Higher Education and Prosperity: From Catholic Missionaries to Luminosity in India*

Amparo Castelló-Climent, Latika Chaudhary and Abhiroop Mukhopadhyay

August 2017

Abstract

This paper estimates the impact of completed higher education on economic prosperity across Indian districts. To address the endogeneity of higher education, we use the location of Catholic missionaries circa 1911 as an instrument. Catholics constitute a very small share of the population in India and their influence beyond higher education has been limited. Our instrumental variable results find a positive effect of higher education on development, as measured by light density. The results are robust to alternative measures of development, and are not driven by lower levels of schooling or other channels by which missionaries could impact current income.

JEL classification: I25, N35,O15

Key words: Human Capital, Catholic Missionaries, Subregional Analysis

*Corresponding author: Amparo Castelló-Climent, Institute for International Economics, Universidad de Valencia, Facultad de Economía, Avenida de los Naranjos, 46022, Valencia, Spain. E-mail: amparo.castello@uv.es. We thank the editor, Joachim Voth, and three anonymous referees for their helpful comments and suggestions. We would also like to thank participants at the X^{th} Annual Conference on Economic Growth and Development, $XVII^{th}$ 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 ECO2015-65263-P 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 Defence or the U.S. Government.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/ecoj.12551

A large and important economics literature has studied the relationship between human capital, development and growth.[1] While early cross-country studies found a positive effect of average human capital on income and growth (Barro, 1991; Mankiw *et al.*, 1992), subsequent work has found weaker effects (Benhabib and Spiegel, 1994; Pritchett, 2001). Attempts to reconcile the different findings have pointed to mis-measurement of average education across countries and problems of reverse causality.[2] More recent work has addressed these concerns using better identification and different measures of human capital.[3] These advances, however, have focused on either cross-country analyses or studies of developed economies such as the United States. Our paper adds to the literature by studying the impact of completed higher education (i.e., secondary and college education) on development in India.

We focus on higher education for three reasons. First, average years of schooling is not a particularly informative measure in poor countries because the large number of people with no education skew the average. Second, services, in particular business services, have played a key role in India's recent growth experience (Kotwal *et al.*, 2011). And, industries associated with services such as software and information technology employ workers with higher levels of education. Castelló-Climent and Mukhopadhyay (2013) also find that only the share of population with completed tertiary education is correlated with state-level economic growth in India. The absence of a statistically robust relationship between the population share with only primary or middle schooling and economic growth suggests that the returns are concentrated in higher education. Third, the importance of highly educated workers is not unique to India. Studies of both contemporary and historical economies suggest higher levels of education can play a disproportionate role in the development process. Meisenzahl and Mokyr (2012) argue that skilled engineers and mechanics played a key role in the British Industrial Revolution while average literacy was relatively low.[4]

[1] See Becker (1964), Uzawa (1965), Nelson and Phelps (1966), Lucas (1988), and Romer (1990).

[2] Krueger and Lindahl (2001) point 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). If the stock of human capital is measured with error, their first differences will bias the estimated coefficient towards zero (Cohen and Soto, 2007; De la Fuente and Doménech, 2006). Bils and Klenow (2000) have highlighted how higher income and faster growth rates can generate more resources to invest in education creating the problem of reverse causality. Sianesi and Van Reenen (2003) offer a great review of this literature.

[3] For example, Acemoglu *et al.* (2014) instrument for average years of schooling using the historical presence of Protestant missionaries around the world. In terms of the composition of human capital, Goldin and Katz (2008) offer a detailed account of the rise in secondary education over the 20^{th} century and its contribution to economic growth and development in the United States.

[4] They highlight the role of the apprentice system and not formal education per se. Their focus, however, is on the contribution of the upper tail of the human capital distribution.

Squicciarini and Voigtländer (2015) show that upper-tail knowledge rather than average human capital, as measured by literacy, also drove French industrialisation. Cantoni and Yuchtman (2014) find that the creation of new universities in medieval Germany, as a result of the papal schism, resulted in greater economic activity in that period.[5] And, using contemporary data, Gennaioli *et al.* (2013) find that managerial education plays a bigger role in explaining regional development than workers human capital.

In this paper, we study the impact of higher education 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 share common governance and national policies thereby reducing some concerns related to omitted variables correlated with education and development. Since institutions are mainly determined at the central and state level, we can account for institutions without the need of an additional instrument.[6] Despite the advantages, 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 as a proxy for local income, in line with the recent literature (Henderson *et al.*, 2012; Micholopoulos and Papaioannou, 2013, 2014; Alesina *et al.*, 2016).

We find a strong positive association between the share of the population with completed higher education (i.e., individuals with completed secondary schooling, graduate-level diplomas or college degrees) and light density at night in OLS models controlling for state fixed effects. This relationship is robust to many factors that may jointly influence higher 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 higher education, however, poses an empirical challenge because higher education and the evolution of income generally go hand in hand. To address this concern we use the historical location of Catholic missionaries in the early 20^{th} century as an instrument for current higher education. We extract the precise location of Catholic missionaries in 1911 from the first edition of the *Atlas Hierarchicus*, over-

[5] In a similar vein, Maloney and Valencia (2017) find the density of engineers at the dawn of the second Industrial Revolution is associated with current levels of income, even controlling for literacy and the number of lawyers.

[6] The difficulty of jointly identifying the effect of human capital and institutions is well known as both are endogenous and the potential instruments for them are likely to be correlated (Glaeser *et al.*, 2004). A new fascinating study by Wantchekon *et al.* (2015) addresses this issue by studying the effect of colonial schools in Benin that were established before formal colonial institutions. They find large, positive and significant effects of human capital on standard of living and political participation of individuals that attended these schools. Interestingly, the effects extend to the children of the treated individuals and extended family members.

laying the historical maps on 2001 district borders using geographic information system (GIS) techniques.

Our reading of the history and empirical analysis suggests the location of Catholic missionaries is a plausible instrument. The location of Catholic missionaries circa 1911 has a large, positive and statistically significant impact on the share of the 2001 adult population with higher education. We find the share of the 2001 adult population with higher education is 2.9 percentage points higher in districts with a Catholic missionary presence, which translates into an economic effect on the order of 16% given mean higher education of 0.18.

What links Catholic missionaries to contemporary higher education? We argue that districts with Catholic missionaries were better placed to respond when the economy demanded more higher education after Indian independence in 1947. India averaged an annual income per-capita growth rate of 0.34% between 1900-01 and 1943-44. Agriculture was the key sector employing just over 75% of the workforce (Sivasubramonian, 1997). Demand for higher education was likely low under such conditions. In comparison growth rates of income per-capita averaged 1.5% per year in the 1960s increasing to over 3% in the 1980s and beyond. Unsurprisingly, higher education also increased in these decades with the share of matriculates (i.e., individuals that have completed secondary schooling and passed a matriculate exam) increasing from 1.6% in 1961 to 18% in 2001.

We quantitatively show how Catholic missionary districts responded to this national demand for higher education in the second half of the 20^{th} century using a Bartik (1991) approach. Specifically, we construct a panel dataset of higher education using decadal census data from 1961 to 2001. We construct a Bartik-style variable for education demand shocks following Autor and Duggan (2003), which excludes each district from its calculation of the national growth rate in higher education. We then interact the Catholic missionary indicator with this Bartik demand shock. Controlling for year fixed effects, district fixed effects and Bartik demand, the interaction between Catholic missionary and the Bartik demand shock is positive and significant. Thus, when national demand for higher education increased, the increase in the share of higher education was larger in places that were exposed to Catholic missionaries in 1911. One reason Catholic missionary districts were better poised to respond was the large presence of Catholic colleges in these districts. Using data on the opening of Catholic colleges, we find a positive and significant correlation between Catholic missionaries and the number of Catholic colleges from 1931 to 2001. Reassuringly, we find the decadal change in the number of Catholic colleges is statistically unrelated to factors such as urbanisation, that are likely to be correlated with demand for education. Taken together, these findings suggest the historical network of Catholic missionaries was a natural platform from which Catholic influence on higher education via Catholic colleges radiated out accounting for our strong first stage.[7]

Apart from a strong first stage, our instrument also has to satisfy the exclusion restriction, i.e., the historical location of Catholic missionaries has to be unrelated to any factor that may impact the subsequent development of districts other than through current higher education. If Catholic missionaries historically located in richer and more educated districts, then our instrument would be invalid. 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 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 3, these quantitative results match historical accounts of Catholic missionaries locating in Portuguese settlements along the coast and in districts with large tribal populations.

Using Catholic missionary location as an instrument, we find a positive and statistically significant effect of higher education on current income proxied by light density at night. A standard deviation increase in the share of the population with higher education increases log light density by 0.51, an economic effect of 12% given mean log light density of 4.24. Although we find significant and positive IV estimates on higher education, one may still be concerned about Catholic missionaries impacting current income via non-education channels.

The existing literature suggests many such 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

[7] Qualitative evidence also suggests Catholics benefitted from 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. 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 colonisation 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. Finally, an early emphasis on Indianising the clergy generated a large pool of Indian priests and nuns giving the Catholic church a unique advantage in post-independence India (Frykenberg, 2008).

(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 (2015) 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, confidence in local institutions, the share of migrants, and the historical presence of Europeans (a proxy for openness to foreign ideas and people, and tolerance more generally). The coefficient on higher education remains positive and statistically significant across these specifications.

We conduct several tests that indicate the model does not suffer from a weak instruments problem. The coefficient on higher education, however, could be picking up the effects of lower levels of education. To address this concern, we perform two tests. First, we control for the share of the adult population with only a primary or middle school education directly in the regressions. We find a negative and insignificant coefficient on this lower level of education. The results on the share of higher education are unchanged. Second, we include additional instruments for the share of individuals with lower levels of education such as the historical share of primary school only completion and an indicator for Protestant missionaries. Acemoglu *et al.* (2014) among others have used similar instruments. We find strong first stage results, but again no significant effect of lower levels of education on light density. This strongly suggests the effects of higher education on income are not driven by lower levels of education.

