

Department of Economics
ISSN number 1441-5429

The Long Shadow of Infrastructure Development: Long Run Effects of Railway Construction in Colonial India

Discussion Paper no. [2021-01](#)

Pushkar Maitra and William Yu

Abstract:

This paper examines the long-term impacts of infrastructural investment. It considers the case of British investment in railway infrastructure in colonial India. Railways had an immediate impact on trade and development in the predominantly agricultural India. In this paper, we show that the positive effects of railways have persisted over more than a century. Districts of the Indian sub-continent that were connected to railways earlier continue to have higher levels of economic prosperity and lower rural poverty rates a century later. Men and women residing districts connected earlier are less likely to be uneducated or malnourished. Districts further away from connected districts are worse off in terms of levels of economic development in 2013. The corresponding IV estimates are larger in magnitude than the OLS estimates indicating that the OLS estimates provide a lower bound to the effect of exposure to railways on long run prosperity. The persistent effects appear to be driven by agglomeration due to early exposure to trade and globalization as a result of connectedness.

Keywords: Infrastructure, Railways, Long Run Prosperity, Colonial India

JEL Classification: O11, N75, O18

Pushkar Maitra: Monash University, Department of Economics (email: pushkar.maitra@monash.edu);
William Yu: Monash University (email: williamyu117@hotmail.com).

© The authors listed. All rights reserved. No part of this paper may be reproduced in any form, or stored in a retrieval system, without the prior written permission of the author.

monash.edu/business/economics

ABN 12 377 614 012 CRICOS Provider Number: 00008C



The Long Shadow of Infrastructure Development: Long Run Effects of Railway Construction in Colonial India*

Pushkar Maitra[†] and William Yu[‡]

JANUARY 2021

Abstract

This paper examines the long-term impacts of infrastructural investment. It considers the case of British investment in railway infrastructure in colonial India. Railways had an immediate impact on trade and development in the predominantly agricultural India. In this paper, we show that the positive effects of railways have persisted over more than a century. Districts of the Indian sub-continent that were connected to railways earlier continue to have higher levels of economic prosperity and lower rural poverty rates a century later. Men and women residing districts connected earlier are less likely to be uneducated or malnourished. Districts further away from connected districts are worse off in terms of levels of economic development in 2013. The corresponding IV estimates are larger in magnitude than the OLS estimates indicating that the OLS estimates provide a lower bound to the effect of exposure to railways on long run prosperity. The persistent effects appear to be driven by agglomeration due to early exposure to trade and globalization as a result of connectedness.

JEL Codes: O11, N75, O18

Key Words: Infrastructure, Railways, Long Run Prosperity, Colonial India

*We thank Shameel Ahmad, Reshad Ahsan, Sascha Becker, Prasad Bhattacharya, Dan Bogart, Lionel Frost, Wayne Geerling, Arghya Ghosh, Gary Magee, Debdulal Mallick, Ahmed Skali, Chinmay Tumbe and seminar participants at Deakin University, University of South Australia, St Xavier's University Kolkata, Presidency University Kolkata and IIT(Kanpur) for discussions and comments. We are grateful to Chinmay Tumbe for providing us access to the district level population data and Debdulal Mallick for giving us access to the data on ruggedness at the district level. Adit Maitra provided excellent research assistance.

[†]Department of Economics, Monash University, Clayton Campus, VIC 3800, Australia.
Pushkar.Maitramonash.edu

[‡]Monash University. williamyu117hotmail.com

1 Introduction

Government decisions on location of public investments have significant economic and political consequences. Regions that benefit from government investment in infrastructure (roads, railways, schools and hospitals), can gain from such investment for decades if not centuries.¹ Relatedly, there is now a large and growing literature documenting the long-term impact of historical institutions: history, through geography, origins of legal structures, institutions, human capital and culture can have persistent impacts. See, for example, [Nunn \(2009\)](#), [Dell \(2010\)](#), [Nunn and Wantchekon \(2011\)](#), [Nunn \(2013\)](#), [Spolaore and Wacziarg \(2013\)](#), [Michalopoulos and Papaioannou \(2013\)](#), [Caicedo \(2019\)](#), among others .

This paper examines the long run impacts of infrastructure development. We consider the case of what is accepted to be one of the largest infrastructural investments in any country in history: the building of the railway infrastructure in colonial India by the British government during the period 1853–1930. The primary objectives of the colonial British government’s investment in railway infrastructure were military, commercial and social, in that order of priority (see, [Sanyal, 1930](#), [Macpherson, 1955](#), [Sharma, 2010](#), [Hurd and Kerr, 2012](#)). Railways changed India in many ways. [Sanyal \(1930\)](#) argued that as a result of the introduction of railways, the *stationary state of India had ceased* (Page 61). Governance became easier as the efficiency of the army and of the civil administration increased immensely. The short term economic impacts were also considerable. [Donaldson \(2018\)](#) argues that the introduction of railways decreased trade costs and inter-regional price differences, increased inter-regional and international trade and most importantly increased contemporaneous real income levels. [Burgess and Donaldson \(2012\)](#) argue that the introduction of railways made real incomes less responsive to local weather shocks. This might have played an important role in reducing the effect of famines in India in the 19th century. [Chaudhary and Fenske \(2020\)](#) show that introduction of railways had positive effects on male and English literacy in colonial India.

In this paper, we show that the positive effects of railways have persisted over more than a century. Using the data collected by [Donaldson \(2018\)](#), we are able to identify the year of introduction of railways in each district of present day India, Bangladesh and Pakistan.² Specifically, [Donaldson](#)

¹For example, [Banerjee et al. \(2020\)](#) find that in China, regions closer to historical transportation networks enjoyed higher levels of GDP per capita, higher income inequality, a higher number of firms and higher average firm profits in the two decades following the opening up of the country to trade in 1979. [Dalgaard et al. \(2018\)](#), consider regions under Roman rule as of CE 117 and show that greater Roman road density in the 2nd century CE is correlated with significant modern road density and greater economic activity in 2010.

²The Indian sub-continent now consists of three separate and independent countries: India, Pakistan and Bangladesh. The British ruled this entire region as a single country, India. In 1947 India (at the time of Indian independence), the country was sub-divided into India and Pakistan. Between 1947 and 1971, Pakistan consisted of two separate geographical regions (West Pakistan and East Pakistan), which were separated by India. In 1971, following a bloody civil war, East Pakistan separated from Pakistan and became the independent country

(2018) geo-coded each segment (approximately 20 km long) of the railway network and also noted the year in which each segment was opened. We start with the coordinates of these locations and then match them to the centroid of the present-day districts of India, Bangladesh and Pakistan. This informs us when railways were introduced in each district and thus gives us a measure of length of exposure to railways at the district level. We show that this length of exposure is associated with an increase in prosperity at the district level in the present, i.e., 150 years after railways were first introduced in the Indian sub-continent.

Our focus is on districts, an administrative level below states and the lowest level at which data is widely available. However, one major problem with a district level analysis is that current income levels at the district level are difficult to obtain. They are often measured with error and data collection is not standardized across India, Bangladesh and Pakistan. To avoid these problems we have used night time lights as a proxy for local income, which provides us with a metric that is consistent. Asher et al. (2021) argue that night lights are a good cross-sectional proxy for local economic development.

We find that the length of exposure to railways at the district level is positively and statistically significantly correlated with district level prosperity more than a century later. In particular, relative to districts not being connected by 1930, districts connected to the railway network prior to 1880 exhibit significantly higher levels of economic prosperity in 2013. The results are robust to the inclusion of a set of pre- and post-independence era controls. Early exposure to railways is associated with higher levels of aggregate bank deposits; a reduction in rural poverty in the district (though urban poverty generally remains unaffected), and improved socio-economic outcomes: the proportion of adult males and females with no schooling is lower in districts that have been exposed longer as is the proportion of underweight (i.e., malnourished) men and women. Using instruments we estimate the causal effect of early railway connectedness on long term prosperity of the district. Like the OLS estimates, the IV estimates of early railway connectedness are positive and statistically significant. The IV estimates are larger, indicating that the OLS estimates provide a lower bound on the effect of exposure to railways on long term prosperity.

We also find strong evidence of geographical spillovers associated with the introduction of railways: an increase in the distance from the nearest connected district is associated with a significant decline in economic activity a century later. As more and more districts are connected, the stronger is the detrimental effect of an increase in the distance of a non-connected district to the nearest connected district. All of these results are indicative of significant beneficial effects of railway connectedness.

of Bangladesh.

In terms of mechanisms we argue that the results are driven by positive agglomeration effects due to increased exposure of connected districts to both international and interregional trade. Consistent with the arguments of [Jedwab and Moradi \(2016\)](#), we find initial infrastructural investments have large effects, leading to structural changes. Subsequent investments though have less significant effects because they are unable to affect the structure of the economy in the manner and to the extent initial infrastructural investments could.

There is a growing literature that examines the impact of institutions developed by the colonial British government on different aspects of the Indian economy. [Banerjee and Iyer \(2005\)](#) examine the land revenue institutions set up by the British in India, and show that differences in historical property rights institutions lead to persistent differences in economic outcomes. Relative to regions where land rights were granted directly to cultivators, regions of India where landlords were granted proprietary rights to land, had significantly lower agricultural investments and productivity in the post-independence period. These areas also experienced lower investments in health and education. [Iyer \(2010\)](#) compares economic outcomes across areas in India which were under direct British colonial rule to areas that were ruled by a local king as a protectorate of the British.³ She finds that quality of governance has long term impacts: areas that experienced direct rule have significantly lower levels of access to schools, health centers and roads in the post-colonial period. [Lehne \(2019\)](#) examines the long-term effects of opium cultivation in India.⁴ Opium growing districts in colonial India enjoyed higher spending on irrigation (water supply was critical for opium cultivation) and rural police (to ensure safety of traders and restrict smuggling), but lower spending on health and educational infrastructure. [Lehne \(2019\)](#) shows that while the complementary investments did not have any persistent effects (there are no significant effects on the percentage of land that is irrigated or density of police station more than a century later), the effects on human capital are persistent over time. [Dar \(2019\)](#) shows that districts in India that benefited from colonial investments in agriculture (i.e., had a canal build pre-1931) experienced successful adoption of green revolution practices post-independence.⁵

³ British rule in India after 1857 was characterized by two kinds of territories: *Presidencies* and *Residencies*. Presidencies, were the direct rule areas where the Crown had complete control over the administration. Residencies (also known as the indirect rule areas or princely states) were territories where the Crown, could only oversee through a resident of the crown (see, for example, [Banerjee and Iyer, 2005](#), [Iyer, 2010](#), [Wiener, 2013](#)). These indirect rule areas were allowed domestic autonomy with preconditions of good governance and were provided with security against foreign invasions. An agent or resident was located in the territory to represent the Crown and to monitor the day-to-day affairs ([Rudolph, 1963](#)). Figure B1 presents the classification of the districts under direct and indirect rule.

⁴One of the largest and most vicious narcotics operation in history was the state sanctioned production of opium in India for export to China. At the peak of this narcotics trade in the middle of the 19th century, opium trade accounted for more than a third of India's exports and about a fifth of the total revenue. To facilitate this trade, first the East India company and later the colonial British government maintained a state-run enterprise (established in 1797 and in place until 1912) and contracted more than one-million farmers to grow poppies. See [Sahu \(1977\)](#).

⁵[Calvi and Mantovanelli \(2018\)](#) and [Calvi et al. \(2020\)](#) examine the long run effect of Protestant medical Missions that spread throughout colonial India. [Calvi and Mantovanelli \(2018\)](#) provide evidence of a strong and robust positive correlation between proximity to a Protestant medical mission and the health outcomes of individuals a century later. [Calvi et al. \(2020\)](#) show that these Protestant Missions were in effect pioneers of female education in

2 Railways and Economic Development in India and around the World

The Battle of Plassey in 1757 is viewed as a landmark event in Indian history. It was here that the East India Company (commanded by Robert Clive) defeated the Nawab of Bengal (Siraj-ud-Daulah) and his French allies. This battle helped the Company seize control of Bengal (which comprises of present-day West Bengal, Bihar, Orissa and parts of Assam in India and the country of Bangladesh). This is viewed as the start of British colonial rule in the Indian sub-continent. In Bengal, after the Battle of Plassey, the British wielded enormous influence and consequently acquired significant concessions for trade. The Company used this revenue to increase their military might and over the next 100 years, the East India Company systematically defeated its rivals in India, including indigenous states such as the Marathas, Mysore, and Sikhs, along with foreign powers such as the Dutch and French. This helped them extend their monopoly over trade while simultaneously expanding the British Empire. After the Sepoy Mutiny of 1857, the governance of India was taken over by the British Parliament.⁶

The idea of a railway network in India was first raised in 1844, and a few experimental railways lines were constructed to test the effectiveness of railway transport. However, it was Lord Dalhousie, who formally proposed the idea of an extended railway network in India. While initially the Board of Directors of the East India Company were not in favour of the proposal, Lord Dalhousie convinced the home authorities of the need for the railways and laid down the trunk lines of their development.

The country's first passenger train ran between Bombay (Mumbai) and Thane in April 1853. Thus, began a railway revolution in the Indian sub-continent. The locomotives were mechanical beasts of steel (or *fire-carriages*) that were not only physically imposing, but also a symbol of British power and rule. As the railway network was extended over the next 70 years, so grew British power, prestige and influence. Starting from 1853, when the first railway line measuring

India: they employed female teachers who could directly interact with women (which men could not) to impart basic literacy skills. They find a strong positive association between historical exposure to Protestant missions and human capital a century later, an effect that is particularly strong for females. [Castelló-Climent et al. \(2018\)](#) in examining the effect of higher education on economic prosperity use location of Catholic missions as an instrument to address the endogeneity of higher education. Unlike the Protestant missions, Catholic missions have had otherwise limited effects beyond education.

⁶The Sepoy Mutiny or the Indian Rebellion of 1857 was a major, but ultimately unsuccessful, uprising in India against the rule of the British East India Company, which functioned as a sovereign power on behalf of the British Crown. It has been described as the first war of independence of India. While the rebellion was ultimately contained, it had far reaching impacts. Indeed, it proved to be an important watershed for India and also the British empire. It led to the dissolution of the East India Company, and forced the British to reorganize the army, the financial system, and the administration in India, through passage of the Government of India Act 1858. India was thereafter administered directly by the British government.

32 km was opened in the Bombay Presidency, India’s railway network expanded to more than 60000 km by 1930 (about 90% of India’s railway network was constructed in this period).⁷

Through out the world, the 19th century witnessed large scale investments in railways. This resulted in significant effects on developing countries around the world. [Herranz-Loncán and Fourie \(2018\)](#) found that railways in the British Cape Colony in South Africa brought great prosperity to areas that were connected (though at the cost of substantial increase in inequality). Specifically, they find that the railways that connected the inner Kimberley diamond fields with the port cities, were responsible for a massive increase in economic growth. Further, [Jedwab and Moradi \(2016\)](#) and [Jedwab et al. \(2017\)](#) using data from different countries in Africa show that railways resulted in the creation of a new spatial equilibrium in Africa. [Jedwab and Moradi \(2016\)](#) argue that even though railroads in Sub-Saharan African countries went into decline after these countries received independence, the initial beneficial effects have persisted over time. Likewise, late 19th century railroads are credited with facilitating factor mobility as well as access to human and financial capital in Japan ([Tang, 2014](#)). By augmenting existing transportation methods, railways in Japan widened the market for inputs and outputs and resulted in more efficient use of inputs in the production process. In Mexico, [Coatsworth \(1979\)](#) finds that the construction of the railroads led to a large increase in short term savings and economic growth, but the focus on the export sector meant that there were no significant effects on human capital investment and no long-term effects. Railroads, however, had significant impact in 19th and 20th century Brazil, and they contributed heavily to the transition from stagnation to growth in the country after 1913 ([Summerhill, 2005](#)). Finally, [Paik and Vechbanyongratana \(2020\)](#) examine the effects of railway construction in the late 19th and early 20th century in Thailand on a range of economic outcomes in the middle of the 20th century. They find that railways increased urbanization and increased agricultural production in the peripheral regions; but the effects were rather localized to regions that benefited from direct access to railways.⁸

In India, the major beneficiary was trade. Prior to the introduction of the railways, most of the transport of commodities was either by river or by bullocks. [Deloche \(1994\)](#) argues that transport using bullocks was slow, traveling approximately 30 km per day along the network of dirt roads of India. While water transport was superior to road transport, it was only feasible on the Brahmaputra, Ganges and Indus river systems. In optimal conditions, downstream river traffic

⁷Figures [B2](#) and [B3](#) presents the present day districts connected by decade and track length in district by decade.

⁸The literature on the contemporaneous effect of railroads on economic growth and development in the US and European countries is large and growing. The seminar work by [Fogel \(1962\)](#) and more recent research by [Atack and Margo \(2011\)](#), [Donaldson and Hornbeck \(2016\)](#) and [Katz \(2018\)](#) examine the impacts of railroads on the US economy in the 19th century. [Hornung \(2015\)](#) examine the effect of railroad access (access to a station) on urban population growth in Prussia. [Berger and Enflo \(2016\)](#) study the impacts of railroads in Sweden. They find that towns that gained access experienced substantial relative increases in population, mainly as a result of relocation of economic activity. The mechanism that we suggest in this paper (see [Section 6](#)) is similar to that suggested by [Berger and Enflo \(2016\)](#).

could cover up to 65 km per day, but upstream traffic was considerably slower. Additionally, both road and water transport were restricted by prevailing weather conditions (for example roads were unusable in monsoon) and there was the perennial fear of piracy. Railways on the other hand could transport the same commodity 600 km a day: naturally the per mile freight rates were substantially lower. [Donaldson \(2018\)](#) finds that the construction of railways decreased trade costs and inter-regional price differentials; they increased both inter-regional and international trade and increased real income levels. [Sanyal \(1930\)](#) and [Bogart and Chaudhary \(2016\)](#) present excellent summaries of the impact of railways on the Indian economy during the colonial era.

3 Data and Descriptive Statistics

We use a number of different data sets for our analysis. The first set of data pertains to the year of introduction of railways in each district of India. Using the GIS database that contains the locations of contemporary railroad, river and coast lines from the Digital Chart of the World, [Donaldson \(2018\)](#) coded each segment (approximately 20 km long) of the railway network according to the year in which it was opened. We matched each location in this data to the current (as of 2013) districts in India, Bangladesh and Pakistan. Each segment is assigned to the district whose centroid is the nearest to the segment. This way, for each district of the Indian subcontinent, we have the year of introduction of railways.⁹ [Figure 1](#) shows the decade of introduction of railways in districts of present-day India, Bangladesh and Pakistan, as of 1956, while [Figure 2](#) presents the cumulative track length in each district of present day Indian sub-continent as of 1930.

Panel A of [Figure 3](#) presents the percentage of districts where railways were introduced in each five-year period starting in 1853. The large majority of districts were connected between 1860 and 1900. The large increase following the Sepoy Mutiny (more than 12% of the districts were connected in the 5-year period 1858–1862) suggests that there was indeed a strong military motive. Panel B of [Figure 3](#) presents the cumulate length of tracks laid over the period 1853–1956. Starting from low numbers in 1853, more than 60,000 kms of railway tracks had been laid by 1956, the majority of this by 1930.

