statistics for the Historical, 1901 - Controls cover the districts in the former British Indian provinces of Bengal, Bihar and Orissa, Bombay and Madras.

Table 2: OLS Results - Log Light Density

	(1)	(2)	(3)	(4)	(5)
Share Tertiary	14.396***	11.203***	10.638***	9.871***	8.019***
	(1.136)	(1.176)	(0.921)	(0.965)	(1.077)
Share Secondary					1.317***
					(0.506)
Share SC		-0.045	-0.405	-0.623	-0.737
		(0.687)	(0.499)	(0.490)	(0.477)
Share ST		-2.109***	-1.638***	-1.445***	-1.358***
		(0.319)	(0.219)	(0.228)	(0.245)
Pop 25+		0.154***	0.094*	0.048	0.068
		(0.059)	(0.054)	(0.052)	(0.052)
Coastal			-0.131*	-0.187***	-0.233***
			(0.072)	(0.070)	(0.074)
Longitude			0.038*	0.044**	0.039*
			(0.022)	(0.022)	(0.022)
Latitude			0.009	-0.006	0.000
			(0.020)	(0.019)	(0.019)
Av. River Length			0.013	0.012	0.011
			(0.008)	(0.008)	(0.008)
Min Dist Big City			-0.132***	-0.128***	-0.119***
			(0.026)	(0.026)	(0.026)
Average Elevation			-0.084***	-0.082***	-0.082***
			(0.005)	(0.006)	(0.006)
Share Urban, 1931				0.532	0.492
				(0.464)	(0.454)
Share Brahmin, 1931				-0.100	-0.532
				(0.958)	(0.976)
Share Tribal, 1931				-0.843*	-0.815*
				(0.481)	(0.470)
Princely State				-0.050	-0.005
				(0.059)	(0.063)
Railway Line, 1909				0.264***	0.254***
				(0.071)	(0.072)
Constant	1.891***	2.476***	2.050	1.991	1.741
	(0.473)	(0.364)	(1.576)	(1.578)	(1.585)
Observations	500	500	500	500	500
R-squared	0.624	0.705	0.793	0.804	0.807
State FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3A: Selection of Catholic Missionary Location?

Dep. Variable - Indicator for Catholic Missionary in 1911			
	(1)	(2)	(3)
Coastal	0.314***	0.345***	0.269**
	(0.079)	(0.110)	(0.118)
Longitude	-0.029**	0.003	-0.031
	(0.013)	(0.038)	(0.046)
Latitude	-0.028*	-0.003	0.052
	(0.015)	(0.039)	(0.047)
Av. River Length	-0.007	0.011	0.003
	(0.005)	(0.013)	(0.014)
Min Dist Big City	-0.001	-0.040	-0.057
	(0.017)	(0.033)	(0.035)
Average Elevation	0.008*	0.060***	0.041
	(0.004)	(0.018)	(0.025)
Railway Line, 1909	0.257***	0.287***	0.258**
	(0.043)	(0.102)	(0.112)
Share Urban, 1901			1.739
			(1.094)
Share Brahman, 1901			2.493
			(2.103)
Share Lower Castes, 1901			-0.399
			(0.639)
Share Tribes, 1901			1.048*
			(0.558)
Ethnic Frac, 1901			-0.249
			(0.611)
Pop, 1901			0.000
			(0.000)
Income-Tax per Capita, 1901			0.771
			(1.193)
Colleges, 1901			-0.020*
			(0.011)
Schools, 1901			0.000
			(0.000)
Constant	3.115***	-0.367	0.828
	(1.047)	(2.391)	(3.138)
Observations	500	156	156
R^2	0.297	0.360	0.408

State FE are included. Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3B: Selection of Protestant Missionary Location?