Finally, we use other proxies for development such as local GDP per-capita, which show that the mechanism from higher education to development operates via the service sector.[8] We find a positive and significant coefficient on higher education for district GDP per-capita similar to our light density results. When we decompose GDP per-capita by sector, we find small and insignificant coefficients on primary (agriculture) and secondary (industry) sector GDP per-capita. But, we find a positive and significant coefficient on tertiary (service) sector GDP per-capita. The service sector employs a larger number of people with higher education. Hence, these results highlight the importance of higher education in increasing service sector GDP per-capita and local incomes.

Apart from the human capital and development literature, our paper contributes to

[8]These measures have been constructed by a private Indian company, Indicus Analytics, and are not without problems. According to Himanshu (2009), there is divergence in the ranking of districts based on economic development indicators between the Indicus data and other official and publicly available data sets (for example, the National Sample Surveys). While Indicus has offered a rebuttal, using these data is still controversial and hence we do not rely on them for our main results.

three other literatures. First, 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. 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. While these papers have established important correlations between income and education around the world, we extend the literature by identifying the effect of higher education on income using sub-national variation across districts in India.

Second, our paper contributes to a burgeoning literature on how history, in particular colonisation, influences current outcomes. Acemoglu *et al.* (2001, 2002) argue that colonies with a more favourable disease environment encouraged more settlement of European colonisers and promoted institutions protecting private property rights. Sokoloff and Engerman (2000) also focus on historical institutional development, but argue that factor endowments such as the land-labor ratio shaped 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 (2016) compute a new measure of the share of European population during the early stages of colonisation. Their findings are in line with the Glaeser *et al.* (2004) view.[9] In our case, history influences the present through the location of Catholic missionaries that were absent from the provision of education in the past but were better placed to respond when the economy demanded higher education after Indian independence.

Third 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 literacy.[10] For example, Becker and Woessmann (2009) find that Protestants had a strong effect on literacy in 19^{th} century Prussia.[11] Mantovanelli (2014) argues that Protestant

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

[10] Other religious communities also have positive effects on the accumulation of human capital. In 1374 Geert Groote founded the Brethren of the Common Life (BCL) in the Netherlands to start a reform of the Roman Catholic Church from within. With the aim that religious people were able to read the Bible and other religious books, the BCL stimulated the accumulation of human capital and increased literacy rates, which contributed to the early development of the Netherlands (Akcomak *et al.*, 2016).

[11] Using data on many cities of the Holy Roman Empire during the years 1300-1900, Cantoni (2015)

missionaries can account for current differences in literacy across India.[12] 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.[13] 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. In contrast, our findings are similar to Neal (1997) that finds positive effects of Catholic schools on high school and college graduation rates in the United States.

Our paper offers a fresh perspective to this literature. First, we identify the relationship between higher education and development. We find that only historic Catholic missionaries are correlated with current higher education in India and not Protestant missionaries. The impact of Catholic missionaries on the supply of higher education increased over the second half of the 20^{th} century as Catholics established colleges across India. Catholic emphasis on higher education is perhaps unsurprising and matches accounts of Catholics, especially the Jesuits, leading the growth of higher education in other parts of the world (Codina 2000). Second, most studies that analyse the long-term consequences of Christianity have focused on Africa and South America (Cage and Rueda, 2016; Caicedo, 2014). Unlike these countries, the Christian population is a minority in India, with only 2 percent of the population 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.

The structure of the paper is as follows. The next section describes the data. In section 2 we present the OLS results and discuss potential biases. Section 3 discusses the instrumental variables strategy. Section 4 presents the main IV results. We describe several robustness checks in section 5, and conclude in section 6.

---

finds no significant difference in the long run growth performance between Catholic and Protestant cities.

[12] As we note in Section 4, Mantovanelli's (2014) results appear to be driven by Protestant missionaries influencing middle school completion. Protestant missionaries do not have a strong correlation with other levels of education.

[13] 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.

1 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.[14] 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.[15]

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.[16] The data on night-light luminosity is recorded worldwide for every pixel by the Operational Linescan System flown on the Defence 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 of 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 district area to calculate light density at night.[17]

[14] 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.

[15] 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.

[16] There are some caveats with the use of night light data. The first problem is censor saturation. As explained below, we address this concern by using radiance calibrated data. Another problem is blooming, since lights tend to look larger over bright surfaces such as water and snow. Chen and Nordhaus (2011) note some of these problems but argue that luminosity is still useful for regional analysis especially when income data are poor.

[17] 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 radiance data, allows measurement of luminosity till 255 and is less susceptible to the saturation criticism. However, it

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 colours 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 higher 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 the highest level of education (i.e., college education) is not censored by age. Using the 2001 census of India, we construct current district-level demographic and education variables.[18] The main independent variable of interest is the share of population 25 years and above who have completed secondary education,[19] higher secondary education or have obtained college degrees or diplomas. We refer to this collective group of individuals with a secondary and post-secondary education as *higher education*. Although 18.25% of the Indian population over 25 has completed higher education, the district-level percentage ranges from 4.6% to 49.8%. *Figure* 2 shows the spatial distribution of completed higher education across Indian districts. Analogous to the figure on luminosity, a lighter colour represents a higher share of adults with higher education. While South India has a larger proportion of people with higher education, again there is significant heterogeneity within states.

The unconditional correlation between district log light density and share of individuals 25 and over in 2001 with higher education is 0.54. *Figure* 3 shows the scatter plot between the two variables. While the picture suggests a large and positive correlation between luminosity and higher 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

is sporadically available and hence less used in research that focuses on growth of luminosity. Since this paper uses cross sectional data, we use the calibrated radiance data set. We use GIS tools to extract luminosity from the raster files provided by DMSP.

[18]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.

[19]Completing secondary education refers to passing Class X exams. Matriculation requires students to pass a country/state wide exam conducted by an all India or state board of examination.

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.[20] 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 education and income.[21] We focus on several historical variables: the urban population share in 1931, the tribal population share in 1931, the population share of Brahmans in 1931, the districts that were historically a part of Princely India, and the presence of railways in 1909.[22] Brahmans typically occupy the top position in the Indian caste system. Although they traditionally worked as priests and teachers, Brahmans also were disproportionately represented among landowners, lawyers and other elite occupations in the colonial era. Thus, the Brahman population share may independently influence both subsequent higher 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

[20]The million plus cities as of the 2001 census are Ahmedabad, Bangalore, Chennai, Delhi, Hyderabad, Jaipur, Kolkata, Mumbai, Pune and Surat.

[21]We use historical data in 1931 because it is the first census year that provides comprehensive and reliable data on these variables for British India and the Princely States. See Section 3 for details on problems with the historical data for the Princely States before 1931.

[22]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 district boundary changes over time.

statistics for the main variables in $Table$ 1.[23]

2 Higher Education and Luminosity: OLS Estimates

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

$$y_{ds} = \alpha_s + \theta higheduc_{ds} + \rho' C_{ds} + \pi' G_{ds} + \delta' H_{ds} + \varepsilon_{ds}. \tag{1}$$

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 *higheduc* 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.[24] We model y as a function of time-invariant characteristics with the vector 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 urbanisation, that are not included in our specification due to potential endogeneity.[25] We estimate the model using robust standard errors.[26]

[23] We describe the data sources in Appendix Table A3: Data Appendix.

[24] 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 (2013) for excellent surveys on the literature.

[25] In Section 5 we analyse the robustness of the results to additional contemporaneous controls.

[26] 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

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 higher education. Unsurprisingly, the addition of contemporaneous and geographic controls reduces the magnitude of the coefficient on higher education. The marginal effect of increasing the share of higher education by one standard deviation (0.074) in the model with no controls (*column* (1)) increases the log light density by 0.51 (mean value is 4.24). The marginal effect drops to 0.35 once we include all controls (*column* (4)). This is an economically significant effect on the order of 8% against the mean of log light density.

One concern with using only the share of higher education is that higher education may differ across districts because of differences in primary and middle schooling or illiteracy. Hence, the coefficient on higher education may be picking up the effect of other levels of schooling. To address this concern, we directly control for the population share with at least primary schooling but no higher than middle school education in *column* (5). For ease of exposition, we refer to this variable as the share of primary schooling in the tables and text. The coefficient on higher education remains positive and statistically significant at the 1 percent level. In fact, the magnitude of the coefficient is almost the same in *columns* (4) and (5). This suggests the findings on higher education are not driven by omitting other levels of schooling.

Our results thus far find a strong correlation between the population share with higher 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 higher education is likely to be endogenous. Reverse causality is a serious concern if individuals with higher education move to districts with higher income or higher light density. Although inter-district migration is low in India (Munshi and Rozensweig, 2009), less is known about the migration of higher 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.

3 Catholic Missionaries as an Instrument

Our empirical exercise uses the location of Catholic missionaries in the early 20$^{\text{th}}$ century as an instrument for the 2001 share of the adult population with higher education. In this

estimated p-values from wild bootstrap suggested by Cameron *et al.* (2008) as a solution to the problem of few clusters. Our results on higher education are statistically significant.

section we discuss why the historical location of Catholic missionaries is a good instrument and assess potential selection concerns regarding their choice of location.

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 (*See Figure A*1).[27] 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* 4 displays Catholic locations after this exercise.

As seen Catholic missionaries were located all over India, but with a larger concentration in the South. 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, a south western state with many Catholic missionaries (see *Figure A*1).[28]

Barring an occasional mention of Catholic priests, church history in India is silent till the arrival of Vasco de Gama and the Portuguese 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*.[29] 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 centre of Portuguese Catholic Church hierarchy. Missionaries also settled early

[27] 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.

[28] 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.

[29] 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 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.

in Portuguese strongholds such as Daman, Diu, Vasai (suburban area north of Mumbai), and Mumbai along the coast.

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 Brahmans 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 centres 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.

Against this backdrop, it is no surprise that a majority of Catholic missionaries located along the coast in former Portuguese colonies. But, a sizeable number also located 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.[30] 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 kilometre, which is likely to be correlated with the historical presence of Christians in a district.[31]

[30] 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 Indian Catholic Church hierarchy in 1885. Not many Catholic missions were established in the early to mid-19$^{\text{th}}$ century and Catholic missionaries were mostly dormant in this period. The older Portuguese missions were set up in the 16$^{\text{th}}$ and 17$^{\text{th}}$ centuries.

[31] We check the robustness of the results using the intensive margin in Table 8.