The second set of data relates to our key outcome variable. It is difficult to obtain consistent data on economic outcomes at the sub-national (district) level. In our context, across India, Bangladesh and Pakistan, data collection is not standardized. To make proper comparisons across districts, we need a consistent metric that we could use across all districts in the three countries. We use night

⁹More railways lines were build in these districts over the following decades, but for the majority of this paper we restrict our analysis to the year of introduction of the railways.

time lights (or night lights).¹⁰ This data comes from the U.S air forces Defense Meteorological Program (DMSP) Operational Line-Scan System (OLS), which is capable of detecting visible and near-infrared (VNIR) emission sources at night. The time series night lights data is available from the National Oceanic and Atmospheric Administrations (NOAA) National Geophysical Data Center (NGDC) as a set of annual composites. The sensors collect light data up to a value of 63 (with 0 being the lower bound representing no lights). There are several steps to transform night-light data into a usable format for economic analysis. First, zonal borders need to be calculated, so that data for a region can be summarized.¹¹ Next each pixel is given a Digital Number from 0–63. Using these Digital Numbers, the total light emitted from a region can be calculated by adding up all the individual values of the pixels. This is *sumlight*. Additionally, the average light of a pixel can be similarly found by summing and averaging over total pixels. This is *meanlight*. While the primary objective of the OLS sensors was to aid in the interpretation of meteorological conditions, collecting data on clouds, the data has been re-purposed by social scientists to collect inexpensive night light data. Indeed [Florida et al. \(2008\)](#), [Henderson et al. \(2011, 2012\)](#) all argue that night lights can be used to provide a single consistent metric of urbanisation or development. [Bhandari and Roychowdhury \(2011\)](#) argue that DMSP-OLS images can predict district level GDP in India quite accurately and indeed can explain both variation in levels and growth rates of district level GDP.¹² Figure 4 presents the spatial distribution of night lights (measured by $\log(1 + \textit{sumlight})$) across the districts of the Indian sub-continent (as of 2013).

We include a set of pre- and post-independence controls. The colonial controls include the real income and total area in a district that was cultivated by 17 major crops. This data is obtained from [Donaldson \(2018\)](#). We also include a dummy for Colonial state, which takes the value of 1 if the district was under the direct rule of the British in the colonial period – as opposed to being under indirect rule, i.e., ruled by an Indian king in the colonial period.¹³

We include the proportion of SC and ST population at the district and the percentage of the population that is literate as of 2011 as additional controls. The data on these two variables are obtained from the 2011 census. Finally we also include a set of geographical variables: the length of river (in km) in the district and percentage of the district with mean elevation of more than 500 metres. These data are obtained from [Duflo and Pande \(2007\)](#).

¹⁰The data is available from <https://datainspace.org/index.php/global-nighttime-lights-at-adm2-level-1992-2013/>

¹¹For zonal statistics calculations, see [Hodler and Raschky \(2014\)](#). The original night time lights data is available from [NOAA \(2014\)](#) and the administrative boundaries are available from [GADM \(2012\)](#).

¹²A number of recent studies that have used night time lights as a proxy for economic outcomes at the sub-national level. See, for example, [Sutton et al. \(2007\)](#), [Noor et al. \(2008\)](#), [Florida et al. \(2008\)](#), [Bhandari and Roychowdhury \(2011\)](#), [Henderson et al. \(2011, 2012\)](#), [Hodler and Raschky \(2014\)](#), [Bruederle and Hodler \(2018\)](#), [Singhal et al. \(2020\)](#).

¹³We acknowledge the potential endogeneity of the colonial rule variable. It is unlikely that the British randomly annexed areas for direct rule. However, the colonial rule variable is not the key variable of interest in our analysis. Also the results are qualitatively and quantitatively similar when the colonial state variable is not included in the set of controls.

We examine the robustness of our results using several alternative measures of district level prosperity. First, instead of using $\log(1 + \textit{sumlight})$ as the primary outcome variable, we use $\log(1 + \textit{meanlight})$ as the outcome variable.¹⁴ Second, we use aggregate bank deposits in a district, which is the sum of all bank deposits at the district level in different banks: state bank and associates, other nationalized banks, foreign banks, regional rural banks, old private sector banks, new private sector banks. Quarterly data on bank deposits has been recently been released by the Reserve Bank of India (RBI). We use data for 2013 and the quarterly data on bank deposits is aggregated over the calendar year to obtain the relevant measure.¹⁵ Finally, we also use rural and urban poverty rates at the district level. This data is obtained from [Chaudhuri and Gupta \(2009\)](#) and are computed using the consumption module of the 60th round of the NSS (survey conducted in 2004).¹⁶

To examine the long-term effects of exposure to railways on socio-economic outcomes, we use Round 2 of the Indian Human Development Survey (IHDS2) data set. The IHDS2, conducted in 2011–2012 is the second wave of a panel data set. The first round of the survey (IHDS1) was conducted in 2004–2005. The surveys collected information on health, education, employment, economic status, marriage, fertility, gender relations, and social capital. While both rounds of the survey collected anthropometric data for women, the corresponding data for men was collected systematically only in IHDS2. The key outcome variables that we consider are proportion of men and women in the district that are underweight (body mass index, defined as the ratio of weight (in kilograms) to height (in meters) squared or $\text{BMI} < 18.5$) and the proportion of men and women in the district that have no schooling. Across all districts, 29% of the population have no schooling; the percentage of men with no schooling is 18.7, while the corresponding percentage for women is a significantly higher 39.3% (difference p-value = 0.00). Similarly, across all districts 15.5% of the population are underweight; this proportion is a high 18% for women and a significantly lower 13% for men (difference p-value = 0.00).

4 Empirical Specification

The key outcome variable to examine the long-term effects of railway construction is $\log(1 + \textit{sumlight}_{dst})$ in district d in state s in year t . A district is denoted as a “treatment” district, if it is connected to railways by 1930 and a “control” district if the district was not connected to

¹⁴Figure [A2](#) presents the spatial distribution for $\log(1 + \textit{meanlight})$ as of 2013.

¹⁵Figure [A3](#) presents the variation in aggregate bank deposits across the different districts of the country as of 2013. See [Cook and Shah \(2020\)](#) on the use of aggregate bank deposits as an indicator of district level prosperity in India.

¹⁶This data is only available for the larger states of India. See Figure [A4](#) for the variation across districts in rural (left panel) and urban (right panel) poverty rates.

railways by 1930. This could include districts that were connected after 1930 or have remained not connected since (as of 1993). Treatment is, however, time varying, depending on the length of exposure to railways. The length of exposure is captured using a set of binary indicator variables:

- $R_1 = 1$ if railway was introduced in the district in the period 1850–1860 (Railway First 1850–1860);
- $R_2 = 1$ if railway was introduced in the district in the period 1861–1870 (Railway First 1861–1870);
- $R_3 = 1$ if railway was introduced in the district in the period 1871–1880 (Railway First 1871–1880);
- $R_4 = 1$ if railway was introduced in the district in the period 1881–1890 (Railway First 1881–1890);
- $R_5 = 1$ if railway was introduced in the district in the period 1891–1900 (Railway First 1891–1900);
- $R_6 = 1$ if railway was introduced in the district in the period 1901–1910 (Railway First 1901–1910);
- $R_7 = 1$ if railway was introduced in the district in the period 1911–1930 (Railway First 1911–1930);

The estimating equation is given by

$$\log(1 + \text{sumlight}_{dst}) = \beta_0 + \sum_{i=1}^7 \beta_i R_i + \gamma \mathbf{X}_{dst} + \lambda_s + \eta_t + \varepsilon_{dst} \quad (1)$$

Where \mathbf{X}_{dst} include the latitude and longitude of the centroid of each district and a set of district specific pre- and post-independence controls; λ_s denotes a set of state dummies to capture any state level observed and unobserved characteristics; and η_t denotes a set of year effects. We do not include the year effects (η_t) when we restrict our analysis to a specific year. Since year of introduction of railways is at the district level, we cannot include any district fixed effects. One advantage of this specification (where the timing of introduction of railways in the district are included as dummies) is that we do not impose any functional form restrictions on how timing of introduction of railways in the district affect long-term economic performance. We compute standard errors in three different ways: state clustered, spatially adjusted for 50 km and spatially adjusted for 100 km.¹⁷

The analysis thus far ignores spatial spillovers. In particular, it assumes that only the districts directly connected to a railway line benefit. However, districts near a railway connected district,

¹⁷The spatially adjusted standard errors (Conley, 1999, Kelly, 2019) are computed using the *acreg* command in STATA (Colella et al., 2019). The regression tables presented in the paper include the state clustered standard errors. The results using the spatially adjusted standard errors for the key results are presented in Tables A1 and A2 and are available for the rest.

could also benefit from geographical proximity to access to railways. To account for geographical spillovers, we need to examine the effect of introduction of railways by a particular year T in a particular district or in a district that is close to the connected district. To do this, we use the following econometric specification

$$\log(1 + \text{sumlight}_{ds}) = \alpha_0 + \alpha_1 \Phi(M_{dsT}) + \gamma \mathbf{X}_{ds} + \lambda_s + \varepsilon_{ds} \quad (2)$$

Here M_{dsT} measures the distance from the centroid of a particular district to that of the nearest connected district as of time T . This distance is 0 for a connected district. A-priori we cannot be sure of the functional form of the relationship between M_{dsT} and local economic development. The Φ function captures the nature of this relationship. We consider connectedness as of $T = 1860, 1870, 1880, 1890, 1900, 1910$ and 1930 . If spatial spillovers are important, then we expect a smaller effect for a district further away from a connected district, i.e., the estimated coefficient α_1 will be negative.

We also estimate a restricted version of equation (2): we assume that a district benefits *only* because it is connected and being close to a connected district has no bearing, i.e., geographical spillover do not have a role to play. In this case the estimating equation is given by

$$\log(1 + \text{sumlight}_{ds}) = \delta_0 + \delta_1 d_{dsT} + \gamma \mathbf{X}_{ds} + \lambda_s + \varepsilon_{ds} \quad (3)$$

Here d_{dsT} is a railway connectedness dummy that takes the value of 1 if district d in state s as of year T is connected and 0 otherwise ($T = 1860, 1870, 1880, 1890, 1900, 1910$ and 1930). The estimated value of δ_1 in equation (3) then gives the overall effect of a district being connected in by year T relative to not being connected by 1930. We expect the estimated δ_1 to be positive.

Notice that the estimated coefficients δ 's from equation (3) are to be interpreted differently to the the estimated coefficients β 's from equation (1). The latter denote the effect of introduction of railways to a district in decade i , and is interpreted as the effect relative to the district not being connected by 1930. Therefore, in equation (1), $R_i = 1$ if the district is connected in decade i and 0 otherwise. In equation (1), the R_i dummies are mutually exclusive. In equation (3), $d_{dsT} = 1$ if district d in state s is connected by year T and 0 otherwise. A district is defined to be connected (i.e., $d_{dsT} = 1$) if it is connected any time before T . As an example, a district is connected as of 1880 if it is connected any time before 1880 (i.e., if $R_1 = 1$ or $R_2 = 1$ or $R_3 = 1$ and 0 otherwise). The estimated δ_1 therefore capture the cumulative effect of the district being connected by T .

5 Results

5.1 Baseline Results

Table 1 presents the baseline results. In all regressions presented in this Table, the explanatory variables are the time of introduction of railways, a set of country dummies, the latitude and longitude of the centroid of the district, a set of state fixed effects (columns 2–6) and a set of year dummies (columns 1 and 2). The state fixed effects control for all other state level observed and unobserved differences. The dependent variable (our measure of district level prosperity) is given by $\log(1 + \textit{sumlight})$. Column 1 presents the OLS regression results on the pooled data over the period 1992–2013 (the latest year for which data is available). The sample includes all districts in present day India, Bangladesh and Pakistan. The coefficient estimates presented in column 1 show that relative to the railways being introduced in the district after 1930, early introduction railways have a positive and statistically significant effect on night lights. While the magnitude of the within state estimates of the impact of the timing of railway construction are lower (see column 2, where we include state fixed effects), qualitatively they are similar to those presented in column 1.¹⁸

Columns 3 and 4 present the corresponding regression results at specific points in time: 1992 (column 3) and 2013 (column 4). In column 5 the dependent variable is the average night light over the period 1992–2013. All of these results essentially tell the same story: relative to the railways being introduced in the district after 1930, early introduction railways, specifically introduction before 1880 have a large, positive and statistically significant effect on night lights in the district.

The dependent variable in column 6 is the coefficient of variation of night lights in the district over the period 1992–2013. This variable can be viewed as a measure of volatility in night light over time within a district. The results in column 6 imply that relative to railways being introduced in the district after 1930, the extent of volatility is significantly lower in districts where railways were introduced prior to 1880. These results suggest that railways can help insure against economic fluctuations.

The baseline results presented in Table 1 therefore imply that early introduction of railways have a much stronger effect on night lights. Early introduction of railways is also associated with a decline in volatility, measured by the coefficient of variation of night lights in the district. However, the results presented in Table 1 do not allow us to account for the fact that there might

¹⁸In Table 1 we present results with state clustered standard errors. The results using spatially corrected standard errors are presented in A1.

be other contemporary or colonial characteristics that drive the level of economic development in the district, i.e., the results in Table 1 could potentially suffer from omitted variable bias.

5.2 Including Additional Controls. India Only

To account for possible omitted variable bias, we include pre- and post-independence controls. However, here we are constrained by data availability (and comparability across Bangladesh, India and Pakistan). We therefore restrict our analysis to district in present day India. The dependent variable continues to be $\log(1 + sunlight_{ds})$ in district d in state s as of 2013.

The regression results are reported in Table 2. Column 1 replicates the results from column 4 in Table 1, though only for the restricted sample. The results show that relative to railways not being introduced in the district pre-1930, early introduction of railways (introduced prior to 1880) has a strong, positive and statistically significant effect on local economic development. The effect is monotonically decreasing, i.e., the length of exposure matters.¹⁹

Columns 2—5 sequentially include a set of post-independence and pre-independence controls. First, column 2 includes the percentage of the population in the district that belong to scheduled castes (SC) and scheduled tribes (ST). The Scheduled castes (SC) and Scheduled Tribes (ST) are the more disadvantaged population groups in India. A greater population share of SC and ST in the district is indicative of lower prosperity. In column 3 we also include the percentage of the population in the district that are literate. A greater proportion of the population being literate is likely to be associated with greater prosperity independent of railway introduction and omitting this variable might over-estimate the effect of railway exposure. In Column 5 we include other geographical characteristics proportion of the district with mean elevation 500m or more and the number of Km of river in the district.²⁰ Finally in column 6 we include a set of pre-independence (as of 1930) characteristics that can potentially affect the level of economic development in the district. These include whether the district was under direct British Rule, real income of the district as of 1930 and area under cultivation in the district as of 1930. Notice however, the number of observations (number of districts in the sample) drops when we include these additional characteristics.

The main story holds even when we include these additional district level controls. While the magnitude of the exposure effects fall (they are quantitatively lower in columns 2 and 3), qualita-

¹⁹In Table 2 we present results with state clustered standard errors. The results using spatially corrected standard errors are presented in A2.

²⁰As an alternative to mean elevation, we re-estimated equation (1) using ruggedness of a district as a geographical control. The results, presented in columns 4 and 5 in Table A3 are similar to those presented in Table 2.

tively the effects are very similar and importantly they continue to remain statistically significant. It is worth noting that an increase in the proportion of ST in the district population is associated with a significant decline in district level prosperity. The regression results in column 3 show that an increase in the proportion of the population in the district that is literate is associated with a significant increase in the district level prosperity, though the effect becomes weaker once we control for geographical characteristics. These geographical characteristics are not correlated with district level prosperity. Finally, the exposure effects are robust to the inclusion of pre-independence controls. The pre-independence characteristics themselves are not correlated with district level prosperity.

One final concern is whether the results driven by the liberalization that the Indian government enacted in 1992. To examine whether this is the case, we re-estimate the specification given by equation (1) using district prosperity as of 1992 as our dependent variable. The results presented in columns 1–3 of Table A3 are qualitatively similar to those presented in Table 2, indicating that the results are not driven by the liberalization.

5.3 The 1880 Effect

One consistent result that we obtain is that connection prior to 1880 is associated with a large, positive and statistically significant effect on long term prosperity. Regions that get access to infrastructural investments early, benefit more from such investments. Why do we see this break at 1880? 1880 turns out to be a fairly water-shed year as far as the development of Indian railways is concerned. We present in this section, how Indian railways changed in two dimensions post–1880 and how they could have impacted long term prosperity. While we cannot draw a causal link with either of these and the long run difference in productivity, they both possibly accounted for this difference in some way.

Change in Financing Arrangements and the New Guarantee System:

The post–1880 period saw the introduction of a new system of financing railways in colonial India: through investment by private companies who had guaranteed returns on their investment. This is termed as the *New Guarantee System*. Private companies building railways in India was nothing new. Even during the period 1853–1870, railways were funded by private companies who were guaranteed returns on their investment. However, the government of India had a major say in designing the trunk routes that were build during this period. These routes were built with British military/political and commercial interests in mind, and connected the major ports of India with

the major provincial cities. There was a state take over of railways during the 1870s, though by the end of this period, the favoured policy regarding further construction of railways was to invite private British companies to invest in railways, again with guaranteed returns. This time, however, the government of India ceased to play an active role in determining where railways should be build. After 1880, making a quick profit became salient, but the actions of the investors had no down-side due to the guarantee on investment. [Dutt \(1900\)](#) argued while the initial trunk routes were built with the trade and economic development of India in mind, railway lines in this period were *pushed through by capitalists and speculators* (see [Sweeny and Hall, 2008](#), page 148).²¹

Break in Gauge:

The second change that happened around this period was what is now known as *break in gauge* in Indian railways. Starting with Lord Dalhousie, the initial preference of the government was to have uniform gauge across the country and the 5'6" (broad) gauge was chosen. However, over time cost issues grew increasingly important. The high cost of constructing broad gauge lines left the government of India with no alternative but to recommend a *less expensive narrow gauge railway for parts of the country which could never remunerate the higher standard and would have otherwise gone without railways altogether* ([Sanyal, 1930](#), page 19). Therefore, Post-1880, the majority of new lines that were built were non-broad gauge; correspondingly the majority of district connected to railways after 1880 were connected by non-broad gauge. This is clear from the left panel of [Figure 5](#). While pre-1880, 76% of first connections were broad gauge, the corresponding proportion was 42% for districts connected between 1881 and 1930; difference statistically significant (p -value = 0.00). There is a persistence in railway gauges within a region as network externalities imply that moving across gauges is expensive. This meant that follow-up lines in the districts used the same gauge as the initial connection. The right panel of [Figure 5](#) makes this clear. For districts where the first connection was broad gauge, majority of follow-up connections were also broad-gauge and the opposite holds for districts where the first connection was non-broad gauge.

[Table 3](#) presents the regression results for the effect of gauge at initial connection in the district on district prosperity in 2013. The reference category in this regression is that the district was not connected by 1930. While the difference (broad gauge — non-broad gauge) is never statistically significant, district prosperity in 2013 is higher in districts initially connected by broad gauge, relative to districts connected by non-broad gauge (i.e., metre gauge and narrow gauge).

²¹This view is shared by [Sanyal \(1930\)](#), [Thorner \(1955\)](#), [Dubey \(1965\)](#), [Satya \(2008\)](#), who all argue that the new guarantee system was an extractive institution. [Bogart and Chaudhary \(2019\)](#) argue that the system did not generate significant returns.