Dep. Variable - Indicator for Protestant Missionary in 1911			
	(1)	(2)	(3)
Coastal	0.327***	0.394***	0.382***
	(0.052)	(0.078)	(0.086)
Longitude	-0.008	-0.054*	-0.018
	(0.016)	(0.033)	(0.041)
Latitude	-0.033**	0.006	-0.012
	(0.016)	(0.030)	(0.039)
Av. River Length	-0.003	0.008	0.004
	(0.006)	(0.011)	(0.012)
Min Dist Big City	0.011	0.034	0.025
	(0.018)	(0.025)	(0.028)
Average Elevation	0.016**	0.040**	0.047**
	(0.006)	(0.016)	(0.022)
Railway Line, 1909	0.212***	0.279***	0.268***
	(0.057)	(0.090)	(0.097)
Share Urban, 1901			-1.834**
			(0.899)
Share Brahman, 1901			2.214
			(2.581)
Share Lower Castes, 1901			0.092
			(0.495)
Share Tribes, 1901			0.224
			(0.632)
Ethnic Frac, 1901			0.996**
			(0.485)
Pop, 1901			0.000
			(0.000)
Income-Tax per Capita, 1901			1.086
			(0.983)
Colleges, 1901			-0.001
			(0.013)
Schools, 1901			-0.000
			(0.000)
Constant	1.507	3.949*	1.161
	(1.180)	(2.020)	(2.605)
Observations	500	156	156
R^2	0.205	0.384	0.423

State FE are included. Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Instrumental Variables Results

	(1)	(2)	(3)	(4)	(5)
Panel A: First Stage - Catholic Missionary on Tertiary Education					
Catholic Missionary	0.027***	0.018***	0.019***	0.015***	0.012***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
F-stat Excluded IV	57.27	26.74	28.85	20.17	17.28
Cragg-Donald Wald F-stat	90.19	37.73	40.37	28.07	16.16
Panel B: Second Stage - Log Light Density					
Share Tertiary	16.246***	14.232***	15.621***	12.968***	12.651**
	(2.264)	(3.514)	(3.089)	(3.788)	(5.010)
Observations	500	500	500	500	500
R^2	0.622	0.700	0.781	0.800	0.801
Controls					
State FE	YES	YES	YES	YES	YES
Current	NO	YES	YES	YES	YES
Geographic	NO	NO	YES	YES	YES
Historical	NO	NO	NO	YES	YES
Schooling	NO	NO	NO	NO	YES

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Current controls include the population aged 25 and above, share of Scheduled Castes and share of Scheduled Tribes in 2001. Geographic controls include indicator for coastal districts, longitude, latitude, average river length, minimum distance to a big city and average elevation. Historical controls include 1931 share of Brahmins, Tribals and urban population, and an indicator for railways in 1909 and an indicator for Princely State. Schooling control includes the population share with some schooling but no higher than secondary education called Share Secondary.

Table 5: Instrumental Variables Results for Tertiary and Secondary Schooling

	(1)	(2)	(3)	(4)
Panel A: First Stage - Christian Missionary on Tertiary Education				
Catholic Missionary	0.025***	0.017***	0.018***	0.014***
	(0.004)	(0.003)	(0.004)	(0.003)
Protestant Missionary	0.005*	0.003	0.003	0.003
	(0.003)	(0.003)	(0.003)	(0.003)
Panel B: First Stage - Christian Missionary on Secondary Schooling				
Catholic Missionary	0.034***	0.023***	0.022***	0.014*
	(0.008)	(0.008)	(0.008)	(0.008)
Protestant Missionary	0.016**	0.017**	0.015**	0.013*
	(0.007)	(0.007)	(0.007)	(0.007)
F-stat Excluded IV	2.14	4.83	3.82	2.49
Cragg-Donald Wald F-stat	1.025	2.286	1.756	1.192
Panel C: Second Stage - Log Light Density				
Share Tertiary	36.846	21.984**	16.084*	13.217[a]
	(22.896)	(11.143)	(9.262)	(9.331)
Share Schooling	-13.734	-4.694	-0.315	-0.218
	(14.686)	(6.016)	(5.484)	(6.537)
Observations	500	500	500	500
R^2	-0.038	0.607	0.778	0.799
Controls				
State FE	YES	YES	YES	YES
Current	NO	YES	YES	YES
Geographic	NO	NO	YES	YES
Historical	NO	NO	NO	YES

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 [a] p<0.15
Current controls include the population aged 25 and above, share of Scheduled Castes and share of Scheduled Tribes in 2001. Geographic controls include indicator for coastal districts, longitude, latitude, average river length, minimum distance to a big city and average elevation. Historical controls include 1931 share of Brahmins, Tribals and urban population, and an indicator for railways in 1909 and an indicator for Princely State.