The presence of Catholic missionaries has to satisfy two conditions to be a valid instrument. First, the instrument has to be correlated with higher education in 2001. In the next section we provide compelling evidence that districts with a historical presence of Catholic missionaries have significantly more higher education in 2001. Further, we also show that 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 could violate the exclusion restriction in two ways. Catholic missionaries may have located in richer or more educated areas that could independently impact current development. And, the exclusion restriction could also be violated if the location of Catholic missionaries impacts contemporary luminosity through channels other than contemporary higher education. In this section we analyse the endogenous location of Catholic missionaries. In Section 5 we examine other channels by which Catholic missionaries could affect 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* 3 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 Brahmans and income tax revenues per-capita, a proxy for income. We also find the location of Catholic missionaries is uncorrelated with the number of schools in 1901. Primary and secondary education in this time came under the purview of local district boards. By most accounts the number of privately managed schools declined as public funding for education increased in the 1920s and 1930s (Chaudhary, 2010).

We find a significant but negative coefficient on the number of colleges, which is also consistent with the historical record. Catholic missionaries were less involved in college education relative to say Protestant missionaries. For example, there were only 9 Catholic colleges (5%) in 1911 compared to 40 Protestant (33%).[32] The remaining 123 colleges (71%) were managed by either private Indian committees or the colonial Government of India (Great Britain, 1914). These correlations are suggestive that the location of Catholic missionaries was not systematically correlated with historical measures of income or education. Taken together, the results indicate the location of Catholic missionaries was not completely random. However, Catholic missionaries did not follow a uniform strategy of positive selection, given that missionaries located in coastal areas and in districts with a higher share of tribal groups.

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$^{\text{th}}$ 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.[33] We use the 1931 census because it has the most detailed and accurate information on the Princely States and districts of British India. Migration and urbanisation 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 Brahmans, urbanised 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. 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.

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

[33] 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 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.

4 Higher Education and Luminosity: IV Estimates

4.1 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 higher education is 0.029. Thus, the share of individuals with higher education is 16% higher in districts with a Catholic missionary given the mean share of higher education of 0.18. This rather large point estimate is highly significant. Reassuring, we do not suffer from a weak instrument problem; the *F stat* is 15.52 and the Cragg-Donald Wald statistic is 19.84.[34]

Second stage results are displayed in *Panel B*. The IV coefficient on higher education is positive and significant at the 1 percent level with a magnitude of 6.92 (*column* (4)). A standard deviation increase in higher education raises log light density by 0.51, an economic effect of 12% given mean log light density. A standard deviation of higher education is equivalent to a 7.4 percentage point greater higher education completion rate, as compared to the omitted share. The omitted share, by construction, is a combination of those who are illiterate, those with some primary and middle schooling.

As noted in our OLS results, one may argue the share of higher education is picking up the impact of other levels of schooling. In *column* (5) we address this issue by including the share of primary schooling. As described earlier, this variable includes the population with completed primary schooling but no higher than middle schooling. The IV coefficient on higher education is positive and significant, while the coefficient on primary schooling is not statistically significant. Moreover, the coefficient on higher education is similar in magnitude to the estimate in *column* (4), which strongly suggests that omitting lowers level of education is not driving our results on higher education.

[34] 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.

4.2 Endogeneity of Lower Levels of Education

Although we control for the share of the population with primary schooling in *Table* 4, it raises a problem because this share is not exogenous. Moreover, the positive correlation between primary and higher education may render the coefficient on primary education insignificant in *column* (5). Hence, we need an instrument for primary schooling. The existing literature offers a few candidates. For example, Acemoglu *et al.* (2014) use the historical location of Protestant missionaries and historical primary school enrolment rates as instruments for average years of schooling.

Many studies have found a strong correlation between Protestant missionaries and literacy (Becker and Woessman, 2008; Mantovanelli, 2014). The principle of "Sola Scriptura" underpins this relationship 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. Following this literature we create an indicator for the location of Protestant missionaries as of 1908 using information in the Statistical Atlas of Christian Missions (1910).

We observe Protestant missionaries in 58% of the districts. Similar to *Table* 3, we ran regressions on the location of Protestant missionaries and historical characteristics. Results, reported in *Appendix Table A*1, show the coastal indicator has a large predictive power on the location of missionaries. Protestants also set up missions in more ethnically diverse districts and those at higher elevation. The correlation between Protestant missionaries and the provision of education is again insignificant. However, using Protestant missions as an instrument does not render significant results in our setting because their location is only weakly correlated with current education outcomes.[35] Thus, it leads to a weak instrument problem.

Unlike Protestant location, historical primary school attainment is strongly correlated with current education outcomes in our context. Acemoglu *et al.* (2014) argue that historical educational enrolment is a suitable instrument for average years of schooling in their cross-country context because the variation in enrolment was often driven by idiosyncratic preferences of individual leaders. Building on their argument, we use the

[35]Interestingly, Protestant missionary location is significantly correlated only with the share of the population with completed middle schooling. Protestant location is statistically uncorrelated with the share of individuals with at most primary schooling. This suggests that Mantovanelli's (2014) results on Protestant missionary location and current literacy arise from middle school completion, rather than just primary school completion.

share of individuals with only primary school completion in 1961 as an instrument for the 2001 share of primary schooling.[36] In the first two decades following Indian independence, public spending on primary education followed idiosyncratic colonial policies (Chaudhary and Garg, 2015). Beginning in the 1970s, however, the Indian government embarked on an ambitious public goods program increasing the number of primary schools in previously under-served communities (Banerjee and Somanathan, 2007). Hence, the variation in primary school completion as of 1961 stems from historical variation that is unrelated to contemporary policy variation and demand conditions.[37]

We present the results instrumenting for both higher education and primary education in *Table* 5. In *column* (1) we show the single instrument results from *Table* 4, *column* (4) for reference. In *column* (2) we instrument for both higher education and primary education using the location of Catholic Missionaries and the share of primary school completion in 1961. Using additional instruments does not fundamentally change the first stage results on the share of higher education; Catholic missionaries continue to be significantly correlated with higher education in 2001, as seen in *Panel A*. In contrast, *Panel B* shows there is no significant relationship between Catholic missionaries and primary education in 2001. Only the historical primary school completion rate is strongly correlated with the contemporary share of primary education. A standard deviation increase (0.046) in the historical rate leads to a 2.3 percentage point increase in the contemporary share of primary education. This translates into an economic effect of 11% given the mean share of 2001 primary education. The second stage results, in *Panel C,* confirm that lower levels of education are uncorrelated with light density. Similar to the OLS estimates, the coefficient on the share of primary education is negative and not statistically significant. This specification passes the weak instruments test at the 20% level. Moreover, the Kleibergen-Paap rk LM statistic value is significant at 1%, rejecting the null that the system is under-identified.

The fundamental relationship between higher education and light density does not change if we use Protestant missionary as an instrument (see *columns* (3) and (4)).[38] Across these specifications, the coefficient on the share of higher education remains signif-

[36] Ideally we want to use primary school enrolment. But these data are unavailable at the district-level. The 1961 census offers the first complete enumeration of primary school completion at the district-level.

[37] We recognise using lagged values as instruments is not without problems. Our main goal here is to just show the results on higher education are robust to controlling and instrumenting for primary education.

[38] In *column* (4), where we include three instruments for two endogenous variables, we cannot reject the Hansen over identification test that our instruments are valid. Our Hansen J statistics is 0.00 while the p-value is 0.9977.

icant and similar in magnitude to *column* (1). Interestingly, these results also suggest our Catholic instrument is not picking up the impact of Christian missionaries more generally. Catholic missionary location is significantly correlated with higher education even after controlling for Protestant location.[39]

Given the robustness of the effect of Catholic missionaries on higher education and the estimated coefficient on higher education in the second stage, in the rest of the paper we focus on the model with only the higher education share instrumented with Catholic missionary location. All the robustness results presented below go through when we control for contemporary primary education and use historical 1961 primary school completion as the additional instrument.

5 Is Catholic Missionary Location a Plausible Instrument?

5.1 Mechanism: Catholic Missionaries to Catholic Colleges

Why does the 1911 location of Catholic missionaries have such an enduring impact on contemporary higher education in India? One explanation is that districts with historic Catholic missionaries were better placed to respond when the demand for higher education increased in the second half of the 20^{th} century. By most accounts the Indian economy was largely stagnant in the first half of the 20^{th} century under colonial rule (Sivasubramonian, 1997). Total national income per-capita grew at just 0.34% per year from 1900 to 1944. A majority of the population was engaged in agriculture (Sivasubramonian, 1997). And, demand for higher education was likely low under such circumstances. Growth picked up after independence in 1947 with GDP per-capita growth averaging 1.5% per year in the 1960s and then over 3% per year in the 1980s and 1990s (Acharya *et al.,* 2003). Against this background the demand for higher education was likely higher in the decades after independence than before.

In fact, Indian independence was an important and positive transition for Catholics. Although the British colonial state had a stated policy of religious neutrality, Protestant missionaries probably received more favourable treatment compared to Catholics. A larger Protestant presence at the post-secondary level in colonial India does suggest some public favouritism. More importantly, the mass departure of Europeans in 1947 did not leave Catholic institutions without leaders because they had been actively promoting Indian clergy from the 19^{th} century (Frykenberg, 2008). India today has 3,851 Jesuit priests,

[39] The use of Protestant missionary location as the instrument for lower levels of schooling leads to a weak instruments problem, evidenced by the lower values of the Cragg-Donald statistics.

more than any other country, despite the fact that less than 2% of the population is Catholic (Frykenberg, 2008). Similar to the enduring impact of Jesuits on human capital in other parts of the world (Caicedo, 2014), many of the Jesuit schools and colleges are among the best education institutions in India.