Add to this the fact that while the first generation of districts connected by the trunk routes were the ports, the key provincial cities and the districts in between, post-1880, the majority of districts that were connected were the more remote districts of India.²² But as a result of the break of gauge, these remote, and later connected districts were less well integrated to the rest of the country: the break of gauge meant an additional transaction cost in trade and transportation, as these district were not directly connected to the ports and major centres.

5.4 Selection Issues

A natural concern when estimating the impacts of infrastructure projects is the possibility of selective program placement: for example, if railways were first introduced in districts that had greater economic potential, this could perhaps explain large contemporaneous and long-term impacts. The lines built prior to 1880 were part of Lord Dalhousie’s trunk routes: it is feasible that the trunk routes were placed in districts that were already well developed and the observed prosperity a century later is simply a persistence effect. In other words, connectivity might be correlated with unobserved district level characteristics resulting in biased estimates of the effect of early connectivity on district level prosperity. We address this question in two different ways. First, we review the existing historical literature on British motivations on railway construction and argue that there is no reason to believe that the placement of lines specifically targeted economically more prosperous districts. Second, we conduct an IV estimation. We use as instrument, the distance of the centroid of a district from the centroid of the nearest district that is on the shortest path joining the major cities.

British Motivations:

The British motives for building railways in India can be categorized into military, commercial and social motives, in roughly that order of importance. Of the different reasons why the colonial government ultimately agreed to build a railway network in India, the political and military aspects of railways (internal administration, reduced military expenditure, defense from external aggression) carried the largest weight in the government’s decision-making process. This was the *military motive*. The military (political) motive for building railways in colonial India always triumphed all other motives (see [Sanyal, 1930](#), [Macpherson, 1955](#), [Sharma, 2010](#), [Donaldson, 2018](#)). The initial burst of railway construction was in the period immediately after the Sepoy

²²This was partly an outcome of the report of the Famine Commission of 1880, which recommended that the railways also play a “protective” as opposed to only a “productive” role. Figure B4 shows that indeed the first connections were non-broad gauge in the more remote districts.

Mutinee (see Panel A of Figure 3). The expansion of railways to the North-West (beyond 1865 and into the 20th century) was driven by strategic motives: part of the Great Game against Russia, in response to Russia's annexation of territories in Central Asia and, the construction of the Transcaspian Railway.

There was of course also British commercial motives for expanding railway system in colonial India. It was the mercantile firms in London and Manchester with trading links with India that initially advocated the development of railways and indeed through the 19th century, railways in colonial India was funded through private capital. For example, industrialists in Britain wanted railways to connect ports in India to the cotton growing regions of India, particularly after the American cotton production stagnated. The interior regions of India were also viewed as important markets for mass produced goods imported from Britain. This was the *commercial motive*. See, for example, Sanyal (1930), Macpherson (1955), Thorner (1955), Sharma (2010), Hurd and Kerr (2012). Finally, the British government also viewed the railways were regarded as a valuable instrument in the improving of Indian social conditions, the spreading of Western civilization and the elevation of backward peoples. The colonial government also hoped that railways would achieve the prevention of local famine and the uniform dispersion of food. This could be thought of as the *social motive*.

However, there is no evidence to suggest that the British specifically targeted prosperous districts for early connection. While it was unanimously agreed that the railways would connect the major cities of the country and the ports, in the early decades of railway construction, there were many debates as to what route should be adopted to connect these cities. Major Kennedy, the chief engineer of India wanted to connect the major cities with a network of lines that followed the gentlest possible gradients, along river gradients and the coastline, wherever possible. Part of the reason for this was cost. Ultimately the Kennedy routes were not adopted and the routes followed roughly a straight line path between two cities, which often ignored geography and hence agricultural productivity.²³

Finally, Donaldson (2018) argues that the networked nature of railways meant that it was hard to target specific locations and also that it was hard for technocrats to agree on plans. Importantly, there is no evidence to suggest that timing of railway connectedness was correlated with actual prosperity. Indeed, Donaldson (2018) argues that an increase in prosperity was a consequence of the introduction of the railway line in the district and not the other way round.

²³The Kennedy routes that were not adopted, would have exposed regions of India that had high agricultural productivity to railways earlier. This is because the gentle gradients that would have allowed low cost construction would be correlated with greater agricultural productivity.

IV estimation:

The question of selective placement of railways prior to 1880 becomes more important given the preference for the trunk routes in the pre-1880 period. Given the military motive, the first aim was to connect the major cities, the exact route between the cities were chosen by balancing the commercial and political interests.²⁴ Ultimately, more direct connections were preferred, despite potentially adding to costs. We construct a hypothetical shortest path between the major cities as the straight line joining pairs of major cities. This is a purely theoretical construct and does not take into account geographical or other factors that might affect railway construction. Districts on the path thus constructed are on the shortest distance path. While the darker shaded districts in the left panel of Figure A1 are the districts actually connected by 1880, the darker shaded districts in the right panel of Figure A1 are the districts on the shortest distance path. These districts are those that would have been connected had the hypothetical shortest distance path been adopted. While not identical, these two sets of districts are close. We use as instrument the minimum distance of (the centroid) of a district to the (centroid) of a district on the shortest distance path. This distance is likely to affect actual construction of railways, but there is no reason why it should affect district prosperity a century later.

We estimate a restricted version of equation (1):

$$\log(1 + \text{sumlight}_{dst}) = \beta_0 + \beta_1 \text{Pre-1880} + \gamma \mathbf{X}_{dst} + \lambda_s + \varepsilon_{dst} \quad (4)$$

The OLS and IV regression results for equation (4) are presented in Table 4. In Columns 1 and 2 the estimating sample includes all districts in present day India, Bangladesh and Pakistan while in columns 3–6, the sample is restricted to districts in present day India. The distance from the line of shortest distance is negatively correlated with a district being connected pre-1880. Also the first stage F values all exceed 10, supporting instrument validity. Finally, while the OLS and the IV regression results both support the argument that pre-1880 connection is associated with a large, positive and statistically significant increase in district prosperity in 2013, the IV estimates are larger in magnitude than the corresponding OLS estimates. This is *per se* not surprising because measurement error in our measure of district level prosperity is likely to lead to attenuation bias, i.e., bias the OLS estimates towards 0. The OLS estimates can, therefore, be thought of as lower bounds of the effects of connection pre-1880 on long term prosperity

Given that we do not find any evidence of selection bias in placement of railway lines, for the rest

²⁴Lord Dalhousie’s Railway Minutes prescribed *trunk lines* that would connect India’s major cities along direct routes and maximize the *political advantages* of a railroad network. Other’s preferred commercial routes. See [Sanyal \(1930\)](#) for more on this debate.

of the paper we will assume that railways were indeed placed randomly.

5.5 Falsification Tests

We also conduct a set of falsification tests. We randomly assign decade of introduction to districts and estimate equation (1) for all districts in present day India, Bangladesh and Pakistan, with the dependent variable as $\log(1 + \text{sumlight}_{ds})$ in district d in state s in 2013. We hold the proportion of districts that were connected in each decade as given. We repeat this exercise 10000 times and this gives us 10000 estimates of β_i 's. The average of these 10000 estimates gives us the simulated value of β_i 's and the confidence interval is computed from the bootstrapped estimates using the Hall's Percentile method. These are presented in the right panel of Figure 6. The left panel of Figure 6 presents the actual estimates of β_i 's, i.e., using the true values of year of introduction of railways in the district. A comparison of the actual and simulated estimates suggest that the results are not driven by any random event, unrelated to the introduction of the railways.

5.6 Spatial Spillovers

Figure 7 presents the non-parametric lowess plots of the association between night time lights as of 2013 and the distance from the nearest connected district in year T . The negative relationship between distance from the nearest connected district and night time lights (local economic development) is clear, irrespective of T . What is interesting, however, is the strength of the relationship over different values of T . In particular, notice that the relationship is considerably weaker in the early decades of railway construction than in the later decades of railway construction in the Indian sub-continent. As more and more districts are connected, the effect of an increase in the distance of a non-connected district to the nearest connected district on night lights becomes stronger. In short, an increase in the distance from the nearest connected district is associated with an adverse effect on local economic development in the long run. The geographical spillovers effects become stronger over time.

The regression results corresponding to the regression specifications (2) and (3) are presented in Table 5. Each cell in this table presents the results from a different regression (each for a different value of T). Columns 1—3 presents the OLS estimates of δ_1 from from equation (3), for the different values of t while columns 4—6 presents the corresponding estimates of α_1 from equation (2). In columns 3 and 6 we include additional post-independence and geographic controls (as in column 5 of Table 2).

Consider first the cumulative effects of the district being connected by a given time T on district level prosperity (results in columns 1—3). The estimated effects are always large, positive and statistically significant, indicating again that connectedness has a beneficial effect on local economic development. Columns 2 and 3 demonstrate that the effects of connectedness decreases over time: being connected as of 1860 has a stronger effect on prosperity in 2013, relative to being connected as of 1870, which in turn has stronger effect relative to being connected as of 1880. Cumulative duration of exposure is thus important. The inclusion of additional post-independence controls does not make any qualitative difference.

Turning now to the results in columns 4—6: the effect of distance from the nearest connected district at a given time t on night lights. The effects are always negative indicating an increase in distance (in 100 kms) from the nearest connected district is associated with a decrease in night lights in 2013. Consistent with Figure 7, the magnitude of the negative impact of being further away from a connected district increases over time.

5.7 Effect of Intensity of Railway Construction

The analysis thus far examines the association between whether or not a district is exposed (and the length of the exposure) and long run prosperity. However, intensity of railway construction could also be important, where intensity is measured by the number of kilometres of railway track in the district as of a particular time T . We estimate a variant of equation (3):

$$\log(1 + \text{sumlight}_{ds}) = \xi_0 + \xi_1 L_{dst} + \gamma \mathbf{X}_{dsT} + \lambda_s + \eta_t + \varepsilon_{ds} \quad (5)$$

where L_{dst} is the number of kilometres of railway track in the district at time T ; ($T = 1860, 1870, 1880, 1890, 1900, 1910$ and 1930). The dependent variable is again $\log(1 + \text{sumlight}_{ds})$ in district d in state s as of 2013. The corresponding regression results are presented in Table 6. Each cell in this table presents the results from a different regression. In column 1, the sample includes all districts in present day Bangladesh, India and Pakistan; in columns 2 and 3, the sample is restricted to districts in present day India. We find that an increase in length of track in a district is associated with a large increase in district level prosperity in 2013. One common element across all three columns is that over the time the effect of intensity of connectivity falls, even though the magnitude remains positive and statistically significant.

5.8 Alternative Measures of District Level Prosperity

We examine the robustness of our main results using alternative measures of district level prosperity.

Mean Light: First, instead of using $\log(1 + \text{sumlight}_{ds})$ as the dependent variable we use $\log(1 + \text{meanlight}_{ds})$. The regression results corresponding the specification given by equation (1) are presented in Table 7. In column 1 the sample includes all districts in present day Bangladesh, India and Pakistan. In columns 2–6, the sample is restricted to all districts in present day India. The results are qualitatively similar to those presented in Table 2. Introduction of railways prior to 1880 has a strong, positive and statistically significant effect on local district prosperity. The estimated effects fall monotonically, consistent with the explanation that length of exposure is important.

Aggregate Bank Deposits: In Table 8, the dependent variables are aggregate bank deposits (columns 1–3) and percentage of rural and urban population categorized as poor (columns 4–6 and 7–9 respectively). Our results in columns 1–3 are similar to our results in Table 2. Aggregate bank deposits are significantly higher in districts exposed prior to 1890 and there is clear evidence that the effects are falling over time: again the length of exposure is important.

Rural and Urban Poverty: Columns 4–6 and 7–9 present the proportion of the rural and urban population in the district categorized as poor as of 2004–2005.²⁵ First, as the regression results presented in columns 7–9 show, connectedness prior to 1930 (irrespective of when connected) does not have a statistically significant effect on the percentage of the urban population categorized as poor. The results in columns 4–6, however, tell a somewhat different story. The effects of early exposure (pre-1900) are large, negative and often statistically significant. The estimates though become imprecise when we include pre-independence controls (see column 6). Overall evidence suggests that districts connected pre-1900 have a significantly lower percentage of the rural population categorized as poor; there are no corresponding effects on urban poverty rates.

²⁵Here we do not include the proportion of the district population that are SC and ST and the proportion of the population that are literate as additional controls.

5.9 Long-Run Impacts on Health and Education

The results presented thus far show that increased exposure to railways is associated with a significantly higher prosperity and lower rural poverty at the district level. The logical subsequent question is thus whether this increased exposure is associated with an improvement in other spheres. In particular, is increased exposure associated with an increase in human capital of adult males and females? We address this question using the IHDS2 data. Note that funding for schools, hospitals and health centres is provided by the state government, the prosperity of a district will not necessarily have a direct effect on health and education in the district. Typically state governments use some form of a budgetary allocation rule to determine how each district is to be funded and this rule might not be correlated with district level prosperity.

We estimate versions of equation (1), though the outcome variables are different. They are the proportion of adult males and females in the district that have no schooling (results presented Table 9) and the proportion of adult males and females that are underweight, which is a measure of malnourishment (results presented in Table 10). In general, evidence suggests that longer exposure to railways reduces the fraction of adult males and females who have no education, though the effect is stronger for women. The estimates are, however, considerably more imprecise when we include pre-independence controls (columns 5 and 6). The results presented in Table 10 show that relative to districts not connected before 1930, districts connected before 1900 have a significantly lower proportion of adult men and women who are underweight. Length of exposure matters as the estimated effects decrease over decades. To the extent being underweight is an indicator of being malnourished, these results suggest that increased exposure to railways is associated with a significant decline in the proportion of the population that is malnourished. It is worth noting that while both men and women benefit from increased exposure to railways, women benefit more than men.

Could these positive effects be driven by migration of more educated and healthier adults migrating to the richer districts? While we cannot explicitly rule this out, we believe that this is unlikely to be driving the results as it is now accepted that the extent of geographical mobility in India is amongst the lowest in the developing world (see, for example [Munshi and Rosenzweig, 2009](#), [Brainerd and Menon, 2014](#)). Data from the Indian census indicates that around 20 percent of the growth in the urban population over the period 1991–2001 can be attributed to migration. However, permanent migration has remained low: most migration is temporary in nature with men moving to the cities in search of low skilled jobs and then moving back to their original (rural) residence during harvest.

6 Mechanism

What explains these results? Recall that the immediate impact of railway construction in colonial India was on trade and development in the predominantly agricultural country. As [Donaldson \(2018\)](#) argues, the introduction of railways led to a significant increase in trade, both inter and intra national. One possibility is that regions connected to railways early (i.e., early exposure districts) experienced positive agglomeration effects or increased returns to scale effects. Districts connected later did not and over time cannot catch up to the early movers. The positive effects persist over a century. This is consistent with the argument by [Jedwab and Moradi \(2016\)](#). They argue that initial infrastructural investments have large effects in developing countries, leading to structural changes in these economies. Subsequent investments though have less significant effects because they are unable to affect the structure of the economy in the manner and to the extent initial infrastructural investments could. We have evidence of similar early mover effects in developed countries as well. [Fernald \(1999\)](#), while examining the impact of the building of roads in the US find that the slow-down in the growth of roads after 1973 had an adverse effect on productivity. Why was 1973 a break-point? The pre-1973 period was characterised by investment in the interstate highway system (construction of the interstate highway system was largely completed by 1973), which was the driver of productivity growth in the US. Building a second network, post-1973 was not productive.

We use population density of a district as our measure of agglomeration effect: the argument being that as regions develop, more people move to these regions. We estimate the effect of pre-1880 railway exposure on population density of a district across different decades. The districts are defined as in 2001.²⁶ We present in Panel A of Figure 8 the probability density function of the distribution of population in averaged across the decades 1901–2001 in each district as defined by the 2001 census of India. Notice that the mass of the density of population (log population density) is to the right for districts connected pre-1880, relative to districts connected between 1880 and 1930. While the sample in the left panel includes all districts, that in the right panel excludes districts that had a large city/town with population more than 100000 in 1881.²⁷ The null hypothesis of equality of distributions is rejected using a Kolmogorov and Smirnov test ($p - value = 0.00$ in both the full and the restricted sample). This gives us the first indication of positive agglomeration effects in districts connected early.

To understand further the role of early connection on population growth in the district we estimate

²⁶See [Tumbe \(2012\)](#) and [Tumbe \(2015\)](#) for more on this data.

²⁷This data is obtained from the Digital South Asia Library (<https://dsal.uchicago.edu/statistics/index.html>).

the following regression:

$$D_{dst} = \alpha_0 + \alpha_1 \text{Pre-1880} + \lambda_s + \eta_t + \varepsilon_{dst} \tag{6}$$

where D_{dst} denotes the log of population density in district d , in state s in decade t ; Pre-1880 is a dummy variable indicating that the district was connected in Pre-1880; λ_s and η_t denote a set of state and decade dummies. The regression results are presented in Table 11. In Column 1 we include all districts, while in column 2 we exclude districts that had a large city/town with population more than 100000 in 1881. Population density is significantly higher in districts connected Pre-1880 further supporting the agglomeration effects.

One could, however, argue that railways were introduced in districts with higher population density in the first place and the effects that we observe in Table 11 is simply the effect of high population to begin with and nothing to do with agglomeration effects. It is difficult to rule this out fully, given that we do not have any data on the district population before railways were introduced in the district. The first British census was conducted in 1871, which was after the railways had been introduced in the country. While we do not have systematic district level population data pre-1871, from Chandler and Fox (1974) we can obtain estimates of population in a set of cities and towns in the Indian sub-continent as of 1850. The 1901 Census of India provides us data on population in these cities and towns as of 1901. Controlling for population in 1850, connectivity pre-1880 leads to a statistically significant higher population in 1901 (p-value = 0.07).²⁸ So, connection pre-1880 leads to city growth in 1901 and beyond 1901, the population density is greater in districts connected pre-1880.

We also conduct the following falsification test. We construct a placebo pre-1880 railway introduction dummy variable and estimate equation (6), including the placebo dummy in the place of the real one. The districts thus assigned the placebo pre-1880 introduction are randomly chosen and we hold the proportion of districts connected pre-1880 constant. We repeat this procedure 10,000 times and save the estimated coefficients. We do this for all districts and excluding the districts with the large cities and towns as of the 1881 census. We present in Panel B of Figure 8 the probability density function of the 10,000 placebo point estimates of α_1 (from equation (6)) and a vertical line indicating our true point estimate. The purpose of this test is to check how many times these randomly generated placebo point estimates happen to be smaller or too close to our true point estimate. If in our main results we were erroneously rejecting the null hypothesis that our coefficient of interest is equal to 0 (i.e., we were attributing to railway introduction a positive effect that does not exist in reality), we should have observed placebo coefficients that are

²⁸The regression specification is given by $\log \text{Pop}_{1901} = \eta_0 + \eta_1 \text{Pre-1880} + \eta_2 \log \text{Pop}_{1850} + \epsilon$. Full set of results available on request. Not surprisingly, population in 1850 is a strong predictor of pre-1880 connection: recall that Lord Dalhousie’s trunk routes, the first set of lines that were built were designed to connect the major cities of the Indian subcontinent.

very close to our true estimate. However, as we see from Panel B of Figure 8 the point estimates generated in the falsification test are almost always to the left of (implying smaller in value than) the true estimated coefficient α_1 . This exercise suggests that pre-1880 railway introduction did result in increased population density in the district.