Table 6: Mechanism from Catholic Missionary to Tertiary Education - Catholic Colleges

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Literacy 1931	English Literacy 1931	Catholic Colleges 1911	Catholic Colleges 1951	Catholic Colleges 1971	Catholic Colleges 2001
Catholic Missionary	0.002	0.000	0.026	0.082*	0.257**	0.427**
	(0.002)	(0.000)	(0.017)	(0.049)	(0.108)	(0.202)
Observations	500	500	327	500	500	500
R^2	0.889	0.795	0.122	0.273	0.422	0.421
State FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In specifications (1) - (6), we include the standard geographic and historical controls. The geographic controls include indicator for coastal districts, longitude, latitude, average river length, minimum distance to a big city and average elevation. Historical controls include 1931 share of Brahmins, Tribals and urban population, and an indicator for railways in 1909 and an indicator for Princely State. In specifications (3)-(6), we also include current controls for population aged 25 and above, and the share of Scheduled Castes and share of Scheduled Tribes in 2001.

Table 7: IV Robustness Checks, Other Channels - Log Light Density

	(1)	(2)	(3)	(4)	(5)	(6)
Share Tertiary	7.883**	13.955***	12.130***	12.990***	13.093***	13.004***
	(3.911)	(4.041)	(4.134)	(3.856)	(3.649)	(3.754)
Observations	462	500	495	500	500	500
R^2	0.810	0.799	0.806	0.800	0.800	0.803
State FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Medical	YES	NO	NO	NO	NO	NO
Religiosity	NO	YES	NO	NO	NO	NO
Infrastructure	NO	NO	YES	NO	NO	NO
Migration	NO	NO	NO	YES	NO	NO
Europeans	NO	NO	NO	NO	YES	NO
Pop-Density, 1931	NO	NO	NO	NO	NO	YES

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
The controls include current, geographic and historical controls. Current controls include the population aged 25 and above, share of Scheduled Castes and share of Scheduled Tribes in 2001. Geographic controls include indicator for coastal districts, longitude, latitude, average river length, minimum distance to a big city and average elevation. Historical controls include 1931 share of Brahmins, Tribals and urban population, and an indicator for railways in 1909 and an indicator for Princely State.

Table 8: Reduced Form IV Test

	(1)	(2)	(3)	(4)	(5)
Panel A: OLS - Catholic Missionary on Log Light Density					
Catholic Missionary	0.447***	0.254***	0.292***	0.200***	0.147**
	(0.073)	(0.067)	(0.062)	(0.066)	(0.063)
Observations	500	500	500	500	500
R^2	0.514	0.648	0.746	0.765	0.790
Panel B: OLS - Catholic Missionary and Tertiary Education on Log Light Density					
Share Tertiary	14.048***	10.963***	10.210***	9.684***	7.793***
	(1.291)	(1.242)	(0.940)	(0.986)	(1.108)
Catholic Missionary	0.060	0.058	0.101*	0.051	0.056
	(0.072)	(0.066)	(0.056)	(0.060)	(0.061)
Observations	500	500	500	500	500
R^2	0.624	0.705	0.794	0.804	0.808
Controls					
State FE	YES	YES	YES	YES	YES
Current	NO	YES	YES	YES	YES
Geographic	NO	NO	YES	YES	YES
Historical	NO	NO	NO	YES	YES
Schooling	NO	NO	NO	NO	YES

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Current controls include the population aged 25 and above, share of Scheduled Castes and share of Scheduled Tribes in 2001. Geographic controls include indicator for coastal districts, longitude, latitude, average river length, minimum distance to a big city and average elevation. Historical controls include 1931 share of Brahmins, Tribals and urban population, and an indicator for railways in 1909 and an indicator for Princely State. Schooling control includes the population share with some schooling but no higher than secondary education called Share Secondary.

Table 9: IV - Other Measures in Logs

	(1)	(2))
	Lights per Capita	Consumption per Capita
Share Tertiary	5.891*	3.551***
	(3.454)	(1.327)
Observations	500	496
R^2	0.757	0.728
State FE	YES	YES
Controls	YES	YES

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
The controls include current, geographic and historical controls. Current controls include the population aged 25 and above, share of Scheduled Castes and share of Scheduled Tribes in 2001. Geographic controls include indicator for coastal districts, longitude, latitude, average river length, minimum distance to a big city and average elevation. Historical controls include 1931 share of Brahmins, Tribals and urban population, and an indicator for railways in 1909 and an indicator for Princely State.