To test whether districts with a Catholic missionary presence responded favourably when demand for higher education increased, we need a measure of an exogenous demand shock in higher education. Bartik (1991) offers such a measure. To calculate this measure and validate our first stage relationship between Catholic missionary location and current higher education we need a panel specification. Hence, we construct a district-level panel dataset using the 1961, 1971, 1981, 1991 and 2001 census data on the share of matriculates as the dependent variable.[40]

Following Autor and Duggan (2003) that build on the original Bartik (1991) demand shock variable, we construct the predicted log change in the share of higher education between year *t* and *t-1* as the *t-1* district share of higher education multiplied by the log change in the national share of higher education between *t* and *t-1*, excluding the individual district's contribution from its national share. Here *t-1* refers to the earlier census year so it is actually *t-10* years. We then interact this predicted change in higher education with the Catholic missionary indicator. The panel regression controls for the Bartik demand shock and includes district and year fixed effects. Since we include district fixed effects, we cannot include the Catholic missionary indicator because it does not vary over time. Reassuringly, the interaction between Catholic missionary and the demand shock in higher education is positive and statistically significant, as seen in *Table 6A*. Once the aggregate demand for higher education increased in the second half of the 20^{th} century, the increment in the share of higher education was larger in places that were exposed to Catholic missionaries in 1911.

Why were districts with former Catholic missionaries better placed to respond? One reason is Catholic colleges are concentrated in districts with a historical Catholic missionary presence. Over the 20^{th} century the number of degree granting Catholic colleges increased from 9 in 1911 to 255 in 2001. Given their concentration in former Catholic missionary districts, an increase in the supply of Catholic colleges is likely driving the Catholic impact on current higher education.[41] We support this conjecture using new

[40] We use the share of matriculates as our measure of higher education here because it is consistently reported between 1961 and 2001.

[41] A higher supply of Catholic colleges may reduce the cost of investing in tertiary education and therefore create incentives to complete secondary education as well. Consistent with this idea, Card (1995) shows that the presence of a nearby 4-year college has a strong effect on completed education.

data on Catholic colleges.

Using the address and date of opening of Catholic colleges reported in the Catholic Directory of India for 2010, we construct the number of degree granting Catholic colleges in each district for the years 1911, 1931, 1951, 1971, 1991 and 2001.[42] In *Table 6B* we study the impact of Catholic missionaries on these colleges. In *column* (1) the coefficient on Catholic missionaries is small and not statistically significant. This is unsurprising because as noted earlier there were only 9 Catholic colleges in India, accounting for 5% of the total in 1911. From 1931 onwards, the coefficients are larger, positive and statistically significant, suggesting that those areas that were exposed to the presence of Catholic missionaries in 1911 have more Catholic colleges.[43]

While we are making a supply side argument for the growth in Catholic colleges, it is possible the expansion of Catholic colleges was driven by demand side factors particularly in the second half of the 20^{th} century when income and hence demand for education picked up. We address this concern partially in Table 6A by showing that places that were exposed to Catholic missionaries responded better when the national demand for higher education increased. Nonetheless, it is possible that the increase in the demand for education was higher in richer districts.

To test the relationship between income and expansion of Catholic colleges, we need estimates of regional income. Since these are unavailable, we use district level urbanisation rates as a proxy for local income. Specifically, we regress the change in the number of Catholic colleges in each decade beginning in 1951 as the dependent variable against the baseline urbanisation share. In addition, we control for the geographic controls, the historical controls, population, the share of Scheduled Castes and Tribes at the beginning of each decade, and the baseline number of Catholic colleges in each decade. *Table 6C* reports these results. Reassuringly, we find that income, as proxied by urbanisation, is not correlated with the expansion in Catholic colleges. We find the baseline number of Catholic colleges is strongly correlated with the absolute increase in Catholic colleges in each decade. This supports a supply side interpretation that Catholic colleges expanded in areas with existing Catholic colleges, which in turn were located in districts with former Catholic missionaries.

[42] We do not have similar information on the time series of secondary and higher secondary schools.

[43] Due to data availability, the number of observations is less in 1911. We have run the regressions for all the years with the reduced sample and the results indicate the findings are not driven by sample selection.

5.2 Exclusion Restriction

Our exclusion restriction assumes that Catholic missionaries impact current light density only via the current adult population share of individuals with higher education. In this sub-section we 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.[44] In *Table* 7, *column* (1) we include infant mortality, our measure of health, as an additional control. The coefficient on higher education remains large and statistically significant.[45] We also test for a direct relationship between infant mortality and the location of Catholic missionaries. These results showing the direct impact of Catholic missionaries on alternative channels are displayed in *Appendix Table A*2. *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 higher 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 higher education is unchanged.[46]

[44] Recent studies have analysed the relationship between religion and health in India. Calvi and Mantovanelli (2015) 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 and McQueeney (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. In the case of Africa, Cage and Rueda (2017) find that proximity to a Protestant or Catholic mission that invested in health decreases contemporaneous HIV prevalence. They do not find, however, a significant relationship between proximity to missions that invested in health and the number of years of schooling.

[45] We find a similar coefficient on higher education in this smaller sample in our standard IV estimation.

[46] 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

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 historic Catholic missionaries. If infrastructure is correlated with education and fosters development, the omission of infrastructure could bias the estimated coefficient. To capture infrastructure, we include the proportion of villages with paved roads, the share of land that is irrigated, and the number of schools per village in the district. Even after including these controls, higher education has a positive and highly significant coefficient.[47]

In *column* (4) we control for the share of inter-district migrants to the total population over 25 in the district. Although migration between states is not too widespread in India (Munshi and Rosenzweig, 2009), we worry people may migrate in larger numbers to districts with a historical Catholic presence because these districts are more welcoming to outsiders and have better economic opportunities. 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 higher education is robust to controlling for migration.[48]

In *columns* (5) - (7) we include proxies for local institutions or culture. Although we always exploit within state variation by including state fixed effects, there are local differences in culture such as stronger cultural norms promoting lower violent crime, for example, that could be correlated with historical Catholic missionary presence and current income. We proxy for institutions with confidence in local courts and village panchayats (Indian village councils with five members normally).[49] We use data from the Human

castes and tribes were more likely to convert to Christianity in India. Our IV results on higher education are unchanged if we include the share of Christians in 1931 as a control.

[47]One may be concerned that districts with a historical Catholic missionary presence were better served by private schools on account of more self-reliance or private entrepreneurship. And, these characteristics may independently influence income. We test these possibilities using information on private schools and measures of learning. Using district-level data on total private schools for 2005-06, we find an insignificant coefficient on Catholic missionary location (p-value of 0.73). In addition, using data on the quality of education for 2008, we find no significant correlation between Catholic missionary location and the average proportion of children aged 12-16 (middle school cohorts) proficient in simple math and vocabulary tests. These tests were conducted in rural India by an Indian NGO (Pratham).

[48]Our results in *Appendix Table A2* suggest that inter-district migration is uncorrelated to Catholic missionary district.

[49]Standard measure of institutions, such as rule of law, are unavailable at the district level. Kaufmann *et al.* (2011) define rule of law as "capturing perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence." Given this definition, one could

Development Survey, where respondents were asked if they feel very confident in these institutions, and aggregate the individual responses to the district-level. We interpret these responses as a measure of the quality of institutions.[50] Controlling for local institutions and culture does not change the results; the coefficient on higher education is similar in magnitude and significance.

In *column* (8) 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 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 higher education is essentially unchanged when we control for the number of Europeans in a district in 1931.[51]

Finally, in *column* (9) we control for the historical population density of a district in 1931. While we always include the 1931 urbanisation 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 higher education.

In *Table* 8 we present additional exercises to check the robustness of our findings. In *column* (1) we control for lagged values of primary education. The idea here is to further control for potential reverse causality. It could be the case that an increase in primary education a few years ago promoted economic growth that in turn lead to an increase in the demand for higher education. Reassuringly, controlling for the lagged value of primary education does not affect the results; the coefficient on the share of higher education is positive, similar in magnitude and statistically significant. In *column* (2) we check the robustness of the results to the intensive form of Catholic missionaries per square km as the instrument, as opposed to the simple indicator for Catholic missionary presence. The

---

interpret confidence in local courts and village panchayats as an imperfect proxy for rule of law at the local level.

[50]The data on local courts and village panchayats are based on the 2004-05 round of the Indian Human Development Survey conducted by the National Council for Applied Economic Research in Delhi and the University of Maryland. These cover a smaller number of districts and while they are strictly not representative of districts, we use them as proxies as the sample sizes are large (41,554 households). The data on violent crimes is from the National Crime Research Bureau, Government of India.

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

coefficient on the share of higher education is similar.

As our results could be driven by atypical observations or extreme values, we check the robustness of our results to the presence of outliers. We first drop districts with light density above the 99^{th} percentile and below the 1^{st} percentile in *column* (3). In *column* (4) we exclude districts with million plus population cities. In *column* (5) we drop districts that include the state capital and in *column* (6) we exclude districts that contain the largest city in the state. Across these specifications, the coefficient on higher education remains positive and statistically significant. In *column* (7) we test if our IV results are driven by a group of compliers. We drop districts that experienced the highest growth in higher education between 1961 and 2001, namely those above the 90^{th} percentile. The results are unchanged when we drop this group of districts.

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. In *column* (8), we drop the state of Kerala. Again, the results on higher education are essentially unchanged.[52]

5.3 Different Measures of Development

In this last sub-section, we use different proxies for development. The literature analysing district-level development in India has been scarce due to a lack 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.

In *Table* 9 we study the effect of higher education on alternative measures of development using our standard IV specification. 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 another measure of lights. *Column* (1) shows the IV estimates on higher education for lights per-capita controlling for current, geographical, and historical variables as before. The coefficient on

[52]In results unreported here, we also checked whether one particular state may be responsible for the results by dropping one state at a time. The results were unchanged. And, we also conducted a matching exercise using the geographic and historical variables. Our results were similar to the IV results.

higher education is positive and statistically significant.[53]

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 higher education on per-capita consumption using data on consumption expenditures from the National Sample Survey Organisation. These data come from quinquennial surveys of a randomly selected sample of households conducted by the Ministry of Statistics and Programme Implementation.[54] The results, displayed in *column* (2), show a positive and highly significant coefficient on higher education. In quantitative terms, a percentage point increase in the population share of individuals with higher education increases consumption per-capita by 1.87 percent.[55] A rough back of the envelope exercise using consumption per-capita suggests a range of macro returns to higher education, which include the Mincer returns found in studies on India (Colglough *et al.*, 2010; Agrawal, 2012; Montenegro and Patrinos, 2014).

Apart from consumption expenditures, there is another measure of district-level GDP. Indicus Analytics, a research firm based in Delhi, has estimated district level GDP per-capita using expenditure and savings habits of households, characteristics and occupations.[56] The correlation between light density and Indicus GDP per-capita for the same year is 0.57.