7 Conclusion

The primary objective of this paper is to understand the long-term impacts of infrastructural development. Related is the role of the governments in choosing which locations they want to invest in. This is because, there exists a remarkable degree of temporal persistence in economic activity. We consider the long-term impact of the building of railways in colonial India during the second half of the 19th century and the early part of the 20th century. While the colonial British government had a number of different reasons to invest in railway infrastructure, there is no reason to believe that railways were selectively placed depending on the actual prosperity of the district. Despite this, the immediate impact of railway construction during that period was on trade and development in the predominantly agricultural India; the introduction of railways resulted in significant increase in trade both domestically, (i.e., within different parts of the country) and internationally, reduction of trade costs and price differences and an increase in real incomes in connected districts. In this paper, we show that these effects were not short lived. Indeed, we see that the effects persist over more than a century. Districts of the Indian sub-continent that were connected to railways early continue to have higher levels of economic prosperity (as measured by night time lights and aggregate bank deposits) and lower rural poverty rates a century later, more men and women have some education and fewer men and women are malnourished. We also find evidence of considerable geographical spillovers: districts further away from connected districts are worse off in terms of levels of economic development in 2013. The evidence is consistent with the idea that the transitory shock of early connectivity to railways results in path dependence in the location of economic activity and permanent effects on economic development, with little evidence of catch-up.

Using instruments we estimate the causal effect of early railway connectedness on long term prosperity of the district. Like the OLS estimates, the IV estimates of early railway connectedness are positive and statistically significant. The IV estimates are larger, indicating that the OLS estimates provide a lower bound on the effect of exposure to railways on long term prosperity.

In terms of mechanisms we argue that the results are driven by the fact that railways significantly benefited trade in the predominantly agricultural India. Regions connected to railways early experienced positive agglomeration effects. Initial infrastructural investments have large effects,

leading to structural changes. Subsequent investments though have less significant effects because they are unable to affect the structure of the economy in the manner and to the extent initial infrastructural investments could. Additionally, while the colonial British government did not target specific districts based on their economic potential, the connected districts were exposed to trade (both inter-regional and international) early. In that sense the results in this paper could also be viewed as the long run impacts of trade and globalization.

References

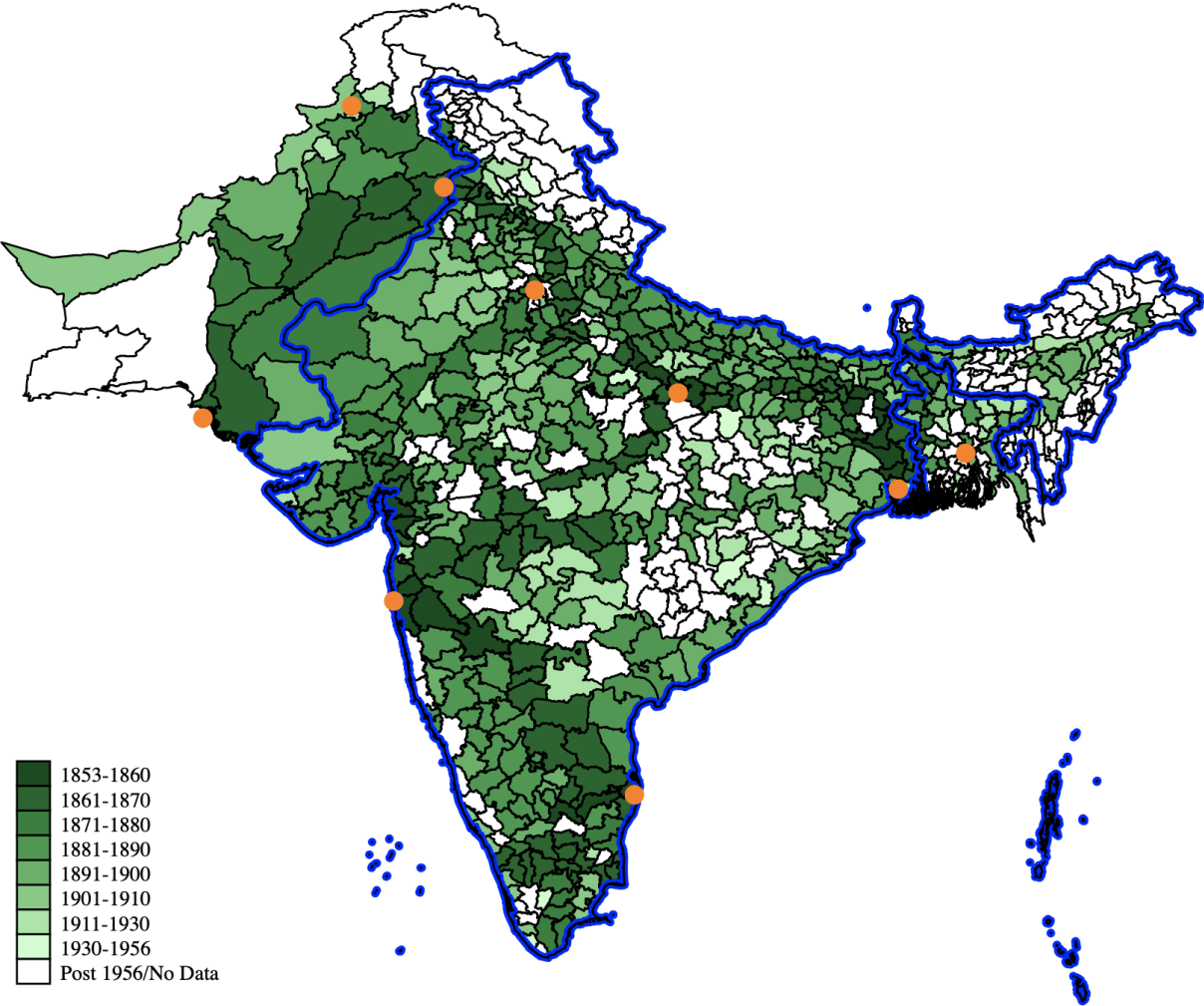
- Asher, S., T. Lunt, R. Matsuura, and P. Novosad (2021). Development research at high geographic resolution: An analysis of night lights, firms, and poverty in india using the shrug open data platform. *World Bank Economic Review Forthcoming*.
- Atack, J. and R. A. Margo (2011). The impact of access to rail transportation on agricultural improvement: The american midwest as a test case, 1850-1860. *Journal of Transport and Land Use*, 4(2), 5 – 18.
- Banerjee, A., E. Duflo, and N. Qian (2020). On the road: Access to transportation infrastructure and economic growth in china. *Journal of Development Economics Forthcoming*.
- Banerjee, A. and L. Iyer (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in india. *American Economic Review* 95(4), 1190 – 1213.
- Berger, T. and K. Enflo (2016). Locomotives of local growth: The short- and long-term impact of railroads in sweden. *Journal of Urban Economics* 98(C), 124 – 138.
- Bhandari, L. and K. Roychowdhury (2011). Night lights and economic activity in india: A study using dmsp-ols night time images. *Proceedings of the Asia-Pac Advanced Network* 32, 218 – 236.
- Bogart, D. and L. Chaudhary (2016). *Railways in Colonial India: An Economic Achievement?* Routledge Press.
- Bogart, D. and L. Chaudhary (2019). Extractive institutions? investor returns to indian railway companies in the age of high imperialism. *Journal of Institutional Economics* 4 15, 751 – 774.
- Brainerd, E. and N. Menon (2014). Seasonal effects of water quality: The hidden costs of the green revolution to infant and child health in india. *Journal of Development Economics* 107, 49 – 84.
- Bruederle, A. and R. Hodler (2018). Nighttime lights as a proxy for human development at the local level. *PLoS ONE* 13(9).
- Burgess, R. and D. Donaldson (2012). Railroads and the demise of famine in colonial india. Report, Mimeo, London School of Economics.
- Caicedo, F. V. (2019). The mission: Human capital transmission, economic persistence and culture in south america. *Quarterly Journal of Economics*, 507 – 556.
- Calvi, R. and F. Mantovanelli (2018). Long-term effects of access to health care: Medical missions in colonial india. *Journal of Development Economics* 135, 285 – 303.
- Calvi, R., F. Mantovanelli, and L. Hoehn-Velasco (2020). The protestant legacy: Missions and human capital in india. *Journal of Human Resources Forthcoming*.

- Castelló-Climent, A., L. Chaudhary, and A. Mukhopadhyay (2018). Higher education and prosperity: From catholic missionaries to luminosity in india. *Economic Journal* 128(616), 3039 – 3075.
- Chandler, T. and G. Fox (1974). *3000 Years of Urban Growth*. Academic Press, New York and London.
- Chaudhary, L. and J. Fenske (2020). Did railways affect literacy? evidence from india. Technical report, Warwick University.
- Chaudhuri, S. and N. Gupta (2009). Levels of living and poverty patterns: A district-wise analysis for india. *Economic and Political Weekly* 44(9), 94 – 110.
- Coatsworth, J. H. (1979). Indispensable railroads in a backward economy: The case of mexico. *Journal of Economic History* 39(4).
- Colella, F., R. Lalive, S. O. Sakalli, and M. Thoenig (2019). Inference with arbitrary clustering. Technical report, IZA Discussion Paper # 12584.
- Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of Econometrics* 92(1), 1 – 45.
- Cook, C. J. and M. Shah (2020). Aggregate effects from public works: Evidence from india. *Review of Economics and Statistics* Forthcoming.
- Dalgaard, C.-J., N. Kaarsen, O. Olsson, and P. Selaya (2018). Roman roads to prosperity: Persistence and non-persistence of public goods provision. Report.
- Dar, A. (2019). “all is water”: Technological complementarities and path dependence in indian agriculture. Technical report, Indian School of Business.
- Dell, M. (2010). The persistent effects of peru’s mining mita. *Econometrica* 78, 1863 – 1903.
- Deloche, J. (1994). *Transport and Communications in India Prior to Steam Locomotion. Volume I: Land Transport*. Oxford: Oxford University Press.
- Donaldson, D. (2018). Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review* 108(4-5), 899 – 934.
- Donaldson, D. and R. Hornbeck (2016). Railroads and american economic growth: A “market access” approach. *Quarterly Journal of Economics* 131(2), 799 – 858.
- Dubey, V. (1965). Railways. In V. B. Singh (Ed.), *Economic History of India*. Allied Publishers.
- Duflo, E. and R. Pande (2007). Dams. *Quarterly Journal of Economics* 122(2), 601 – 646.
- Dutt, R. C. (1900). *Open Letters to Lord Curzon on Famines and Land Assessments in India*. Gian Publishing House, 1986.
- Fernald, J. G. (1999). Roads to prosperity? assessing the link between public capital and productivity. *American Economic Review* 89(3), 638.
- Florida, R., T. Gulden, and C. Mellander (2008). The rise of the mega-region. *Cambridge Journal of Regions, Economy and Society* 1(3), 459 – 476.
- Fogel, R. W. (1962). A quantitative approach to the study of railroads in american economic growth: A report of some preliminary findings. *Journal of Economic History* XXII(2), 163 – 197.
- GADM (2012). Database of global administrative areas.
- Henderson, V., A. Storeygard, and D. N. Weil (2011). A bright idea for measuring economic growth. *American Economic Review* 101(3), 194 – 199.
- Henderson, V., A. Storeygard, and D. N. Weil (2012). Measuring economic growth from outer space. *American Economic Review* 102(2), 994 – 1028.

- Herranz-Loncán, A. and J. Fourie (2018). “for the public benefit”? railways in the british cape colony. *European Review of Economic History* 22(1), 73 – 100.
- Hodler, R. and P. A. Raschky (2014). Regional favoritism. *The Quarterly Journal of Economics* 129(2), 995 – 1033.
- Hornung, E. (2015). Railroads and growth in prussia. *Journal of European Economic Association* 13(4), 699 – 736.
- Hurd, J. and I. J. Kerr (2012). *India’s Railway History : A Research Handbook*. BRILL.
- Iyer, L. (2010). Direct versus indirect colonial rule in india: Long-term consequences. *Review of Economics and Statistics* 92(4), 693 – 713.
- Jedwab, R., E. Kerby, and A. Moradi (2017). History, path dependence and development: Evidence from colonial railways, settlers and cities in kenya. *Economic Journal* 127, 1467 – 1494.
- Jedwab, R. and A. Moradi (2016). The permanent effects of transportation revolutions in poor countries: evidence from africa. *Review of Economics and Statistics* 98(2), 268 – 284.
- Katz, O. (2018). Railroads, economic development, and the demographic transition in the united states. Technical report, Brown University.
- Kelly, M. (2019). The standard errors of persistence. Technical report, University College Dublin.
- Lehne, J. (2019). Irrigation vs education: The long-run effects of opium cultivation in british india. Report, Paris School of Economics.
- Macpherson, W. J. (1955). Investment in indian railways, 1845-1875. *The Economic History Review, New Series* 8(2), 177 – 186.
- Michalopoulos, S. and E. Papaioannou (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica* 81, 113 – 152.
- Munshi, K. and M. R. Rosenzweig (2009). Why is mobility in india so low? social insurance, inequality, and growth. Technical report, NBER, Working paper w14850.
- NOAA (2014). National oceanic and atmospheric administration version 4 dmsp-ols nighttime lights time series, boulder co: National geophysical data center.
- Noor, A., V. Alegana, P. Gething, A. Tatem, and R. Snow (2008). Using remotely sensed nighttime light as a proxy for poverty in africa. *Population Health Metrics* 6, 5.
- Nunn, N. (2009). The importance of history for economic development. *Annual Review of Economics* 1, 65 – 92.
- Nunn, N. (2013). *Historical Development*. North-Holland.
- Nunn, N. and L. Wantchekon (2011). The slave trade and the origins of mistrust in africa. *American Economic Review* 101, 3221 – 3252.
- Paik, C. and J. Vechbanyongratana (2020). Reform, rails, and rice: Thailand’s political railroads and economic development in the 20th century. Technical report, NYU, Abu Dhabi.
- Rudolph, S. H. (1963). The princely states of rajputana: Ethic, authority and structure. *The Indian Journal of Political Science* 24(1), 14 – 32.
- Sahu, A. C. (1977). Genesis and growth of indo-chinese opium monopoly under east india company. *Proceedings of the Indian History Congress* 38, 529 – 533.
- Sanyal, N. (1930). *Development of Indian Railways*. University of Calcutta.
- Satya, L. (2008). British imperial railways in nineteenth century south asia. *Economic and Political Weekly* 43(47), 69 – 77.

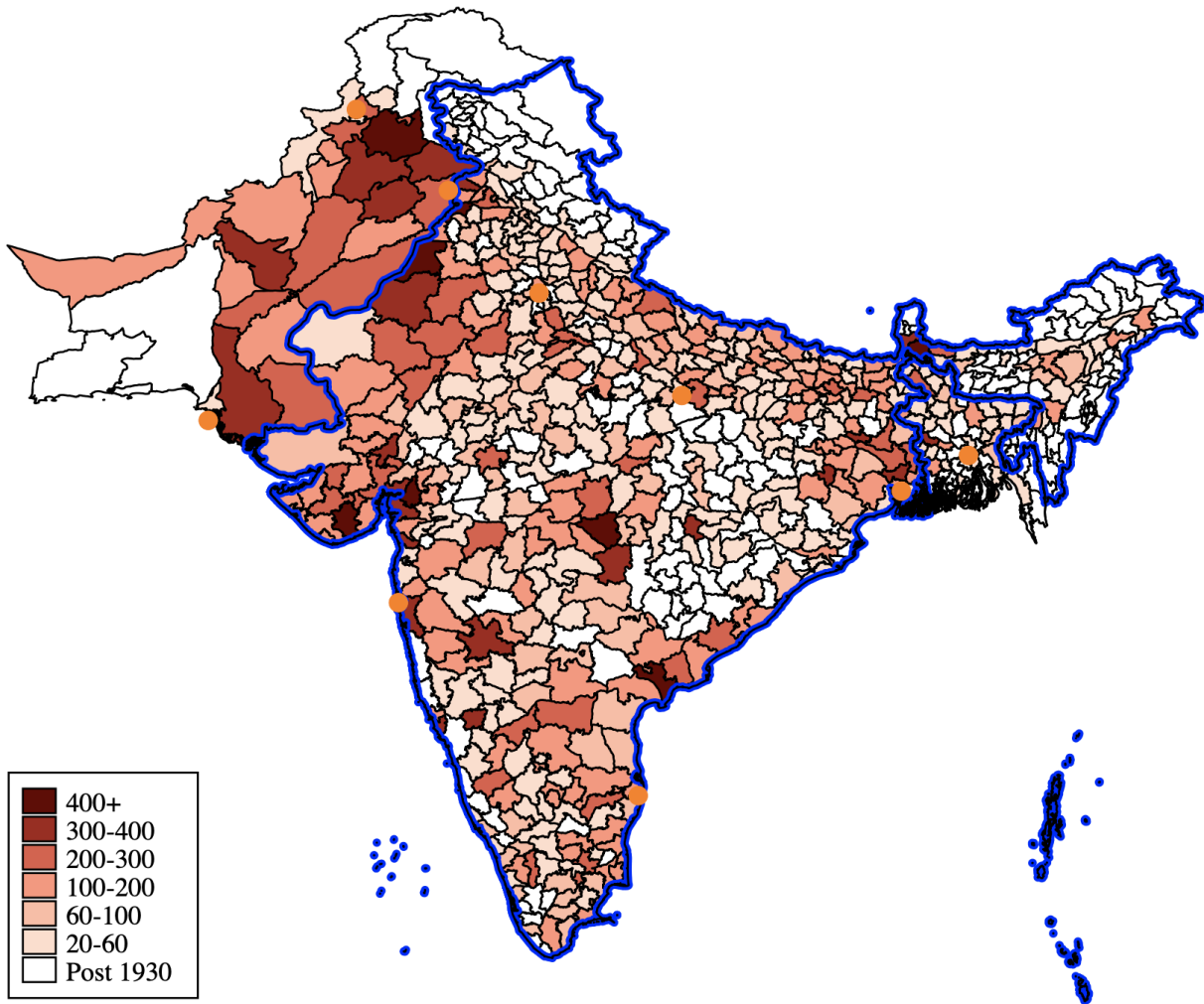
- Sharma, A. K. (2010). “fire-carriages” of the raj: The indian railway and its rapid development in british india. *Essays in History* 43.
- Singhal, A., S. Sahu, S. Chattopadhyay, A. Mukherjee, and S. N. Bhanja (2020). Using night time lights to find regional inequality in india and its relationship with economic development. *PLoS ONE*.
- Spolaore, E. and R. Wacziarg (2013). How deep are the roots of economic development? *Journal of Economic Literature* 51, 325 – 369.
- Summerhill, W. R. (2005). Big social savings in a small laggard economy: railroad-led growth in brazil. *Journal of Economic History* 65(1), 72 – 102.
- Sutton, P., C. Elvidge, and T. Ghosh (2007). Estimation of gross domestic product at sub-national scales using nighttime satellite imagery. *International Journal of Ecological Economics and Statistics*. 8(S07), 5–21.
- Sweeny, S. and M. Hall (2008). Indian railways and famine 1875 - 1914: Magic wheels and empty stomachs. *Essays in Economic and Business History* XXVI, 147 – 158.
- Tang, J. (2014). Railroad expansion and industrialization: Evidence from meiji japan. *Journal of Economic History* 74(3), 863 – 886.
- Thorner, D. (1955). The pattern of railway development in india. *Far Eastern Quarterly*, XIV, 201 – 206.
- Tumbe, C. (2012). *Migration and Remittances in India: Historical, Regional, Social and Economic Dimensions*. Ph. D. thesis, Indian Institute of Management, Bangalore.
- Tumbe, C. (2015). Missing men, migration and labour markets: Evidence from india. *Indian Journal of Labour Economics* 58(2), 245 – 267.
- Wiener, M. J. (2013). The idea of ‘colonial legacy’ and the historiography of empire. *Journal of the Historical Society* 13(1), 1 – 32.