Table 10: IV Robustness Checks, Dropping Outliers - Log Light Density

	(1)	(2)	(3)	(4)	(5)
Share Tertiary	11.727***	12.915***	13.481***	12.856***	13.151***
	(3.600)	(3.787)	(4.959)	(4.370)	(4.018)
Observations	490	495	486	484	486
R^2	0.780	0.801	0.798	0.800	0.796
State FE	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES
Drop Light Density Outliers	YES	NO	NO	NO	NO
Drop Million+ City	NO	YES	NO	NO	NO
Drop State Capital	NO	NO	YES	NO	NO
Drop Largest City	NO	NO	NO	YES	NO
Drop Kerala	NO	NO	NO	NO	YES

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The controls include current, geographic and historical controls. Current controls include the population aged 25 and above, share of Scheduled Castes and share of Scheduled Tribes in 2001. Geographic controls include indicator for coastal districts, longitude, latitude, average river length, minimum distance to a big city and average height. Historical controls include 1931 share of Brahmins, Tribals and urban population, and an indicator for railways in 1909 and an indicator for Princely State.

Appendix Table 1: Catholic Missionary on Alternative Channels

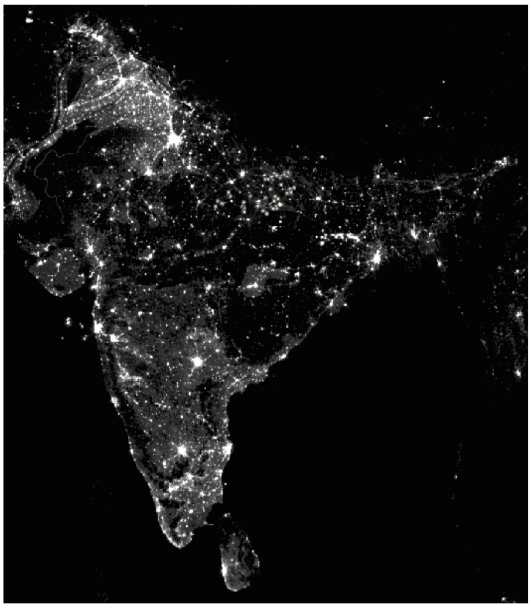
	(1)	(2)	(3)	(4)	(5)	(6)
	Under 5 Mortality Rate	Share Christians	Share Villages with Paved Road	Share Irrigated	Government Schools	Share of Migrants
Catholic Missionary	-3.196	0.011**	0.015	0.662	-134.124**	0.002
	(2.173)	(0.006)	(0.013)	(1.577)	(66.361)	(0.009)
Observations	462	500	500	500	495	500
R-squared	0.816	0.529	0.838	0.646	0.702	0.482
State FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

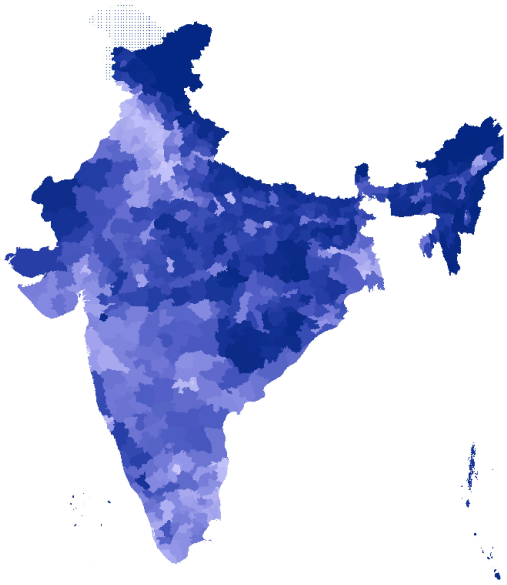
The controls include current, geographic and historical controls. Current controls include the population aged 25 and above, share of Scheduled Castes and share of Scheduled Tribes in 2001. Geographic controls include indicator for coastal districts, longitude, latitude, average river length, minimum distance to a big city and average elevation. Historical controls include 1931 share of Brahmins, Tribals and urban population, and an indicator for railways in 1909 and an indicator for Princely State.