In *columns* (3) to (6), we study the relationship between current higher education

[53] 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.

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

[55] As with luminosity, we are unable 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 (2016) also suggest that night lights come closer to measures of income per-capita per national accounts compared to survey estimates.

[56] Indicus uses survey data from NSSO surveys, National Data Survey of Savings Patterns of Indians (NDSPPI), District Level Household Survey, Census of India, RBI dataset and other sources. The quality of the data computed by Indicus, however, is under debate. See, for example, the concerns about the data raised by Himanshu (2009) and the reply by Bhandari (2009).

and the Indicus measure of local GDP per-capita (total and by sector). We find that more higher education leads to higher GDP per-capita with the effect primarily operating through income in the service sector. The coefficient on higher education is small and statistically insignificant for agriculture and industry sector GDP per-capita. In theory, missionaries could have transferred the technology from their countries of origin. And, the use of more advanced irrigation techniques or fertilisers could directly affect development by increasing agricultural productivity in the districts where missionaries located. Likewise, missionaries could have promoted the skills of the people in manual work and could have created the basis for development in the industry sector. The insignificant results in *columns* (4) and (5) suggest this was not the case. The result in *column* (6) on service sector GDP is also consistent with Castelló-Climent and Mukhopadhyay (2013) that find a large effect of higher education, namely tertiary education in their case, on the growth of the service sector. These findings also support the broad macro-patterns on the importance of the tertiary/service sector for India's recent growth and development (see Eichengreen and Gupta, 2011).

6 Conclusion

This paper investigates if higher levels of education lead to higher levels of development. Although the literature has studied this question extensively, the evidence thus far has been mixed because of common identification problems of 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).[57] Moreover, education can also be correlated with omitted variables that influence development.

We use history to address this problem by exploiting the location of Catholic missionaries in colonial India. This is similar in spirit to the empirical literature that looks at history to identify the causal impact of institutions on development (for example, Acemoglu *et al.*, 2000, 2001). Christian missionaries in general, and Catholic missionaries in particular are associated with high quality education around the world. However, the novelty of our approach is to focus on India, a country where Christians form a small minority and have relatively limited influence over other institutions correlated with development.

[57]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.

We use district-level data with state fixed effects to control for institutions and cultural characteristics. In our analysis, we focus on higher education because this level of education is particularly important in the Indian context. It is the main educational variable correlated with economic growth at the state level (Castelló-Climent and Mukhopadhyay, 2013).[58] The focus on one country and a particular level of education is therefore part of our strategy to better identify the effect of education on development.

We find that exposure to Catholic missionaries at the beginning of the 20th century has had a long-term impact on the current composition of education. 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 over the 20^{th} century. Using the location of Catholic missionaries in 1911 as an instrument, we find a strong and positive effect of the adult population share with higher education on contemporary levels of development, as measured by night light density.

A broad array of sensitivity analyses indicate that it is unlikely that Catholic missionaries influenced current development through channels other than higher education. We also show the results hold with a rich set of geographical, historical and current controls. The effects of higher 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.

Our main goal in this paper has been to identify the causal effect of higher 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. One explanation is perhaps the poor quality of education at the primary level. Future research may want to analyse whether the link from higher education to development is specific to India or more general. This will provide policy makers with better tools to assess if, for example, the quality of current primary schooling is sufficient to promote development or to design policies to increase those levels of education that are more growth and development enhancing.

[58] 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).

Amparo Castelló-Climent: *University of Valencia*
Latika Chaudhary: *Naval Postgraduate School*
Abhiroop Mukhopadhyay: *Indian Statistical Institute*

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 and 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*, vol. 91, pp. 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*, vol 118, pp. 1231-94.

[5] Acharya, S., I. Ahluwalia, K. L. Krishna and I. Patnaik (2003). 'India: economic growth, 1950-2000', in Kirit S. Parikh ed., Explaining Growth in South Asia, New Delhi: Oxford University Press.

[6] 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.

[7] Agrawal, T (2012). 'Returns to education in India: some recent evidence', *Journal of Quantitative Economics*, vol. 10(2), pp. 131-51.

[8] Akcomak, S., D, Webbink and B. ter Weel (2016). 'Why did the Netherlands develop so early? the legacy of the Brethren of the Common Life'. *Economic Journal*, vol. 126, pp. 821-60.

[9] Alesina, A., S. Michalopoulos and E. Papaioannou (2016). 'Ethnic inequality', *Journal of Political Economy*, vol. 124(2), pp. 428-88 .

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

[11] 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.

[12] Autor, D. and M.G. Duggan (2003). 'The rise in the disability rolls and the decline in unemployment', *Quarterly Journal of Economics*, vol. 118(1), pp. 157-205.

[13] 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 Association*, vol. 13(4), pp. 669-98.

[14] 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.

[15] Banerjee, A. and R. Somanathan (2007). 'The political economy of public goods: some evidence from India', *Journal of Development Economics*, vol. 82(2), pp. 287-314.

[16] Barro, R. (1991). 'Economic Growth in a Cross Section of Countries', *Quarterly Journal of Economics*, vol. 106 (2), pp. 407-43.

[17] Barro, R. and R.M. McCleary (2003). 'Religion and economic growth across countries', *American Sociological Review*, vol. 68(5), pp.760-81.

[18] Bartik, T.J. (1991). *Who Benefits from State and Local Development Policies?* Kalamazoo, Michigan: W.E. UpJohn Institute for Employment Research.

[19] Becker, G. (1964). 'Human capital', 1^{st} ed. New York: Columbia University Press for the National Bureau of Economic Research. 2^{nd} edn. Columbia University Press, New York, 1975 and 3^{rd} 1994.

[20] 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.

[21] 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. 531-96.

[22] Benhabib, J. and M. Spiegel (1994). 'The role of human capital in economic development', *Journal of Monetary Economics*, vol. 34 (2), pp. 143-74.

[23] Besley, T. and R. Burgess (2000). 'Land reform, poverty reduction, and growth: evidence from India', *Quarterly Journal of Economics*, vol. 115(2), pp. 389-430.

[24] Bhandari, L. (2009). 'Socio-economic performance of constituencies: a response', *Economic and Political Weekly,* vol. 44(40), pp. 61-63.

[25] Bharadwaj, P., A. Khwaja and A. Mian (2008). 'The big march: migratory flows after the partition of India', *Economic and Political Weekly,* vol. 43(35), pp.39-49.

[26] Bils, M. and P. Klenow (2000). 'Does schooling cause growth', *American Economic Review*, vol. 90(5), pp.1160-83.

[27] Cagé, J., and V. Rueda (2016). 'The long-term effects of the printing press in sub-Saharan Africa,' *American Economic Journal: Applied Economics*, vol. 8(3), pp. 69–99.

[28] Cagé, J., and V. Rueda (2017). 'Sex and the mission: the conficting effects of early Christian investments on sub-Saharan Africa's HIV epidemic', mimeo, Oxford University.

[29] Caicedo, F.V. (2014). 'The mission: economic persistence, human capital transmission and culture in South America', mimeo Bonn University.

[30] Calvi, R. and F. Mantovanelli (2015). 'Long-term effects of access to health care: medical missionaries in Colonial India', mimeo Boston College.

[31] 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.

[32] Cantoni, D. (2015). 'The economic effects of the Protestant Reformation: testing the Weber hypothesis in the German Lands'. *Journal of the European Economic Association*, vol. 13(4), pp. 561-98.

[33] Cantoni, D. and N. Yuchtman (2014). 'Medieval universities, legal institutions, and the commercial revolution'. *Quarterly Journal of Economics*, vol. 129(2), pp. 823-87.

[34] Card, D. (1995). 'Using geographic variation in college proximity to estimate the return to schooling', in "Aspects of labour market behaviour: essays in honor of John Vanderkamp,' eds. Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky, pp. 201–22, Toronto: University of Toronto Press.

[35] 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. 303-20.

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

[37] Chaudhary, L. (2009). 'Determinants of Primary Schooling in British India', *Journal of Economic History*, vol. 69(1), pp. 269-302.

[38] Chaudhary, L. (2010). 'Land revenues, schools and literacy: a historical examination of public and private funding of education', *Indian Economic and Social History Review*, vol. 47(2), pp. 179-204.

[39] Chaudhary, L. and M. Garg (2015). 'Does history matter? colonial education investments in India', *Economic History Review*, vol. 68 (3), pp. 937-61.

[40] Chen, X. and W. D. Nordhaus (2011). 'Using luminosity data as a proxy for economic statistics', *Proceedings of the National Academy of Sciences*, vol. 108 (21), pp. 8589-94.

[41] Codina, G. S.J. (2000). 'A century of Jesuit education, 1900-2000', in Jesuits: Yearbook of the Society of Jesus.

[42] Cohen, D. and M. Soto (2007). 'Growth and human capital: good data, good results', *Journal of Economic Growth*, vol. 12, pp. 51–76.

[43] Colglough C., G. Kingdon and H. Patrinos (2010). 'The changing pattern of wage returns to education and its implications', *Development Policy Review*, vol. 28(6), pp. 733-47.

[44] 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.

[45] 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*, vol. 4(1), pp. 1-36.

[46] Easterly, W. and R. Levine (2016). 'The European origins of economic development', *Journal of Economic Growth*, vol. 21(3), pp. 225-57.

[47] Eichengreen, B. and P. Gupta (2011). 'The service sector as India's road to economic growth', NBER Working Paper Nº 16757.

[48] Frykenberg, R. E. (2008). 'Christianity in India: from beginnings to the present', Oxford University Press.

[49] 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.

[50] 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.

[51] 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.

[52] Goldin, C. and L. F. Katz. (2008). *The Race between Education and Technology.* Cambridge, Mass: Harvard University Press.

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

[54] 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.

[55] Himanshu (2009). 'Electoral politics and the manipulation of statistics', *Economic & Political Weekly*, vol. 44 (19), pp. 31-35.

[56] Iyer, L. (2010). 'Direct versus indirect Colonial rule in India: long-term consequences', *Review of Economics and Statistics*, vol. 92(4), pp. 693-713.

[57] Kaufmann, D., A. Kraay and M. Mastruzzi (2011). 'The worldwide governance indicators: methodology and analytical issues', *Hague Journal on the Rule of Law*, vol.3(2), pp. 220-46.