Figure 1: Introduction of Railways in Each District of the Indian Subcontinent



Notes: Sample restricted to districts in present day Bangladesh, India and Pakistan. No data implies district not connected as of 1956.

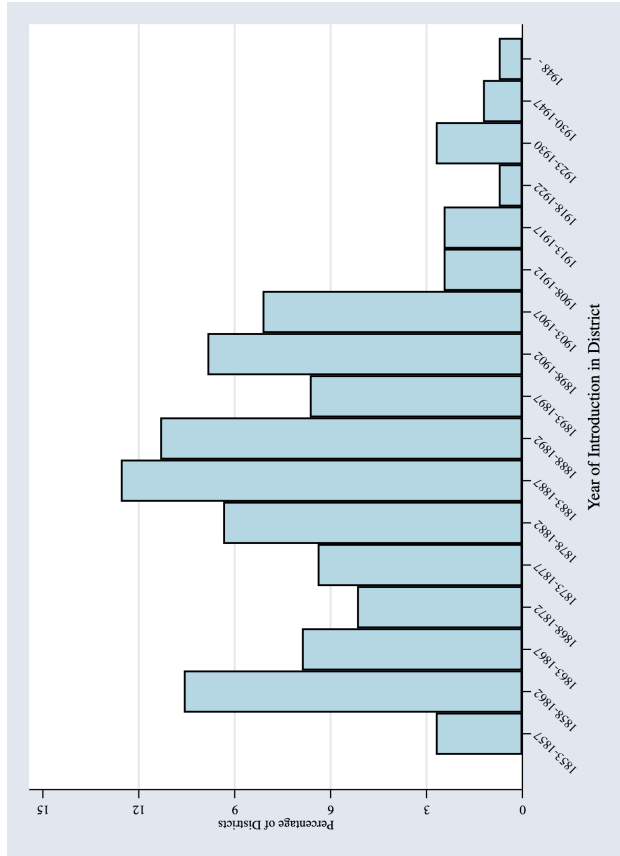
Figure 2: Cumulative Track Length in District as of 1930



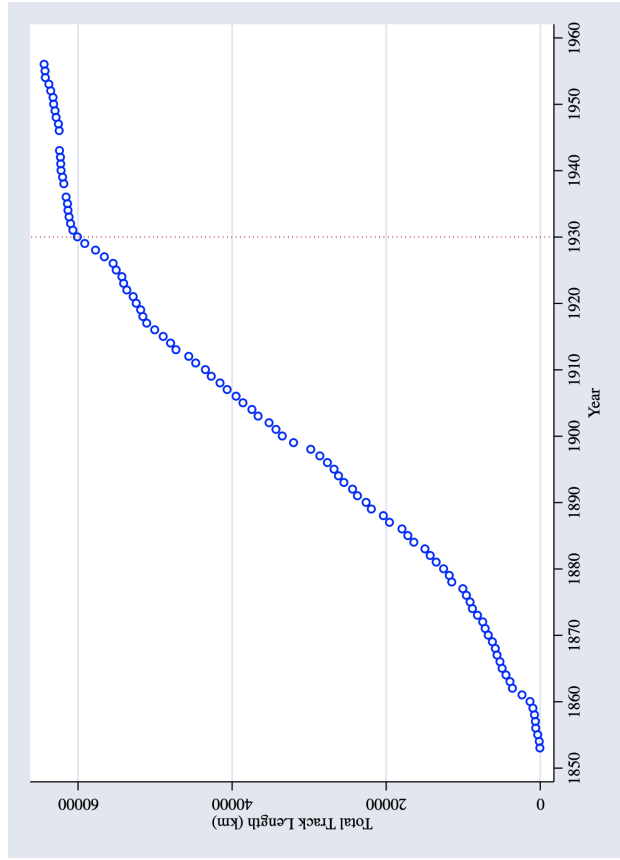
Notes: Data does not include Andaman and Nicobar Islands. Sample restricted to districts in present day Bangladesh, India and Pakistan. No data implies district not connected as of 1956.

Figure 3: Percentage of Districts connected by Railways and Total Length of Tracks in each year.

Panel A: Percentage of District Connected

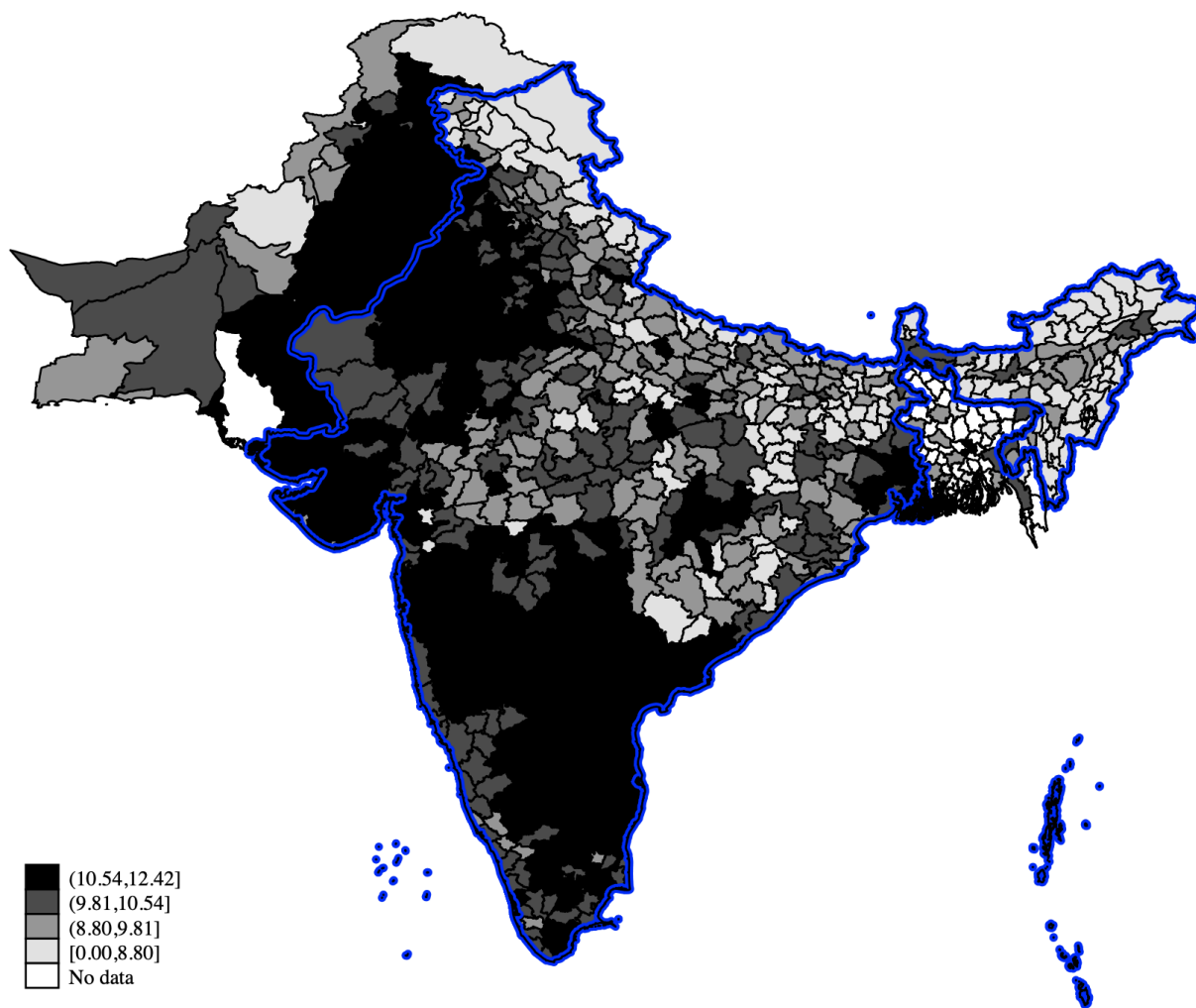


Panel B: Total Length of Tracks



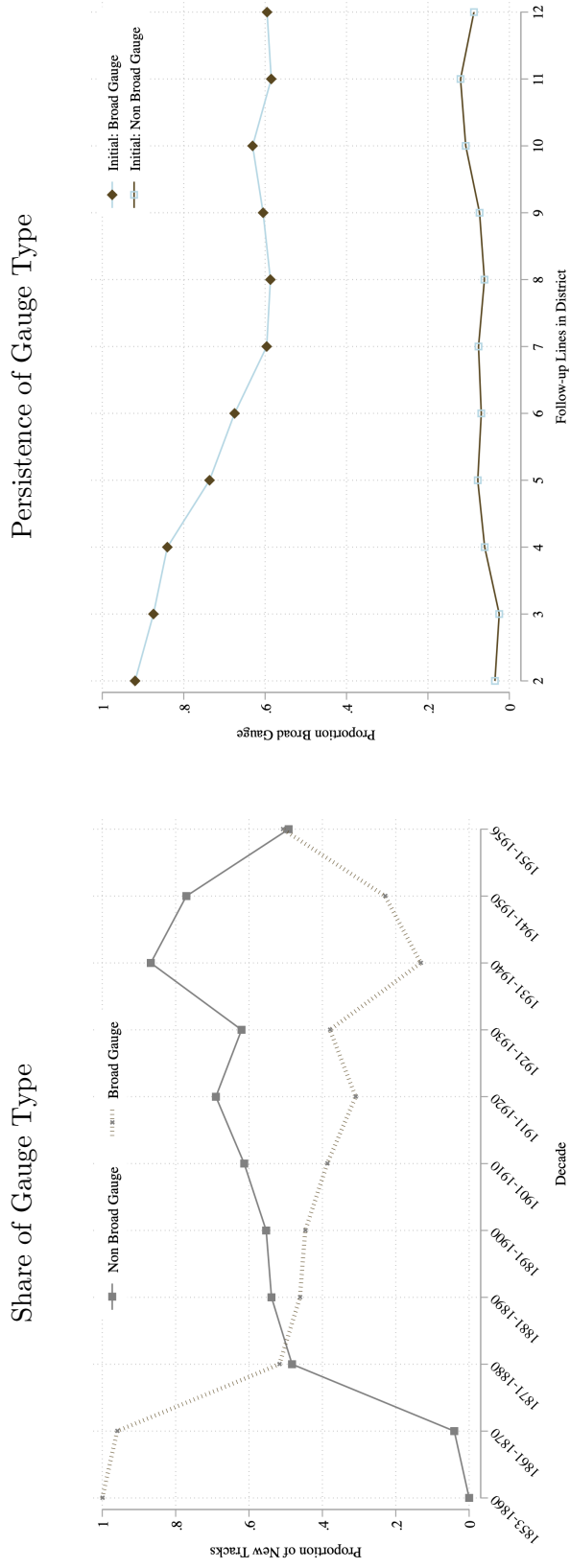
Notes: Sample Includes districts in present day Bangladesh, India and Pakistan where railways were constructed before 1956.

Figure 4: Geographical Dispersion of Night lights (Sum light). 2013. Each District of the Indian Subcontinent



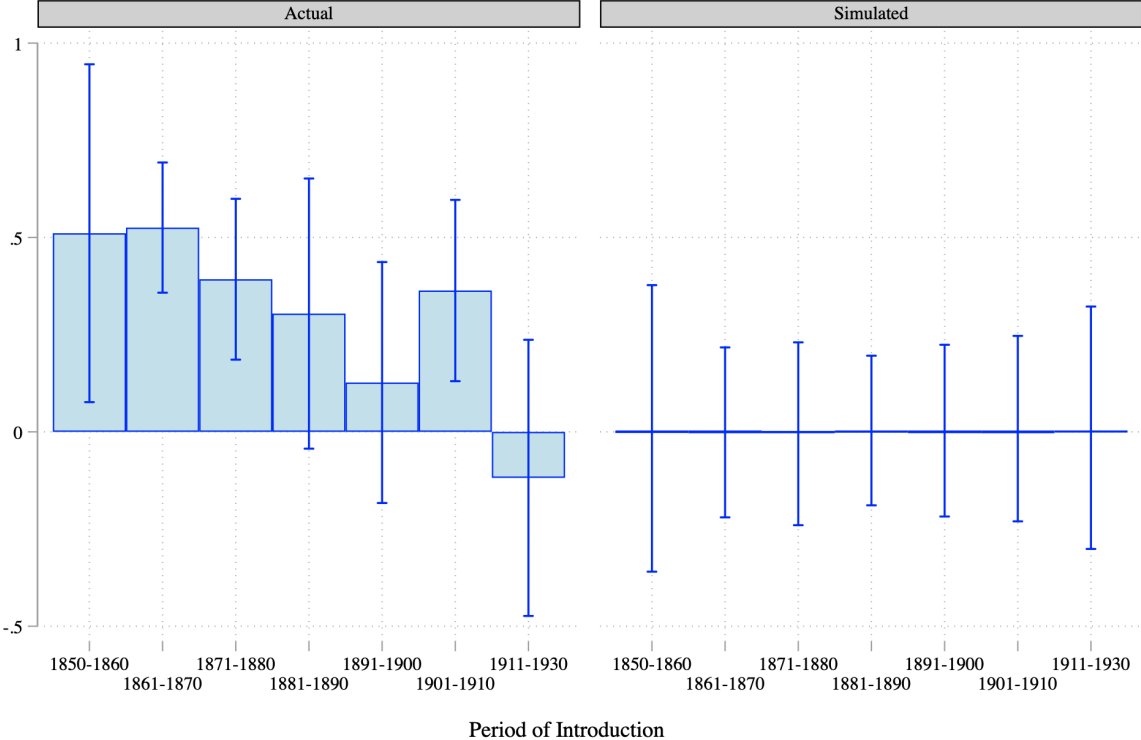
Notes: Outcome variable is $\log(1+sumlight)$. Sample includes all districts in present day Bangladesh, India and Pakistan. No data denotes night lights data missing for that district.

Figure 5: Railway Gauge across Districts



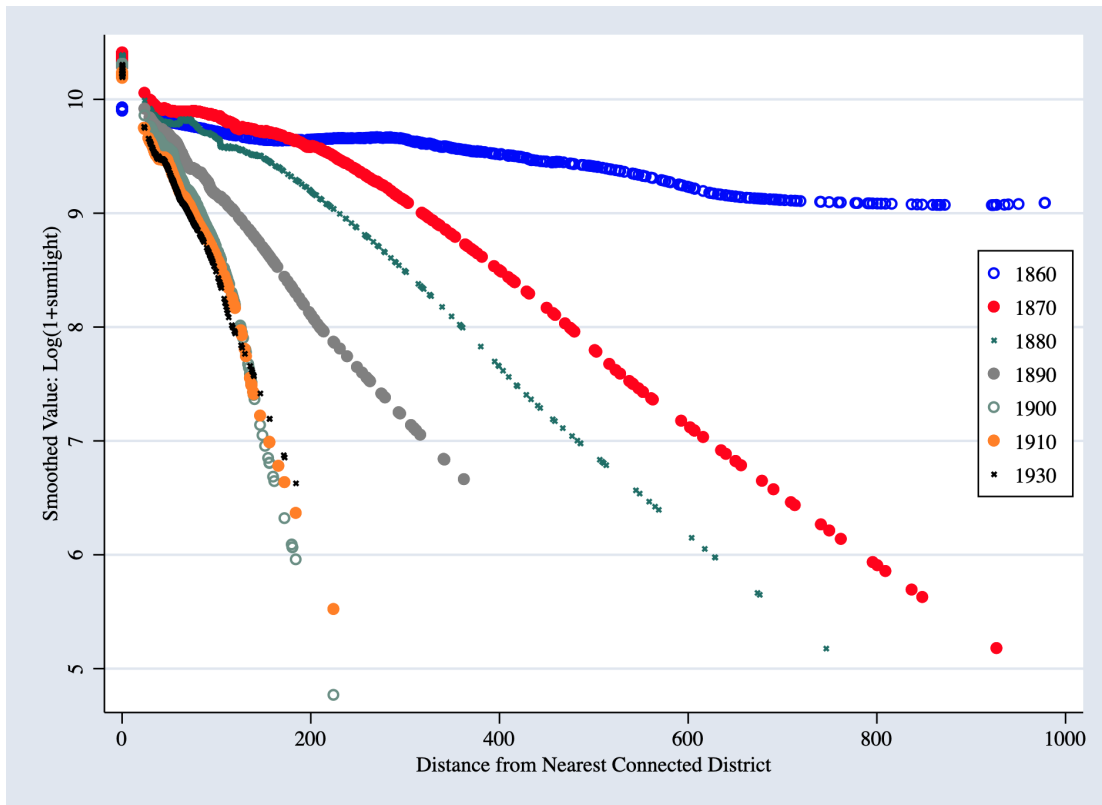
Notes: Sample Includes all districts in present day Bangladesh, India and Pakistan. Non-broad gauge include both narrow and metre gauge tracks. The left panel presents the average share of all new constructions that are broad and non-broad gauge. The right panel presents evidence of persistence of gauge type within the district (gauge type of follow-up railway constructions within the district).

Figure 6: Falsification Test: Actual and Simulated Effect



Notes: Dependent variable $\log(1 + sunlight_{ds})$ in district d in state s in 2013. Sample includes all districts in present day Bangladesh, India and Pakistan. Regression includes state fixed effects. Standard errors are clustered at the district level. 90% confidence interval presented. The simulated effect is the average from 10000 iterations. Bootstrapped 90% confidence intervals for the simulated effects obtained using Hall’s percentile method. The average simulated impacts are zero, suggesting that randomly allocated decade if connection does not have any long term effect on district prosperity. See Figure ?? for a distribution of the simulated effects.