Figure 1: Night Lights

(a) Night Lights Map

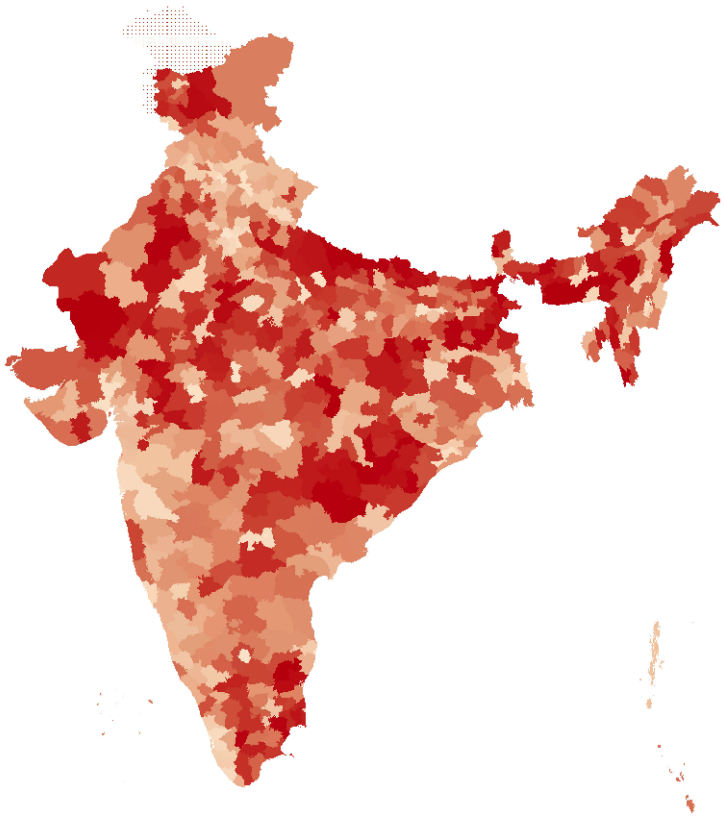


(b) District-Level Luminosity



Note: Dark colors correspond to lower luminosity and light colors correspond to higher luminosity.

Figure 2: District-Level Share of Tertiary Education



Note: Dark colors correspond to lower tertiary education and light colors correspond to higher tertiary education.

Figure 3: Correlation Light Density and Tertiary Education

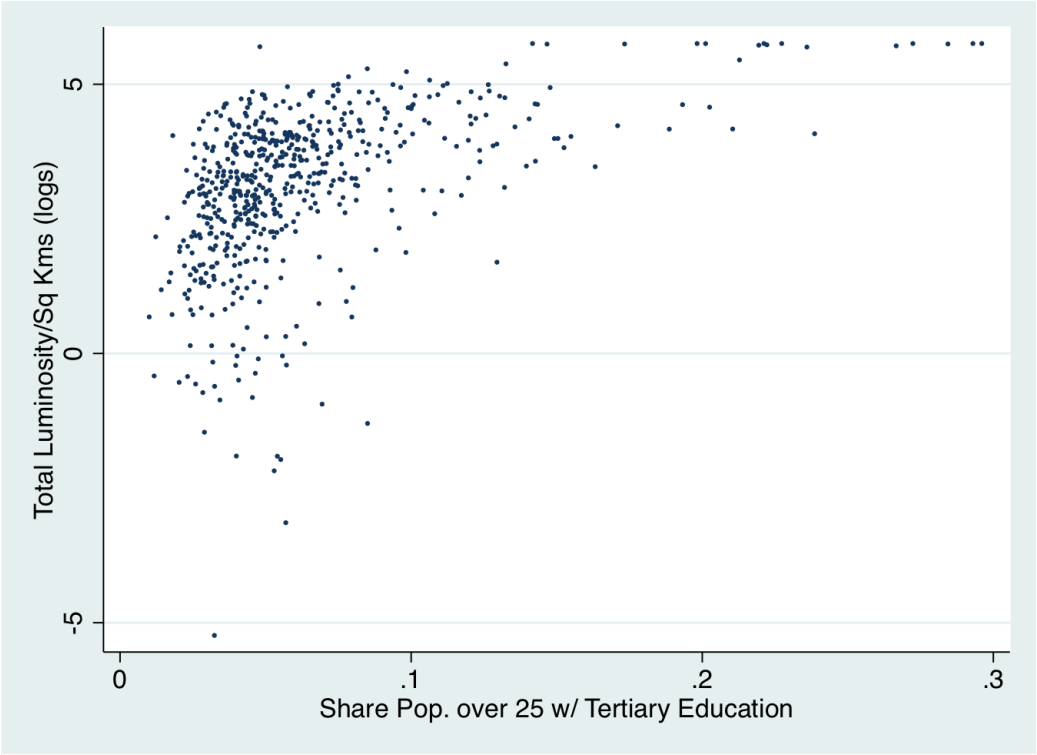


Figure 4: Location of Catholic Missaries

(a) Historical Map, North

(b) Historical Map, South

Figure 5: Catholic Missionary Location, 1911