[58] Kotwal, A., B. Ramaswami and W. Wadhwa (2011). 'Economic liberalization and Indian economic growth: what's the evidence?', *Journal of Economic Literature*, vol. 49(4), pp. 1152-99.

[59] Krueger, A. and M. Lindahl (2001). 'Education for growth: why and for whom?', *Journal of Economic Literature*, vol. XXXIX, pp. 1101-36.

[60] Kumar, H. and R. Somanathan (2009). 'Mapping Indian districts across census years, 1971-2001', *Economic and Political Weekly* , vol. 36(9), pp. 783-92.

[61] Lucas, R. (1988). 'On the mechanics of economic development', *Journal of Monetary Economics*, vol. 22, pp. 3-42.

[62] Maloney, W.F. and F. Valencia Caicedo (2017). 'Engineering growth: innovative capacity and development in the Americas', CES ifo Working Paper, Nº 6339.

[63] Mankiw, N. G., D. Romer and D. Weil (1992). 'A contribution to the empirics of economic growth', *Quarterly Journal of Economics,* vol. 107, pp. 407-37.

[64] Mantovanelli, F. G. (2014). 'The Protestant legacy: missionaries and literacy in India', mimeo Boston College.

[65] McCleary, R.M. and Barro, R. (2006). 'Religion and economy', *Journal of Economic Perspectives*, vol. 20(2), pp.49-72.

[66] Meisenzahl, R., and J. Mokyr (2012). 'The rate and direction of invention in the British Industrial Revolution: incentives and institutions', in Josh Lerner, and Scott Stern eds., The Rate and Direction of Inventive Activity Revisited, Chapter 9. Chicago: University of Chicago Press, pp. 443-479.

[67] Menon, N. and K. McQueeney (2015). 'Christianity and infant health in India', IZA DP Nº 9177.

[68] Michalopoulos, S. and E. Papaioannou (2013) 'Pre-colonial ethnic institutions and contemporary African development', *Econometrica*, vol. 81(1), pp. 113-52.

[69] Michalopoulos, S. and E. Papaioannou (2014). 'National institutions and subnational development in Africa', *Quarterly Journal of Economics*, pp. 151-213.

[70] Montenegro, C. E. and H. Patrinos (2014). 'Comparable estimates of returns to schooling around the world', Policy Research Working Paper, Nº 7020. The World Bank.

[71] Munshi, K. and M. Rosenzweig (2009). 'Why is mobility in India so low? social insurance, inequality, and growth', NBER Working Paper Nº 14850.

[72] Neal, D. (1997). 'The effects of Catholic secondary schooling on educational achievement', *Journal of Labor Economics*, vol. 15(1), pp. 98-123.

[73] Nelson, R., and E. Phelps (1966). 'Investment in humans, technological difusion, and economic growth', *American Economic Review*, vol. 51(2), pp. 69-75.

[74] Nunn, N. (2010). "Religious conversion in Colonial Africa,' *American Economic Review Papers and Proceedings*, vol. 100 (2), pp. 147-52.

[75] 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.

[76] Nunn, N. (2014). 'Historical development,' *Handbook of Economic Growth*, Chapter Seven, vol. 2A, pp. 347-402.

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

[78] Pinkovskiy, M. and X. Sala-i-Martin (2016). 'Lights, camera,...income! Illuminating the national accounts-household surveys debate', *Quarterly Journal of Economics*, vol. 131(2), pp. 579-631.

[79] Pritchett, L. (2001). 'Where has all the education gone?', *World Bank Economic Review*, vol. 15(3), pp. 367-91.

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

[81] Romer, P. M. (1990). 'Endogenous technological change', *Journal of Political Economy*, vol. 98(5), S71-S102.

[82] Sachs, J. (2003), 'Institutions don't rule: direct effects of geography on per capita income', NBER Working Paper Nº 9490.

[83] Sianesi, B. and J. Van Reenen (2003). 'The returns to education: macroeconomics', *Journal of Economic Surveys*, vol. 17(2), pp. 157-200.

[84] Sivasubramonian (1997). 'Revised estimates of the national income of India, 1900-1901 to 1946-4', *Indian Economic and Social History Review*, vol. 34(2), pp. 113-67.

[85] 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.

[86] Spolaore, E. and R. Wacziarg (2009). 'The diffusion of development', *Quarterly Journal of Economics*, vol. 124(2), pp. 469-529.

[87] Spolaore, E. and R. Wacziarg (2013). 'How deep are the roots of economic development?', *Journal of Economic Literature*, vol. 51(2), pp. 325-69.

[88] Squicciarini, M. P. and N. Voigtländer (2015). 'Human capital and industrialization: evidence from the age of enlightenment', *Quarterly Journal of Economics,* vol. 130(4), pp. 1825-83.

[89] 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.

[90] Uzawa, H. (1965). 'Optimal technical change in an aggregate model of economic growth', *International Economic Review*, vol. 6, pp. 18-31.

[91] Vandenbussche, J., P. Aghion and C. Meghir (2006). 'Growth, distance to frontier and composition of human capital', *Journal of Economic Growth*, vol.11, pp. 97-127.

[92] Wantchekon, L., M. Klasnja and N. Novta (2015). 'Education and human capital externalities: evidence from Colonial Benin', *The Quarterly Journal of Economics*, vol. 130(2), pp. 703-57.

[93] 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
Log GDP per-capita	498	-3.924	0.540	-5.482	-1.914
Log Primary Sec. GDP per-capita	498	-5.176	0.580	-7.776	-3.330
Log Secondary Sec. GDP per-capita	498	-5.614	0.989	-8.561	-2.325
Log Tertiary Sec. GDP per-capita	498	-4.746	0.570	-6.292	-3.042
Share Higher Educ, 25+	500	0.183	0.074	0.046	0.498
Share Primary Educ, 25+	500	0.214	0.061	0.070	0.457
Share SC	500	0.162	0.079	0.000	0.501
Share ST	500	0.110	0.177	0.000	0.938
Pop 25+	500	874,426	613,867	41,358	4,670,683
Geography					
Coastal	500	0.098	0.298	0	1.000
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	619.580	3.967	4941.724
Historical					
Catholic Missionary	500	0.300	0.459	0.000	1.000
Share Primary Only, 1961	500	0.062	0.046	0.004	0.275
Protestant Missionary	500	0.582	0.494	0	1.000
Share Urban, 1931	500	0.108	0.075	0.000	0.495
Share Brahman, 1931	500	0.056	0.043	0	.270
Share Tribal, 1931	500	0.034	0.088	0	.69
Princely State	500	0.34	0.474	0	1.000
Railway Line, 1909	500	0.782	0.413	0	1.000
Historical, 1901 - Controls in Table 3					
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 Historical, 1901 - Controls cover 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 Higher Educ, 25+	6.879***	5.375***	5.029***	4.712***	4.856***
	(0.481)	(0.519)	(0.411)	(0.434)	(0.459)
Share Primary Educ, 25+					-0.464
					(0.749)
Share SC		-0.236	-0.558	-0.747	-0.714
		(0.667)	(0.484)	(0.470)	(0.468)
Share ST		-2.098***	-1.624***	-1.459***	-1.487***
		(0.336)	(0.223)	(0.238)	(0.250)
Pop 25+		0.168***	0.115**	0.069	0.062
		(0.058)	(0.054)	(0.052)	(0.053)
Coastal			-0.146**	-0.184**	-0.167**
			(0.074)	(0.071)	(0.078)
Longitude			0.027	0.036	0.036*
			(0.022)	(0.022)	(0.022)
Latitude			0.007	-0.009	-0.012
			(0.020)	(0.019)	(0.019)
Av. River Length			0.011	0.011	0.011
			(0.008)	(0.008)	(0.008)
Min Dist Big City			-0.127***	-0.122***	-0.125***
			(0.026)	(0.025)	(0.026)
Average Height			-0.083***	-0.081***	-0.081***
			(0.006)	(0.006)	(0.006)
Share Urban, 1931				0.463	0.446
				(0.444)	(0.446)
Share Brahman, 1931				-0.682	-0.603
				(0.936)	(0.965)
Share Tribal, 1931				-0.813*	-0.820*
				(0.473)	(0.474)
Princely State				-0.011	-0.021
				(0.059)	(0.063)
Railway Line, 1909				0.268***	0.271***
				(0.072)	(0.072)
Constant	1.291***	2.009***	2.470	2.278	2.415
	(0.479)	(0.371)	(1.565)	(1.574)	(1.629)
Observations	500	500	500	500	500
R^2	0.634	0.714	0.799	0.809	0.809
State FE	YES	YES	YES	YES	YES

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

Table 3
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 4
Instrumental Variables Results

	(1)	(2)	(3)	(4)	(5)
Panel A: First Stage - Catholic Missionary on Share Higher Education					
Catholic Missionary	0.053***	0.034***	0.037***	0.029***	0.026***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
F-stat Excluded IV	54.2	22.65	25.2	15.52	15.05
Cragg-Donald F-stat	69.19	27.91	31.12	19.84	20.5
Panel B: Second Stage - Log Light Density					
Share Higher Educ, 25+	8.382***	7.375***	7.948***	6.917***	7.225***
	(1.165)	(1.847)	(1.593)	(2.049)	(2.196)
Share Primary Educ, 25+					-2.131
					(1.625)
Observations	500	500	500	500	500
R^2	0.627	0.703	0.778	0.798	0.8
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 Brahmans, Tribals and urban population, an indicator for railways in 1909, and an indicator for Princely State. Schooling control includes the population share 25 and above with primary schooling but no higher than middle schooling that we call Share Primary Educ, 25+.