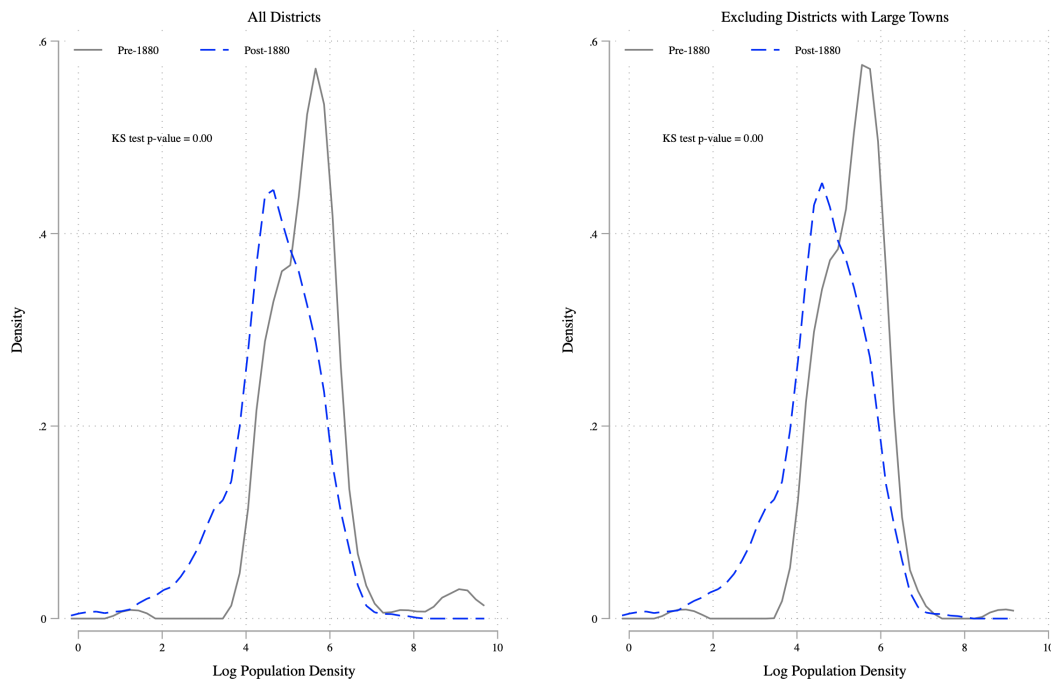
Figure 7: Distance from Nearest Connected District and Night time lights. Lowess Plots



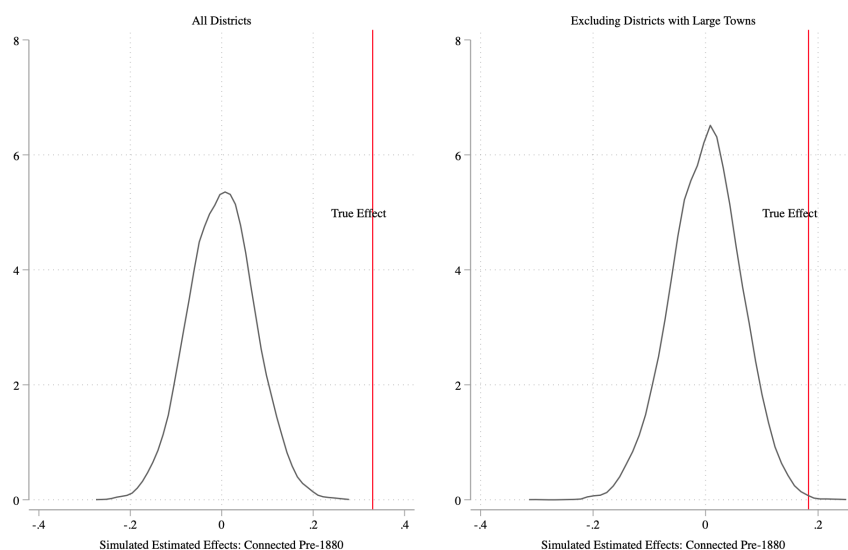
Notes: Lowess plots from a non-parametric regression of $\log(1 + \text{sunlight})$ on distance from the nearest connected district in each of the decades. Sample excludes districts more than 1000 kms from the nearest connected district.

Figure 8: Distribution of Population Density for Districts connected pre-1880 and post-1880

Panel A: Population Density in District 1901–2001



Panel B: Simulated and Actual Effect of pre-1880 introduction of railways in District



Notes: In **Panel A** the sample includes districts in India as of 2001. Districts without large towns refers to those districts that had a town in 1881 with population more than 100000. District population density for 1901 and beyond included in sample. **Panel B** presents the probability density function of the estimated coefficients of connected Pre-1880 obtained by randomly allocating year of connection and iterating 10,000 times. The vertical line indicates our true point estimates of pre-1880 connection dummy. Districts without large towns refers to those districts that had a town in 1881 with population more than 100000.

Table 1: Time of Introduction of Railways and Night Lights. Bangladesh, India and Pakistan

	Pooled 1992—2013 (1)	Pooled 1992—2013 (2)	1992 (3)	2013 (4)	Average 1992—2013 (5)	Coefficient of Variation (6)
Railway First 1850—1860	1.398*** (0.446)	0.893*** (0.201)	1.097*** (0.286)	0.752*** (0.233)	0.893*** (0.212)	-0.023** (0.009)
Railway First 1860—1870	1.311*** (0.333)	0.830*** (0.148)	1.049*** (0.223)	0.676*** (0.146)	0.830*** (0.156)	-0.026** (0.010)
Railway First 1870—1880	1.062*** (0.290)	0.741*** (0.147)	0.934*** (0.198)	0.547*** (0.174)	0.741*** (0.154)	-0.024** (0.010)
Railway First 1880—1890	1.015*** (0.295)	0.516** (0.207)	0.629** (0.279)	0.424* (0.229)	0.516** (0.218)	-0.020 (0.014)
Railway First 1890—1900	0.986*** (0.338)	0.432** (0.184)	0.662*** (0.230)	0.242 (0.207)	0.432** (0.194)	-0.017* (0.010)
Railway First 1900—1910	0.922*** (0.296)	0.523*** (0.186)	0.630** (0.255)	0.474*** (0.176)	0.523** (0.195)	-0.017* (0.009)
Railway First 1910—1930	0.941** (0.394)	0.086 (0.209)	0.168 (0.258)	0.026 (0.203)	0.086 (0.220)	-0.008 (0.009)
Latitude	-0.133*** (0.025)	-0.089*** (0.025)	-0.083** (0.033)	-0.064 (0.042)	-0.089*** (0.026)	0.001 (0.002)
Longitude	-0.046** (0.019)	-0.012 (0.042)	-0.022 (0.058)	0.018 (0.045)	-0.012 (0.045)	0.006 (0.005)
Constant	19.451*** (2.046)	15.491*** (1.926)	15.111*** (2.683)	13.976*** (3.448)	16.143*** (1.988)	-0.142 (0.108)
Sample Size	11,858	11,858	539	539	539	538
Number of Districts	539	539	539	539	539	538
Year Dummies	Yes	Yes	No	No	No	No
Country Dummies	Yes	No	No	No	No	No
State Fixed Effects	No	Yes	Yes	Yes	Yes	Yes

Notes: All districts in present day Bangladesh, Pakistan and India included. Dependent variable is $\log(1 + \text{sumlight}_{dst})$ in district d in state s in time t . OLS regression results presented in column 1. OLS regressions with state fixed effects presented in columns 2–6. Regression in column 1 include a set of country dummies. Regressions include state fixed effects. State clustered standard errors presented in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are presented in [A1](#).

Table 2: Time of Introduction of Railways and Night Lights. India. 2013

	(1)	(2)	(3)	(4)	(5)
Railway First 1850—1860	0.752*** (0.236)	0.602*** (0.219)	0.532** (0.226)	0.743*** (0.250)	0.807*** (0.171)
Railway First 1860—1870	0.640*** (0.144)	0.503*** (0.114)	0.403*** (0.142)	0.528*** (0.152)	0.279** (0.118)
Railway First 1870—1880	0.593*** (0.181)	0.405*** (0.141)	0.352** (0.149)	0.406** (0.157)	0.307** (0.143)
Railway First 1880—1890	0.387 (0.239)	0.241 (0.212)	0.254 (0.187)	0.214 (0.169)	0.049 (0.143)
Railway First 1890—1900	0.211 (0.201)	0.082 (0.163)	0.108 (0.166)	0.085 (0.148)	-0.272 (0.190)
Railway First 1900—1910	0.468** (0.186)	0.311** (0.129)	0.241* (0.142)	0.214 (0.164)	0.083 (0.134)
Railway First 1910—1930	0.137 (0.194)	0.110 (0.148)	0.069 (0.157)	0.428* (0.207)	0.321 (0.294)
Latitude	-0.066 (0.043)	-0.045 (0.044)	-0.053 (0.039)	-0.050 (0.040)	-0.077*** (0.024)
Longitude	0.029 (0.045)	-0.006 (0.043)	0.018 (0.043)	0.024 (0.035)	0.045 (0.043)
Proportion SC in District		0.176 (1.192)	0.323 (1.177)	-0.719 (0.763)	0.024 (0.809)
Proportion ST in District		-2.143*** (0.468)	-1.754*** (0.430)	-1.854*** (0.518)	-1.071 (0.805)
Proportion literate in District			2.843** (1.074)	1.757 (1.448)	3.560** (1.590)
Percentage of District with mean elevation 500 m or more				0.229 (0.176)	0.162 (0.280)
River Km in District				-0.010 (0.084)	-0.174 (0.235)
Direct British Rule					-0.171 (0.180)
Log Real Income 1930					-0.032 (0.103)
Log are under Cultivation 1930					0.308*** (0.096)
Constant	13.890*** (3.579)	13.409*** (3.102)	11.713*** (2.572)	12.426*** (2.757)	9.408*** (2.714)
Number of Districts	491	491	491	294	166

Notes: Sample restricted to present day districts in India. Dependent variable is $\log(1 + \text{sumlight}_{ds})$ in district d in state s in 2013. OLS regression results presented. Regressions include state fixed effects. State clustered standard errors presented in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are presented in [A2](#).

Table 3: Initial Gauge Type and Night Lights in the District in 2013

	All (1)	India (2)	India (3)	India (4)
Broad Gauge	0.501*** (0.152)	0.477*** (0.154)	0.318* (0.166)	0.326* (0.179)
Non-Broad Gauge	0.395** (0.196)	0.397* (0.197)	0.251 (0.173)	0.210 (0.204)
Latitude	-0.086* (0.043)	-0.071 (0.044)	-0.058 (0.041)	-0.098** (0.034)
Longitude	0.027 (0.045)	0.032 (0.046)	0.033 (0.039)	0.027 (0.054)
Proportion SC in District			-0.793 (0.851)	0.007 (0.864)
Proportion ST in District			-1.745*** (0.511)	-1.155 (0.806)
Proportion literate in district			2.337 (1.434)	3.554** (1.635)
Percent of District with mean elevation 500 m or more			0.203 (0.176)	0.258 (0.251)
River KM in district			0.011 (0.077)	-0.132 (0.211)
Direct British Rule				-0.171 (0.164)
Log Real Income 1930				-0.066 (0.128)
Log Area under Cultivation 1930				0.286** (0.116)
Constant	15.556*** (3.632)	14.196*** (3.735)	12.558*** (2.825)	12.244** (4.207)
Difference by Gauge Type	0.106 (0.127)	0.080 (0.129)	0.067 (0.096)	0.116 (0.149)
Number of Districts	539	491	294	166

Notes: Dependent variable is $\log(1 + \text{sumlight}_{ds})$ in district d in state s in 2013. OLS regression results presented. In Column 1 sample includes all districts in present day India, Bangladesh and Pakistan; in Columns 2–4, the sample includes all districts in present day India. *Broad Gauge* dummy denotes the initial connection in the district was broad gauge; *Non-Broad Gauge* dummy denotes the initial connection in the district was non-broad gauge (metre or narrow gauge). Reference category is that the district was not connected by 1930. Regressions include state fixed effects. State clustered standard errors presented in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are available on request.

Table 4: Selection: OLS and IV Regression Results

	All	All: IV	India	India: IV	India	India: IV
	(1)	(2)	(3)	(4)	(5)	(6)
Railway First: Pre-1880	0.385*** (0.062)	0.878** (0.417)	0.364*** (0.065)	0.972** (0.452)	0.335*** (0.062)	0.851** (0.396)
Latitude	-0.080* (0.040)	-0.073* (0.039)	-0.065 (0.041)	-0.061 (0.040)	-0.054 (0.037)	-0.039 (0.041)
Longitude	0.034 (0.043)	0.024 (0.045)	0.040 (0.045)	0.030 (0.047)	0.021 (0.037)	-0.007 (0.038)
Proportion of SC in District					-0.698 (0.826)	-0.946 (0.691)
Proportion of ST in District					-1.831*** (0.530)	-1.909*** (0.441)
Proportion literate in District					2.006 (1.386)	1.193 (1.619)
Proportion of district with mean elevation 500 m or more					0.202 (0.186)	0.247 (0.190)
River Km in district					0.004 (0.080)	-0.017 (0.082)
Constant	15.075*** (3.340)	13.135*** (3.677)	13.787*** (3.437)	11.937*** (3.788)	12.773*** (2.657)	13.755*** (2.762)
<i>First Stage Regression: Dependent variable District connected Pre-1880</i>						
Minimum Distance from Line of Shortest Distance		-0.121*** (0.026)		-0.125*** (0.027)		-0.124*** (0.033)
First Stage F		22.71***		20.19***		13.98***
Number of Districts	539	539	491	491	294	294

Notes: Dependent variable is $\log(1 + \text{sumlight}_{ds})$ in district d in state s in 2013. OLS regression results presented in Columns 1 and 2, sample includes all districts in present day India, Bangladesh and Pakistan. In Columns 3–6, the sample includes all districts in Regressions include state fixed effects. State clustered standard errors presented in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are available on request.

Table 5: Connectedness, Distance from Connected Districts and Night lights. 2013

	Connected District			Distance to Nearest Connected District		
	All (1)	India (2)	India (3)	All (4)	India (5)	India (6)
Connected by 1860	0.426*** (0.143)	0.480*** (0.149)	0.480*** (0.149)	-0.062 (0.053)	-0.119** (0.052)	-0.073* (0.040)
Connected by 1870	0.508*** (0.059)	0.478*** (0.072)	0.478*** (0.072)	-0.210** (0.082)	-0.216*** (0.059)	-0.219*** (0.047)
Connected by 1880	0.458*** (0.045)	0.450*** (0.061)	0.450*** (0.061)	-0.212** (0.098)	-0.231*** (0.061)	-0.303*** (0.060)
Connected by 1890	0.416*** (0.089)	0.352*** (0.058)	0.352*** (0.058)	-0.366*** (0.093)	-0.276*** (0.068)	-0.328*** (0.060)
Connected by 1900	0.426*** (0.093)	0.315*** (0.054)	0.315*** (0.054)	-0.674*** (0.146)	-0.338*** (0.092)	-0.390*** (0.095)
Connected by 1910	0.401*** (0.092)	0.286*** (0.065)	0.286*** (0.065)	-0.650*** (0.162)	-0.317*** (0.088)	-0.333** (0.131)
Connected by 1930	0.416*** (0.096)	0.305*** (0.083)	0.305*** (0.083)	-0.670*** (0.175)	-0.312*** (0.088)	-0.354** (0.163)
Number of Districts	568	310	310	568	520	310
Additional Controls	No	No	Yes	No	No	Yes

Notes: OLS regression results presented. Dependent variable $\log(1 + \text{sumlight}_{d,s})$ in district d in state s in 2013. Sample in regression results presented in columns 1 and 4 include all districts in present day Bangladesh, India and Pakistan. Sample in regression results presented in columns 2, 3, 5 and 6 include all districts in present day India. Each cell presents the results from a separate regression. Columns 1–3 presents the regression results corresponding to equation (3), while Columns 4–6 presents the regression results corresponding to equation (2). Regressions in columns 3 and 6 include additional controls (as in Table 2) Regressions include state fixed effects. Standard errors clustered at the state level are presented in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are available on request.

Table 6: Track length and Night lights. 2013

	All (1)	India (2)	India (3)
Length of Track in district as of 1860	0.021*** (0.007)	0.021*** (0.007)	0.022*** (0.005)
Length of Track in district as of 1870	0.017*** (0.006)	0.016*** (0.006)	0.020*** (0.007)
Length of Track in district as of 1880	0.014** (0.006)	0.018*** (0.005)	0.016*** (0.005)
Length of Track in district as of 1890	0.010* (0.006)	0.013** (0.005)	0.011*** (0.003)
Length of Track in district as of 1900	0.008* (0.004)	0.010** (0.004)	0.007** (0.003)
Length of Track in district as of 1910	0.009** (0.004)	0.010** (0.004)	0.008** (0.003)
Length of Track in district as of 1930	0.009*** (0.003)	0.010*** (0.003)	0.008*** (0.002)
Number of Districts	408	371	252
Additional Controls	No	No	Yes

Notes: OLS regression results presented. Dependent variable $\log(1 + \text{sumlight}_{ds})$ in district d in state s in 2013. Sample in regression results presented in column 1 includes all districts in present day Bangladesh, India and Pakistan. Sample in regression results presented in columns 2 and 3 include all districts in present day India. Each cell presents the results from a separate regression. Regressions in column 3 include additional controls (as in Table 2) Regressions include state fixed effects. Standard errors are clustered at the state level are presented in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are available on request.

Table 7: Time of Introduction of Railways and District Prosperity. Mean Lights

	All (1)	India (2)	India (3)	India (4)	India (5)	India (6)
Railway First 1850–1860	0.726*** (0.167)	0.691*** (0.167)	0.589*** (0.160)	0.509*** (0.128)	0.517*** (0.147)	0.548*** (0.169)
Railway First 1860–1870	0.464*** (0.097)	0.403*** (0.096)	0.337*** (0.076)	0.223** (0.089)	0.223* (0.111)	0.224** (0.081)
Railway First 1870–1880	0.333*** (0.095)	0.307*** (0.101)	0.213** (0.084)	0.153* (0.085)	0.162* (0.084)	0.157 (0.109)
Railway First 1880–1890	0.198 (0.177)	0.136 (0.181)	0.058 (0.166)	0.073 (0.129)	0.070 (0.102)	0.026 (0.121)
Railway First 1890–1900	0.131 (0.128)	0.096 (0.131)	0.040 (0.126)	0.069 (0.126)	-0.032 (0.108)	-0.188 (0.227)
Railway First 1900–1910	0.248** (0.117)	0.232* (0.121)	0.141 (0.093)	0.062 (0.102)	0.047 (0.120)	-0.000 (0.116)
Railway First 1910–1930	0.087 (0.111)	0.076 (0.116)	0.088 (0.107)	0.042 (0.112)	0.030 (0.163)	0.058 (0.149)
Latitude	-0.009 (0.039)	-0.011 (0.039)	0.005 (0.042)	-0.005 (0.034)	-0.002 (0.031)	-0.012 (0.024)
Longitude	0.024 (0.034)	0.027 (0.035)	0.007 (0.032)	0.034 (0.034)	0.035 (0.036)	0.072** (0.024)
Proportion SC in District			-1.076 (0.682)	-0.909 (0.670)	-1.721** (0.670)	-1.814*** (0.556)
Proportion ST in District			-1.493*** (0.228)	-1.052*** (0.250)	-1.021** (0.387)	-1.331*** (0.374)
Proportion literate in district				3.216*** (0.739)	3.582*** (0.988)	4.028*** (1.090)
Percent of District with mean elevation 500 m or more					-0.154 (0.220)	-0.173 (0.281)
River KM in district					-0.123 (0.076)	-0.229** (0.097)
Direct British Rule						-0.275* (0.130)
Log Real Income 1930						0.038 (0.078)
Log Area under Cultivation 1930						-0.058 (0.108)
Constant	1.591 (2.959)	1.738 (3.034)	1.395 (2.983)	-0.523 (2.320)	-0.647 (2.320)	-0.477 (1.633)
Number of Districts	539	491	491	491	294	166

Notes: Column 1, sample restricted to all districts in present day India, Bangladesh and Pakistan. In columns 2–5 sample restricted to present day districts in India. Dependent variable is $\log(1 + \text{meanlight}_{ds})$ in district d in state s in 2013. Regressions in columns 1 and 2 include a set of country dummies. State clustered standard errors are presented in parentheses. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are available on request.