Table 5

Instrumental Variables Results for Higher and Primary Education

	(1)	(2)	(3)	(4)
Panel A: First Stage - Instruments on Share Higher Educ, 25+				
Catholic Missionary	0.029***	0.028***	0.026***	0.024***
	(0.007)	(0.007)	(0.007)	(0.007)
Share Primary Only, 1961		0.769***		0.797***
		(0.135)		(0.135)
Protestant Missionary			0.010*	0.013**
			(0.006)	(0.006)
Panel B: First Stage - Instruments on Share Primary Educ, 25+				
Catholic Missionary		0.004	0.002	0.001
		(0.004)	(0.004)	(0.004)
Share Primary Only, 1961		0.503***		0.520***
		(0.092)		(0.093)
Protestant Missionary			0.006	0.008**
			(0.004)	(0.004)
Cragg-Donald Wald F-stat	19.84	4.242	0.748	2.88
Kleibergen-Paap rk		8.53***	1.79	9**
Panel C: Second Stage - Log Light Density				
Share Higher, 25+	6.917***	7.770***	7.765**	7.770***
	(2.049)	(2.528)	(3.215)	(2.532)
Share Primary, 25+		-5.904	-5.868	-5.902
		(4.430)	(13.109)	(4.444)
Observations	500	500	500	500
State FE	YES	YES	YES	YES
Controls	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 Brahmans, Tribals and urban population, an indicator for railways in 1909, and an indicator for Princely State.

Table 6A

Bartik Demand Shock, Panel Data, 1961-2001

	(1)
	Share of Matriculates
Catholic Missionary*Predicted Demand Shock in Higher Education	0.9456***
	(0.2802)
Predicted Demand Shock in Higher Education	0.6145
	(0.5294)
Year FE	Yes
District FE	Yes
Observations	1,156

Robust standard errors clustered at the district-level in parentheses. ** p<0.01, ** p<0.05. We run the panel regression on aggregated districts that we can follow over time because of boundary changes in the 1961-2001 period.

Table 6B

Catholic Missionaries to Catholic Colleges

	(1)	(2)	(3)	(4)	(5)	(6)
	Catholic Colleges 1911	Catholic Colleges 1931	Catholic Colleges 1951	Catholic Colleges 1971	Catholic Colleges 1991	Catholic Colleges 2001
Catholic Missionary	0.026	0.047**	0.143***	0.385***	0.574***	0.427**
	(0.017)	(0.019)	(0.041)	(0.099)	(0.153)	(0.202)
Observations	327	500	486	500	486	500
R^2	0.122	0.146	0.255	0.409	0.382	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

The controls include geographic and historical controls. Geographic controls include an indicator for coastal districts, longitude, latitude, average river length, minimum distance to a big city and average height. In specification (1), we only focus on districts in former British India. Here the historical controls are 1901 share of Brahmans, Tribals and urban population, and an indicator for railways in 1909. In specifications (2)-(6), the historical controls include 1931 share of Brahmans, Tribals and urban population, and an indicator for railways in 1909, and an indicator for Princely State. Specifications (3) - (6) also include current controls for that year namely the population, share of Scheduled Castes and share of Scheduled Tribes. Data for Scheduled Caste and Scheduled Tribes is missing for Jammu and Kashmir in 1951 and 1991.

Table 6C

Dynamics of Increase in Catholic Colleges

	(1)	(2)	(3)	(4)	(5)
	Dep Var: Change in Catholic Colleges				
	1951-1961	1961-1971	1971-1981	1981-1991	1991-2001
Num Catholic Colleges, 1951	0.291***				
	(0.099)				
Num Catholic Colleges, 1961		0.365**			
		(0.155)			
Num Catholic Colleges, 1971			0.328***		
			(0.115)		
Num Catholic Colleges, 1981				0.261***	
				(0.069)	
Num Catholic Colleges, 1991					0.203***
					(0.030)
Share Urban, 1951	-0.235				
	(0.184)				
Share Urban, 1961		-0.359			
		(0.421)			
Share Urban, 1971			-0.215		
			(0.232)		
Share Urban, 1981				-0.168	
				(0.281)	
Share Urban, 1991					0.374
					(0.297)
Observations	486	500	500	477	486
R^2	0.446	0.520	0.549	0.681	0.690
State FE	Yes	Yes	Yes	Yes	Yes
Controls	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 total population, share of Scheduled Castes and share of Scheduled Tribes at the beginning of each decade (i.e. 1951 for specification (1), 1961 for specification (2), etc.). 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 Brahmans, Tribals and urban population, and an indicator for railways in 1909, and an indicator for Princely State. The data for Share SC/ST is missing for Kashmir districts in 1991 and 1951.

Table 7

IV Robustness Checks, Other Channels - Log Light Density

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Share Higher Educ, 25+	4.497**	7.798***	5.080***	6.951***	6.678***	5.361***	5.559***	6.994***	6.935***
	(2.252)	(2.287)	(1.852)	(2.075)	(1.878)	(1.788)	(1.787)	(1.994)	(2.027)
Observations	462	500	470	500	488	340	340	500	500
R^2	0.811	0.795	0.809	0.799	0.800	0.784	0.781	0.798	0.801
State FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Medical	YES	NO	NO	NO	NO	NO	NO	NO	NO
Religiosity	NO	YES	NO	NO	NO	NO	NO	NO	NO
Infrastructure	NO	NO	YES	NO	NO	NO	NO	NO	NO
Migration	NO	NO	NO	YES	NO	NO	NO	NO	NO
Violent Crime	NO	NO	NO	NO	YES	NO	NO	NO	NO
Confidence Courts	NO	NO	NO	NO	NO	YES	NO	NO	NO
Confidence Village Councils	NO	NO	NO	NO	NO	NO	YES	NO	NO
Europeans, 1931	NO	NO	NO	NO	NO	NO	NO	YES	NO
Pop-Density, 1931	NO	NO	NO	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 Brahmans, Tribals and urban population, an indicator for railways in 1909, and an indicator for Princely State.

Table 8

Dropping Outliers and Robustness Checks - Log Light Density

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share Higher Educ, 25+	5.871***	5.645***	6.203***	6.854***	6.857***	6.628***	7.379***	7.213***
	(2.093)	(1.142)	(1.891)	(2.028)	(2.553)	(2.256)	(2.121)	(2.224)
Observations	500	500	490	495	486	484	450	486
R^2	0.798	0.807	0.784	0.799	0.797	0.800	0.770	0.794
State FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Add Lagged Primary Educ Share	YES	NO	NO	NO	NO	NO	NO	NO
Intensive IV - #Catholic Miss/Area	NO	YES	NO	NO	NO	NO	NO	NO
Drop Lights Outliers	NO	NO	YES	NO	NO	NO	NO	NO
Drop Million+ City	NO	NO	NO	YES	NO	NO	NO	NO
Drop State Capital	NO	NO	NO	NO	YES	NO	NO	NO
Drop Largest City	NO	NO	NO	NO	NO	YES	NO	NO
Drop Higher Educ. Growth Outliers	NO	NO	NO	NO	NO	NO	YES	NO
Drop Kerala	No	NO	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 Brahmans, Tribals and urban population, an indicator for railways in 1909, and an indicator for Princely State.

Table 9

Alternate Measures of Income in Logs

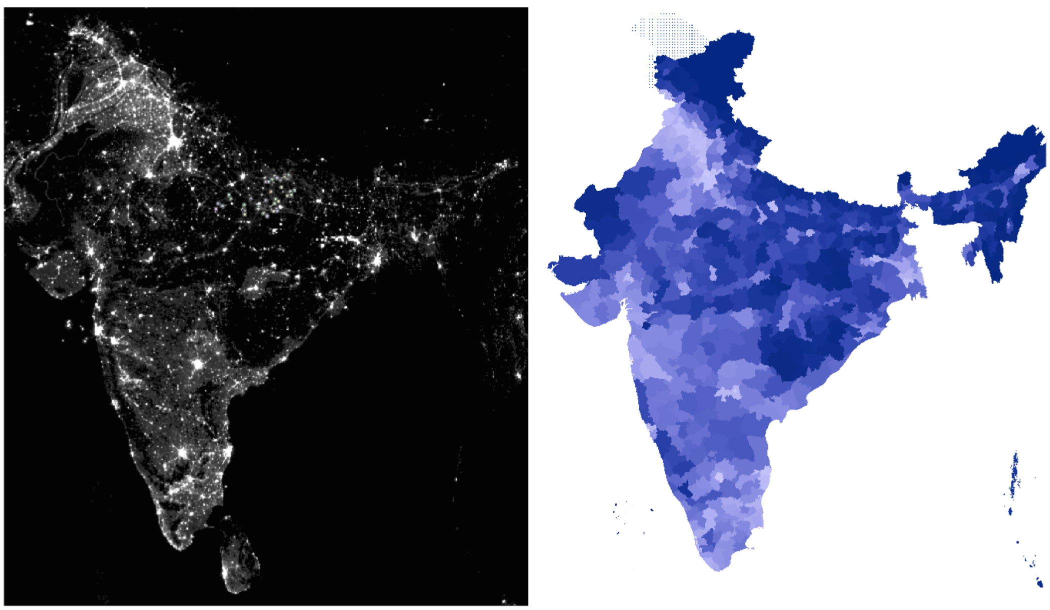
	(1)	(2)	(3)	(4)	(5)	(6)
	Lights per-capita	Consumption per-capita	Total GDP per-capita	Primary GDP per-capita	Secondary GDP per-capita	Tertiary GDP per-capita
Share Higher Educ, 25+	3.142*	1.874***	3.525***	-1.170	1.378	8.686***
	(1.811)	(0.670)	(1.145)	(1.878)	(2.479)	(1.440)
Observations	500	496	498	498	498	498
R^2	0.763	0.739	0.729	0.433	0.593	0.729
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 dependent variables in columns (4) to (6) refer to sectoral GDP per-capita. 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 Brahmans, Tribals and urban population, an indicator for railways in 1909, and an indicator for Princely State.

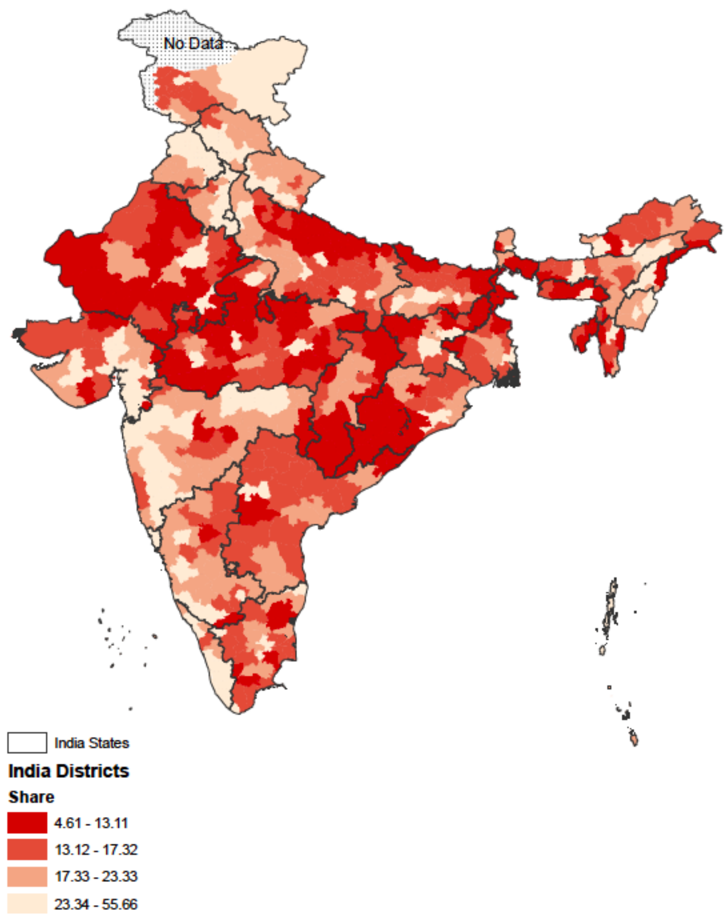
Figure 1:
Night Lights

(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 Adult Higher Education



Note: Light colors correspond to more higher education and dark colors correspond to less higher education.

Figure 3:
Correlation Light Density and Higher Education

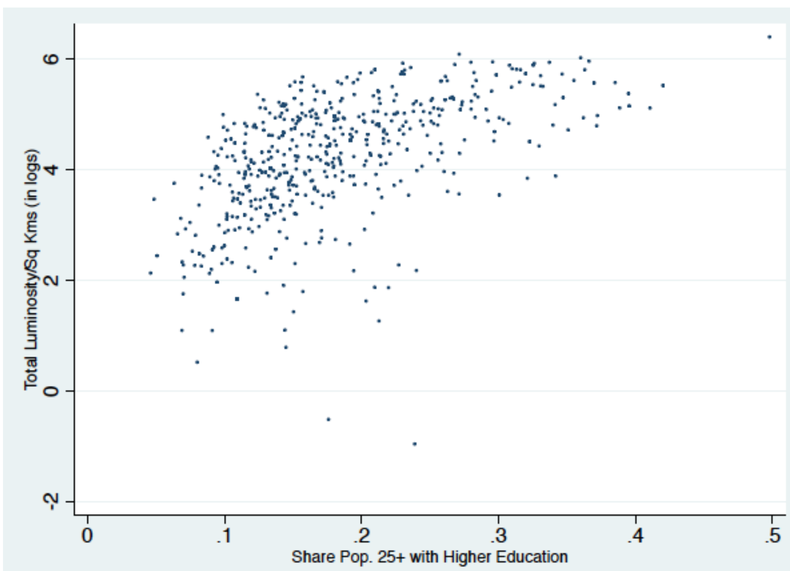


Figure 4:
Catholic Missionary Location, 1911

Appendix Table A1

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

Appendix Table A2

Catholic Missionary on Alternative Channels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Under 5 Mortality Rate	Share Chris-tians	Share Villages with Paved Road	Share Irrigated	Schools per Village	Share of Migrants	Violent Crimes per Capita	Confidence Courts	Confidence Panchay-ats
Catholic Missionary	-3.196 (2.173)	0.011** (0.006)	0.015 (0.013)	0.662 (1.577)	0.026 (0.174)	0.002 (0.009)	0.029 (0.038)	0.020 (0.016)	0.015 (0.019)
Observations	462	500	500	500	495	500	488	340	340
R-squared	0.816	0.529	0.838	0.646	0.702	0.482	0.369	0.402	0.442
State FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	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 Brahmans, Tribals and urban population, an indicator for railways in 1909 and an indicator for Princely State.

Appendix Table A3

Data Appendix, Variable Descriptions and Sources

Outcomes	
Log light density (area)	Log of light luminosity per square kilometre. Luminosity is measured in radians and ranges from 0 to 255 radians. Source: https://www.ngdc.noaa.gov/eog/dmsp/download_radcal.html
Log lights per capita	Log of light luminosity per person Source: Luminosity: as above, Population: Census, 2001
Log consumption per capita	Log of district-level average consumption expenditure per person Source: Consumption Expenditure Survey, National Sample Survey, 2004
Log GDP per capita	Log of district GDP per capita Source: Indicus Analytics, 2004
Log primary sector GDP per capita	Log of district primary sector GDP per capita. Primary sector includes agriculture, mining and quarrying, fishing, forestry Source: Indicus Analytics, 2004
Log secondary sector GDP per capita	Log of district secondary sector GDP per capita. Secondary sector includes construction, manufacturing, electricity, gas, and water supply Source: Indicus Analytics, 2004
Log of tertiary sector GDP per capita	Log of district tertiary sector GDP per capita. Tertiary sector includes trade, repair, hotels and restaurants, transport, storage, communication and services related to broadcasting, financial, real estate prof services, community, social and personal services Source: Indicus Analytics, 2004
Number of Catholic Colleges	Stock of catholic colleges for years 1911-2001. Source: Catholic Directory of India
2001 Education Variables	
Share of Higher Educ, 25 +	(Only Secondary school completed (25+) + Only Higher Secondary school completed (25+)+ University completed (25+) + Only Technical and Non Technical Diplomas (25+)) / Population (25+)

	Source: Census 2001
Share Primary Educ, 25+	(Only primary school completed (25+) + Only middle school completed (25+))/ Population (25+) Source: Census, 2001
Instruments	
Catholic Missionary	1 if a Catholic missionary was located in a district (2001 boundary) in 1911, 0 otherwise Source: Karl Streit (1913), Atlas Hierarchichus
Share Primary Only, 1961	Share of population that completed only primary education in 1961 Source: Census, 1961
Protestant Missionary	1 if a Protestant missionary was located in a district (2001 boundary) in 1908, 0 otherwise Source: Statistical Atlas of Christian Missions (1910)
Current Controls	
Pop 25+	Total district population aged 25 years and above Source: Census, 2001
Share SC	Share of Scheduled Castes Population Source: Census, 2001
Share ST	Share of Scheduled Tribes Population Source: Census, 2001
Geographical Controls	
Coastal	1 if district is on the coast, 0 otherwise Source: Map of India
Longitude	Longitude coordinates of the centroid of a district Source: Map of India, processed by Arc GIS
Latitude	Latitude coordinates of the centroid of a district Source: Map of India, processed by Arc GIS
Av. River Length	Average length of rivers that pass through the district (kms) Source: http://www.diva-gis.org/

Average Height	Average elevation of the district (kms) Source: http://www.diva-gis.org/
Min. Dist Big City	The distance of the district centroid to the nearest city with million plus population (Ahmedabad, Bangalore, Chennai, Delhi, Hyderabad, Jaipur, Kolkata, Mumbai, Pune and Surat) Source: Map of India, processed by Arc GIS
Historical Controls	
Share Urban, 1931	Share of urban population in 1931 Source: Census 1931
Share Brahman, 1931	Share of Brahman population 1931 Source: Census 1931
Share Tribal, 1931	Share of tribal population 1931 Source: Census 1931
Princely State	Indicator for districts that were in Princely India in 1931 Source: Census 1931
Railway Line, 1909	Indicator if a railway line passed through a district in 1909 Source: Indian Administrative Report on Railways, map was processed by Arc GIS
Share Urban 1901	Share of urban population in 1901 Source: Census 1901
Share Brahman, 1901	Share of Brahman population in 1901 Source: Census 1901
Share Lower Castes, 1901	Share of population recorded as lower caste in 1901 based on each province's enumeration of lower castes Source: Census 1901 and Chaudhary (2009)
Share Tribal, 1901	Share of tribal population in 1901 Source: Census 1901
Ethnic Frac 1901	A Herfindahl-based measure of caste and religious fractionalisation ranging from 0 to 1 Source: Census 1901 and Chaudhary (2009)
Total Population 1901	Total population of a district in 1901

	Source: Census 1901
Income-Tax per Capita 1901	Income tax revenues per capita in 1901 Source: District Gazetteers of India
Colleges 1901	Number of colleges in 1901 Source: District Gazetteers of India
Schools 1901	Number of schools in 1901 Source: District Gazetteers of India
Controls for 1951-1991	
Share Urban, 1951-1991	Share Urban Population 1951/61/71/81/91 Source: Census 1951, All India District Database 1961-1991
Share SC, 1951-1991	Share of Scheduled Castes Population 1951/61/71/81/91 Source: Census 1951, All India District Database1961-1991
Share ST, 1951-1991	Share of Scheduled Tribes Population 1951/61/71/81/91 Source: Census 1951, All India District Database 1961-1991
Population, 1951-1991	Total district population 1951/61/71/81/91 Source: Census 1951, All India District Database 1961-1991
Robustness Checks	
Medical	Under 5 Mortality Rates, 2001 Source: Census, 2001
Religiosity	Share of Christian population in 2001 Source: Census, 2001
Villages w/ paved road	Share of villages with all weather metalled roads Source: Census, 2001
Share of irrigated area	Share of land that is irrigated Source: Census, 2001
Schools per village	Total number of schools/number of villages in district Source: District Information System for Education, 2004
Violent Crime per 10,000 pop	Total Violent Crimes per 10,000 pop, 2004. Violent crimes include murder, attempt to commit murder, culpable homicide not amounting to murder, attempt to commit culpable homicide deaths, kidnapping and abduction

	Source: National Crime Record Bureau, India, 2004
Confidence in Courts	Share of households that report that they have a great deal of confidence in courts to meet out justice Source: Indian Human Development Survey, 2004-05
Confidence in Panchayats	Share of households that report that they have a great deal of confidence in village panchayats/nagarpalika (local councils) to implement public projects Source: Indian Human Development Survey, 2004-05
Migration	Share of Population (25+) who are migrants Source: Census, 2001
Europeans, 1931	Number of Europeans in a district in 1931 Source: Census, 1931
Pop-Density. 1931	Total population of a district divided by area of the district in 1931 Source: Census, 1931

Figure A1:
Location of Catholic Missionaries

(a) *Historical Map, North*

(b) *Historical Map, South*