Table 8: Time of Introduction of Railways and District Prosperity. Alternative Measures

	(1)	Bank Deposits			Rural Poverty			Urban Poverty		
	(1)	(2)	(3)	(4)	(5)	(7)	(8)	(9)		
Railway First 1850–1860	1.848*** (0.556)	1.560*** (0.348)	1.534*** (0.440)	-2.498 (3.971)	-7.535 (6.787)	-9.634*** (2.921)	-7.457 (7.753)	-6.990 (4.207)	-2.360 (9.144)	
Railway First 1860–1870	1.001*** (0.215)	0.820*** (0.180)	0.780*** (0.316)	-5.189* (2.848)	-0.327 (3.323)	-7.623*** (3.241)	3.160 (3.629)	-7.040 (4.052)	1.755 (6.002)	
Railway First 1870–1880	1.074*** (0.148)	0.978*** (0.163)	1.117*** (0.310)	-7.289** (3.136)	-0.717 (3.551)	-10.875*** (3.249)	-0.395 (3.930)	-6.851 (4.904)	-0.258 (7.008)	
Railway First 1880–1890	0.640*** (0.184)	0.531*** (0.105)	0.553* (0.266)	-3.249 (2.938)	2.793 (3.828)	-8.613*** (3.180)	5.231 (3.604)	-4.784 (4.112)	8.574 (7.067)	
Railway First 1890–1900	0.551*** (0.202)	0.428** (0.200)	0.323 (0.360)	-5.732 (3.329)	0.491 (3.195)	-6.904 (4.204)	4.659 (3.408)	-5.412 (5.916)	10.529 (6.878)	
Railway First 1900–1910	0.508*** (0.169)	0.149 (0.211)	0.346 (0.261)	-2.103 (3.735)	1.331 (3.652)	-4.537 (5.032)	2.900 (4.345)	1.004 (5.172)	2.103 (6.060)	
Railway First 1910–1930	0.033 (0.164)	0.252 (0.235)	0.605** (0.279)	0.548 (3.679)	2.748 (3.786)	-0.824 (4.972)	5.084 (3.612)	2.702 (6.243)	11.614** (4.919)	
Latitude	-0.066 (0.045)	-0.048 (0.032)	-0.043 (0.053)	1.604 (1.372)	0.015 (0.809)	1.346 (1.088)	-0.147 (0.693)	2.236* (1.159)	-0.581 (1.181)	
Longitude	-0.030 (0.049)	0.012 (0.039)	0.037 (0.066)	-0.383 (1.247)	-0.556 (1.294)	-0.348 (1.233)	0.147 (1.048)	0.077 (1.324)	-0.200 (0.532)	
Proportion SC in District		-3.971*** (0.487)	-3.271*** (0.652)							
Proportion ST in District		-1.213* (0.593)	-1.675* (0.843)							
Proportion literate in District		5.768*** (0.876)	6.087*** (1.104)							
Percent of District with mean elevation 500 m or more		-0.144 (0.315)	-0.060 (0.465)							
River KM in district		0.066 (0.146)	0.171 (0.281)							
Direct British Rule			-0.174 (0.348)							
Log Real Income 1930			0.274 (0.198)							
Log Area under Cultivation 1930			-0.320 (0.186)							
Constant	14.670*** (4.052)	9.895*** (2.704)	8.537 (5.737)	-87.780 (106.939)	43.922 (57.325)	-67.131 (79.785)	38.065 (58.154)	-107.165 (101.858)	99.875 (117.985)	
Number of Districts	629	334	166	506	506	332	329	164	165	

Continued ...

Time of Introduction of Railways and District Prosperity. Alternative Measures

(1)	Bank Deposits (2)	Rural Poverty (5)	Urban Poverty (9)
	(3)	(4)	(7)
		(8)	(9)

Notes: OLS regression results presented. The table presents the estimated β 's from equation (1). Dependent variable in column 1 log (total bank deposits) in district d , state s in all banks (state bank and associates, other nationalized banks, foreign banks, regional rural banks, old private sector banks, new private sector banks). Dependent variable in columns 2 and 3 are percentage of rural and urban population respectively categorized as poor. Data on rural and urban poverty rates by district obtained from Chaudhuri and Gupta (2009). Regressions include state fixed effects. State clustered standard errors are presented in parentheses. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are available on request.

Table 9: Timing of Introduction of Railways and Proportion of Adult Males and Females with no education (IHDS2)

	Male (1)	Female (2)	Male (3)	Female (4)	Male (5)	Female (6)
Railway First 1850–1860	-0.062* (0.036)	-0.119** (0.049)	-0.070** (0.025)	-0.131** (0.048)	-0.031 (0.045)	-0.089 (0.064)
Railway First 1860–1870	-0.048 (0.034)	-0.081* (0.047)	-0.056* (0.028)	-0.102** (0.047)	-0.053 (0.045)	-0.112* (0.053)
Railway First 1870–1880	-0.079*** (0.027)	-0.108*** (0.035)	-0.088*** (0.024)	-0.128*** (0.037)	-0.061** (0.028)	-0.158*** (0.039)
Railway First 1880–1890	-0.046 (0.030)	-0.080* (0.039)	-0.050** (0.024)	-0.104** (0.041)	-0.048 (0.039)	-0.127* (0.061)
Railway First 1890–1900	-0.079** (0.036)	-0.117*** (0.041)	-0.073** (0.027)	-0.134*** (0.036)	-0.065 (0.060)	-0.147* (0.080)
Railway First 1900–1910	-0.026 (0.026)	-0.035 (0.029)	-0.046* (0.022)	-0.065 (0.040)	-0.039 (0.041)	-0.078 (0.058)
Railway First 1910–1930	-0.042 (0.037)	-0.053 (0.039)	-0.048** (0.020)	-0.063* (0.034)	-0.032 (0.039)	-0.077** (0.029)
Latitude	-0.002 (0.006)	0.002 (0.008)	-0.003 (0.004)	-0.003 (0.008)	0.003 (0.006)	0.008 (0.010)
Longitude	0.004 (0.006)	0.016** (0.007)	0.012** (0.004)	0.020*** (0.006)	0.019* (0.010)	0.025* (0.013)
Proportion SC in District			0.089 (0.203)	0.337 (0.309)	-0.078 (0.205)	-0.043 (0.183)
Proportion ST in District			0.100 (0.073)	0.190 (0.115)	-0.050 (0.195)	0.189 (0.238)
Proportion literate in District			0.053* (0.029)	0.048 (0.039)	0.074* (0.041)	0.136** (0.054)
Percent of District with mean elevation 500 m or more			0.023 (0.015)	0.036** (0.016)	0.049 (0.048)	0.049 (0.068)
River KM in district					-0.015 (0.039)	-0.053 (0.035)
Direct British Rule					-0.029 (0.039)	0.001 (0.043)
Log Real Income 1930					0.039 (0.040)	0.023 (0.042)
Constant	0.327 (0.538)	-0.073 (0.658)	0.118 (0.407)	0.180 (0.683)	-0.509 (0.550)	-1.014 (0.895)
Number of Districts	358	358	240	240	112	112

Notes: Dependent variable: proportion of adult males and females aged 15 and higher who report no schooling. Sample restricted to present day districts in India. OLS regression results using IHDS2 data presented. Regressions include state fixed effects. Date clustered standard errors are presented in parentheses. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are available on request.

Table 10: Timing of Introduction of Railways and Proportion of Adult Males and Females categorized as Underweight (IHDS2)

	Male (1)	Female (2)	Male (3)	Female (4)	Male (5)	Female (6)
Railway First 1850–1860	-0.053 (0.036)	-0.071** (0.027)	-0.049 (0.030)	-0.075*** (0.022)	-0.060* (0.031)	-0.105*** (0.029)
Railway First 1860–1870	-0.032 (0.021)	-0.049** (0.020)	-0.035 (0.020)	-0.056*** (0.018)	-0.046 (0.034)	-0.077** (0.035)
Railway First 1870–1880	-0.035** (0.017)	-0.045*** (0.014)	-0.039** (0.017)	-0.055*** (0.014)	-0.063 (0.036)	-0.083*** (0.025)
Railway First 1880–1890	-0.036 (0.022)	-0.049** (0.024)	-0.031 (0.019)	-0.049** (0.019)	-0.046 (0.027)	-0.069** (0.024)
Railway First 1890–1900	-0.048*** (0.015)	-0.038* (0.019)	-0.040** (0.014)	-0.052*** (0.018)	-0.069 (0.040)	-0.084** (0.034)
Railway First 1900–1910	-0.021 (0.012)	-0.021* (0.011)	-0.013 (0.016)	-0.009 (0.017)	-0.034 (0.024)	-0.021 (0.029)
Railway First 1910–1930	0.011 (0.020)	-0.007 (0.028)	0.020 (0.016)	-0.001 (0.031)	0.036* (0.020)	-0.050 (0.034)
Latitude	0.008* (0.004)	0.009*** (0.003)	0.006* (0.003)	0.004 (0.003)	0.010* (0.005)	0.008** (0.003)
Longitude	0.003 (0.003)	0.007* (0.004)	0.004 (0.003)	0.007** (0.003)	0.005 (0.006)	0.002 (0.005)
Proportion SC in District			0.110 (0.114)	0.164 (0.123)	0.131 (0.173)	0.156 (0.216)
Proportion ST in District			0.110* (0.053)	0.150*** (0.045)	0.199* (0.113)	0.321** (0.133)
Proportion literate in District			-0.042** (0.015)	-0.022 (0.015)	-0.022 (0.026)	0.012 (0.038)
Percent of District with mean elevation 500 m or more			0.018 (0.011)	0.020* (0.010)	0.043 (0.042)	0.043 (0.043)
River KM in district					0.008 (0.019)	0.030* (0.017)
Direct British Rule					-0.014 (0.014)	-0.013 (0.020)
Log Real Income 1930					0.017 (0.015)	0.012 (0.021)
Constant	-0.569 (0.376)	-0.650** (0.255)	-0.434 (0.290)	-0.276 (0.231)	-0.811** (0.370)	-0.494* (0.236)
Number of Districts	358	358	240	240	112	112

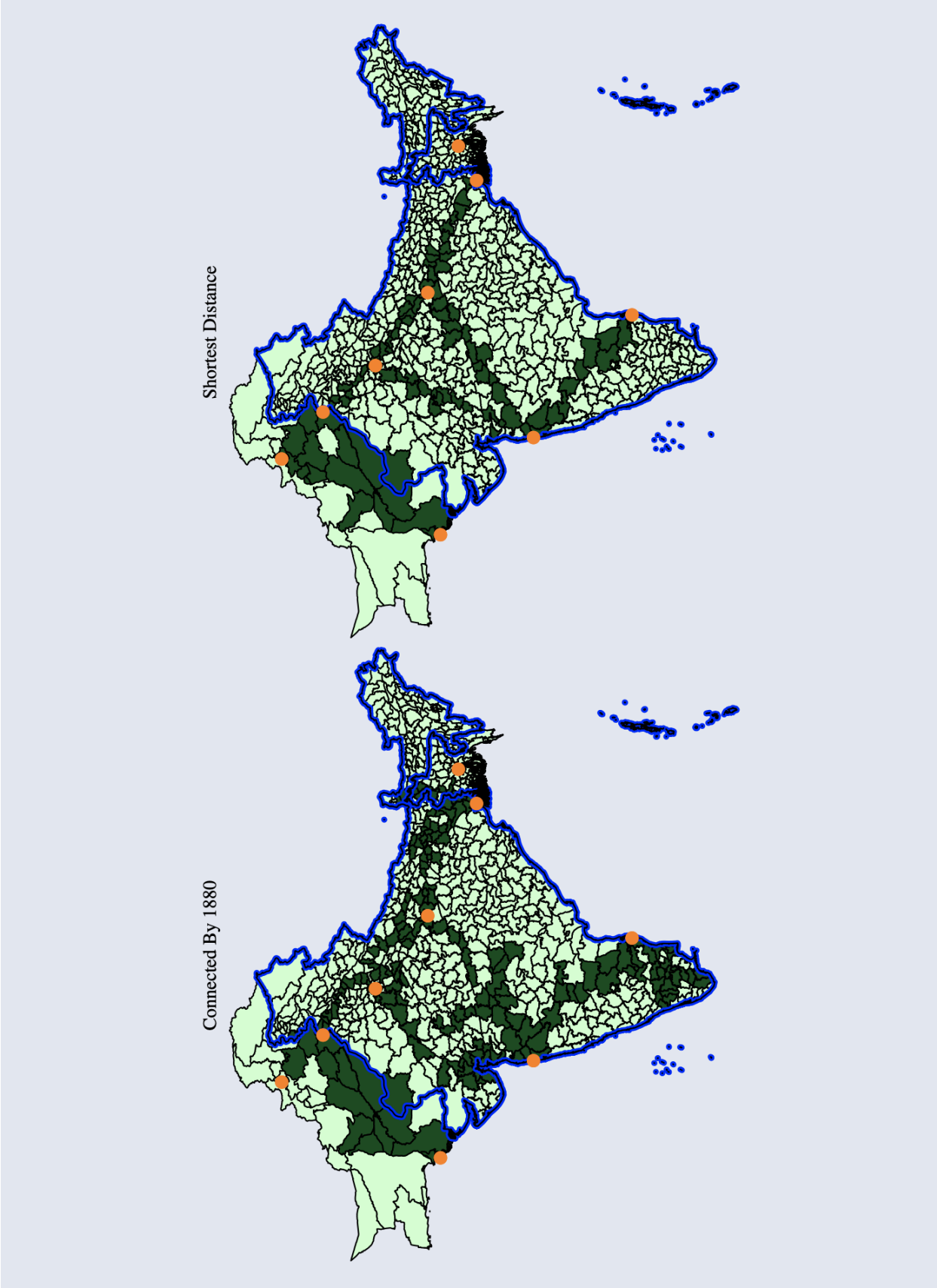
Notes: Dependent variable: proportion of adult males and females aged 15 and higher who are underweight ($BMI \leq 18.5$). Sample restricted to present day districts in India. OLS regression results using IHDS2 data presented. Regressions include state fixed effects. Date clustered standard errors are presented in parentheses. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are available on request.

Table 11: Connection Before 1880 and Population Density

	All Districts (1)	Districts without Large Towns (2)
Pre-1880	0.330*** (0.069)	0.183*** (0.043)
Latitude	0.021 (0.049)	0.033 (0.041)
Longitude	-0.019 (0.042)	-0.010 (0.040)
Constant	2.954 (4.155)	1.733 (3.257)
Number of Observations	6,012	5,814

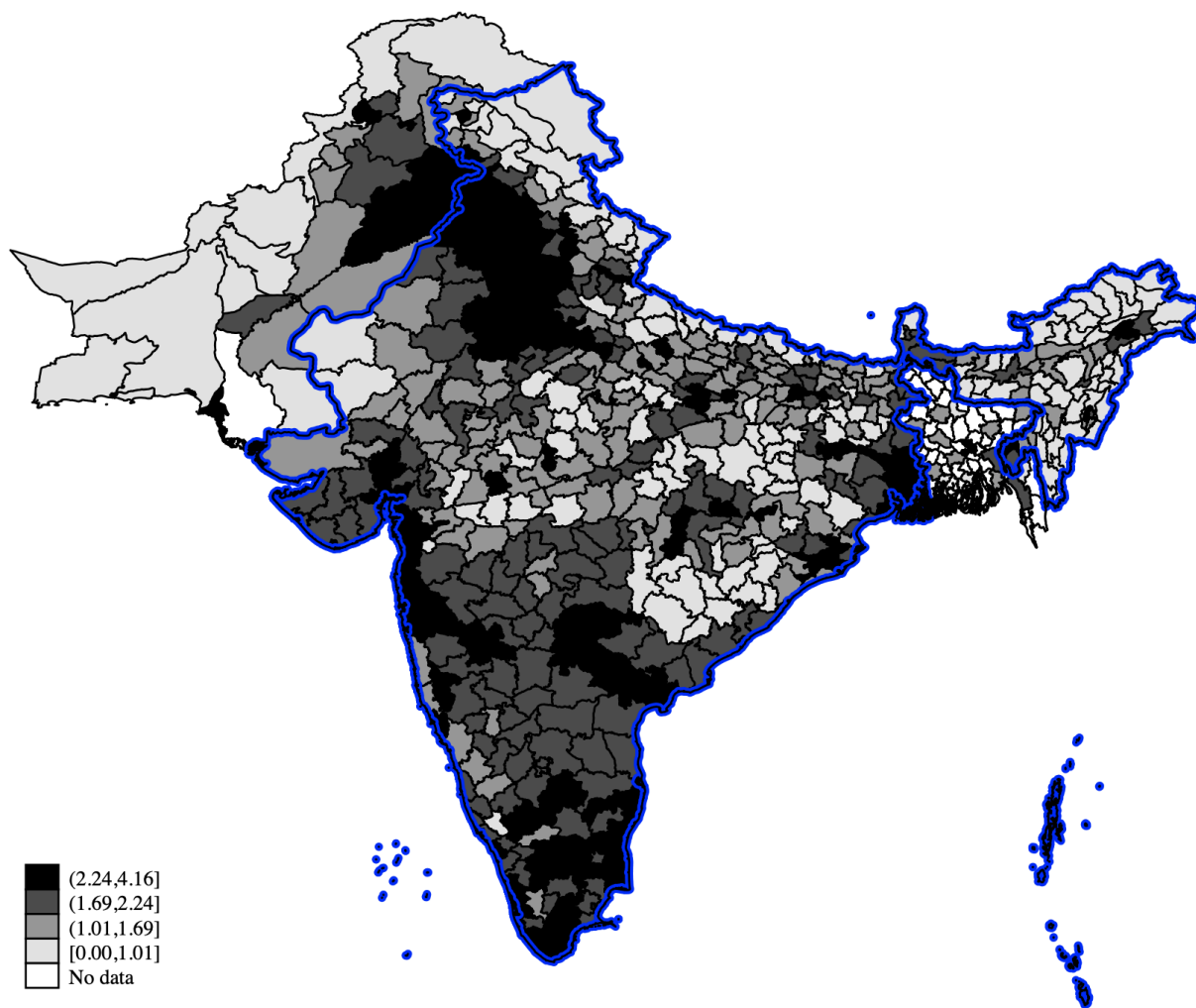
Notes: Dependent variable: $\log(\text{Population Density})$ in district d in state s in decade T . OLS regression results presented. Sample includes districts in India as of 2001. OLS regression results presented. Districts without large towns refers to those districts that had a town in 1881 with population more than 100000. Regressions include a set of decade dummies and a set of state fixed effects. Standard errors clustered at the state level are presented in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. The results using spatially corrected standard errors are available on request.

Figure A1: Actual and Shortest Distance Connections



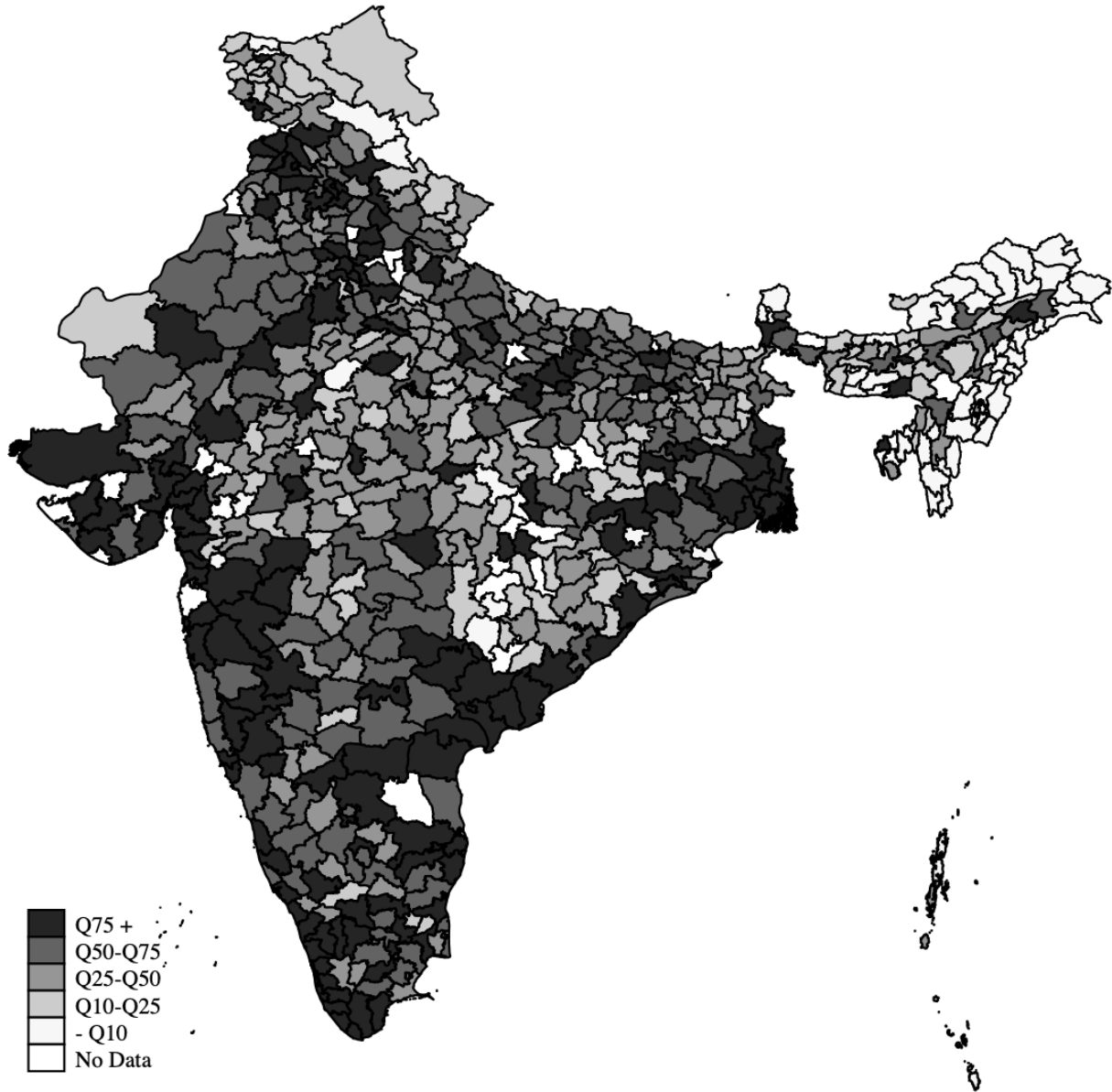
Notes: The darker shaded districts in the left panel denote the present day districts of India, Bangladesh and Pakistan that were connected as of 1880. The darker shaded districts in the right panel denote the districts that would be connected if a hypothetical shortest distance route between the key cities was adopted.

Figure A2: Geographical Dispersion of Night lights (Mean light). 2013. Each District of the Indian Subcontinent



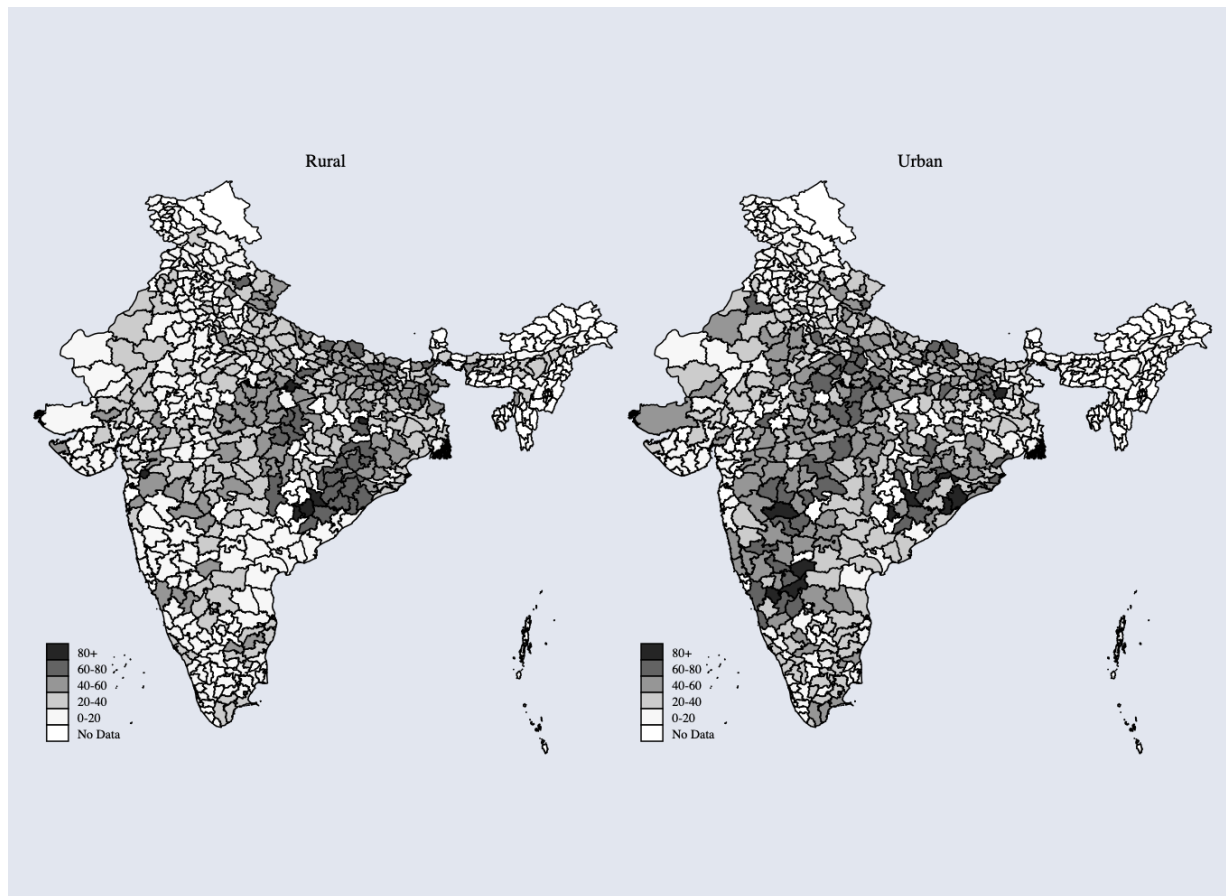
Notes: Outcome variable is $\log(1 + \text{meanlight})$. Sample includes all districts in present day India, Bangladesh and Pakistan. No data denotes night lights data missing for that district.

Figure A3: Aggregate Bank Deposits by District, 2013



Notes: Sample includes all districts in present day India. Aggregate Bank deposits include deposits in state bank and associates, other nationalized banks, foreign banks, regional rural banks, old private sector banks and new private sector banks.

Figure A4: Rural and Urban Poverty Rate by District, 2004–05



Notes: Sample includes all districts in present day India. Percentage of rural and urban population respectively categorized as poor presented. Data on rural and urban poverty rates by district obtained from [Chaudhuri and Gupta \(2009\)](#).

Table A1: Time of Introduction of Railways and Night Lights. Bangladesh, India and Pakistan. Alternative Clustering of Standard Errors

	Pooled 1992–2013 (1)	Pooled 1992–2013 (2)	1992 (3)	2013 (4)	Average 1992–2013 (5)	Coefficient of Variation (6)
Railway First 1850–1860	1.398*** (0.446) {0.309} [0.425]	0.893*** (0.201) {0.231} [0.197]	1.097*** (0.286) {0.280} [0.157]	0.752*** (0.233) {0.232} [0.180]	0.893*** (0.212) {0.231} [0.197]	-0.023** (0.009) {0.008} [0.006]
Railway First 1860–1870	1.311*** (0.333) {0.194} [0.228]	0.830*** (0.148) {0.137} [0.129]	1.049*** (0.223) {0.165} [0.150]	0.676*** (0.146) {0.142} [0.115]	0.830*** (0.156) {0.137} [0.129]	-0.026** (0.010) {0.008} [0.008]
Railway First 1870–1880	1.062*** (0.290) {0.204} [0.231]	0.741*** (0.147) {0.163} [0.197]	0.934*** (0.198) {0.177} [0.202]	0.547*** (0.174) {0.178} [0.201]	0.741*** (0.154) {0.163} [0.197]	-0.024** (0.010) {0.007} [0.008]
Railway First 1880–1890	1.015*** (0.295) {0.180} [0.198]	0.516** (0.207) {0.152} [0.177]	0.629** (0.279) {0.188} [0.207]	0.424* (0.229) {0.143} [0.168]	0.516** (0.218) {0.152} [0.177]	-0.020 (0.014) {0.011} [0.011]
Railway First 1890–1900	0.986*** (0.338) {0.209} [0.226]	0.432** (0.184) {0.162} [0.172]	0.662*** (0.230) {0.184} [0.210]	0.242 (0.207) {0.174} [0.171]	0.432** (0.194) {0.162} [0.172]	-0.017* (0.010) {0.007} [0.008]
Railway First 1900–1910	0.922*** (0.296) {0.178} [0.221]	0.523*** (0.186) {0.157} [0.173]	0.630** (0.255) {0.189} [0.191]	0.474*** (0.176) {0.157} [0.166]	0.523** (0.195) {0.157} [0.173]	-0.017* (0.009) {0.006} [0.006]
Railway First 1910–1930	0.941** (0.394) {0.298} [0.308]	0.086 (0.209) {0.181} [0.183]	0.168 (0.258) {0.225} [0.223]	0.026 (0.203) {0.159} [0.176]	0.086 (0.220) {0.181} [0.183]	-0.008 (0.009) {0.008} [0.009]
Constant	19.451*** (2.046)	15.491*** (1.926)	15.111*** (2.683)	13.976*** (3.448)	16.143*** (1.988)	-0.142 (0.108)
Sample Size	11,858	11,858	539	539	539	538
Number of Districts	539	539	539	539	539	538
Year Dummies	Yes	Yes	No	No	No	No
State Fixed Effects	No	Yes	Yes	Yes	Yes	Yes

Notes: All districts in present day Bangladesh, Pakistan and India included. Dependent variable is $\log(1 + \text{sumlight}_{dst})$ in district d in state s in time t . OLS regression results presented in column 1. OLS regressions with state fixed effects presented in columns 2–6. All regressions control for latitude and longitude. Regression in column 1 include a set of country dummies. Regression results presented in columns 1 and 2 include a set of year dummies. Three standard errors are reported: state clustered standard errors are given in parentheses “()”; Spatially adjusted (50km) standard errors are given by braces “{ }”; and spatially adjusted (100km) standard errors are given by brackets “[]”. Regressions include state fixed effects. Significance (using state clustered standard errors) *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table A2: Time of Introduction of Railways and Night Lights. India. 2013. Alternative Standard Errors

	(1)	(2)	(3)	(4)	(5)
Railway First 1850–1860	0.752*** (0.236) {0.257} [0.359]	0.602*** (0.219) {0.237} [0.322]	0.532** (0.226) {0.225} [0.285]	0.743*** (0.250) {0.277} [0.292]	0.807*** (0.171) {0.294} [0.299]
Railway First 1860–1870	0.640*** (0.144) {0.183} [0.215]	0.503*** (0.114) {0.178} [0.202]	0.403*** (0.142) {0.178} [0.200]	0.528*** (0.152) {0.176} [0.191]	0.279** (0.118) {0.264} [0.262]
Railway First 1870–1880	0.593*** (0.181) {0.191} [0.224]	0.405*** (0.141) {0.175} [0.210]	0.352** (0.149) {0.176} [0.206]	0.406** (0.157) {0.189} [0.203]	0.307** (0.143) {0.276} [0.282]
Railway First 1880–1890	0.387 (0.239) {0.171} [0.200]	0.241 (0.212) {0.156} [0.173]	0.254 (0.187) {0.159} [0.170]	0.214 (0.169) {0.167} [0.172]	0.049 (0.143) {0.265} [0.278]
Railway First 1890–1900	0.211 (0.201) {0.193} [0.215]	0.082 (0.163) {0.188} [0.201]	0.108 (0.166) {0.192} [0.196]	0.085 (0.148) {0.187} [0.197]	-0.272 (0.190) {0.304} [0.330]
Railway First 1900–1910	0.468** (0.186) {0.163} [0.207]	0.311** (0.129) {0.160} [0.197]	0.241* (0.142) {0.157} [0.189]	0.214 (0.164) {0.178} [0.191]	0.083 (0.134) {0.263} [0.276]
Railway First 1910–1930	0.137 (0.194) {0.268} [0.291]	0.110 (0.148) {0.223} [0.238]	0.069 (0.157) {0.223} [0.238]	0.428* (0.207) {0.241} [0.252]	0.321 (0.294) {0.267} [0.269]
Constant	13.890*** (3.579)	13.409*** (3.102)	11.713*** (2.572)	12.426*** (2.757)	9.408*** (2.714)
Latitude and Longitude	Yes	Yes	Yes	Yes	Yes
Population Distribution	No	Yes	Yes	Yes	Yes
Education	No	No	Yes	Yes	Yes
Geographic Features of District	No	No	No	Yes	Yes
Colonial Characteristics	No	No	No	No	Yes
Number of Districts	491	491	491	294	166

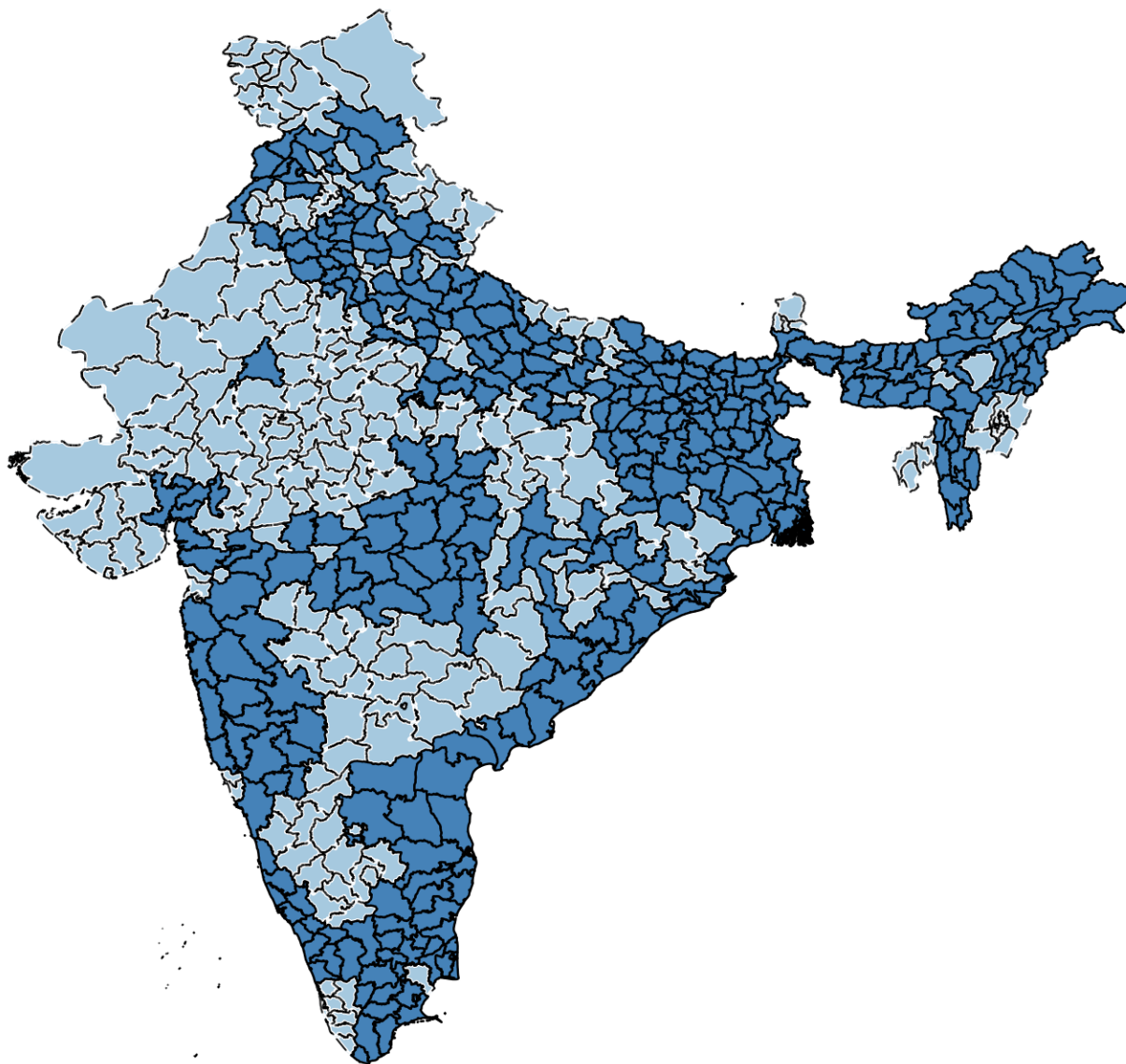
Notes: Sample restricted to present day districts in India. Dependent variable is $\log(1 + \text{sumlight}_{dst})$ in district d in state s in time t . Three standard errors are reported: state clustered standard errors are given in parentheses “()”; Spatially adjusted (50km) standard errors are given by braces “{ }”; and spatially adjusted (100km) standard errors are given by brackets “[]”. Regressions include state fixed effects. State clustered standard errors presented in parenthesis. Significance (using state clustered standard errors) *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table A3: Timing of Introduction of Railways and Night Light. Robustness of Results

	Night Light 1992		Alternative Explanatory Variable		
	(1)	(2)	(3)	(4)	(5)
Railway First 1850–1860	1.097*** (0.295)	1.316*** (0.318)	1.289*** (0.288)	0.928*** (0.181)	0.765*** (0.182)
Railway First 1860–1870	1.019*** (0.236)	1.012*** (0.271)	0.990*** (0.185)	0.531*** (0.163)	0.268** (0.119)
Railway First 1870–1880	0.987*** (0.208)	0.881*** (0.266)	0.676*** (0.218)	0.424** (0.172)	0.301* (0.146)
Railway First 1880–1890	0.609** (0.298)	0.569** (0.268)	0.423* (0.197)	0.249 (0.172)	0.031 (0.144)
Railway First 1890–1900	0.645** (0.239)	0.621* (0.306)	0.477 (0.278)	0.128 (0.162)	-0.292 (0.186)
Railway First 1900–1910	0.610** (0.271)	0.316 (0.304)	-0.151 (0.258)	0.237 (0.156)	0.075 (0.142)
Railway First 1910–1930	0.222 (0.284)	0.604* (0.287)	0.628* (0.307)	0.417** (0.177)	0.305 (0.297)
Latitude	-0.086** (0.034)	-0.042 (0.041)	-0.016 (0.028)	-0.048 (0.039)	-0.081*** (0.022)
Longitude	-0.018 (0.060)	0.016 (0.052)	-0.061 (0.052)	0.034 (0.036)	0.031 (0.045)
Proportion SC in District				-1.090 (0.761)	0.020 (0.805)
Proportion ST in District				-1.663*** (0.507)	-0.861 (0.781)
Proportion literate in District				1.811 (1.287)	3.643** (1.593)
Percent of District with mean elevation 500 m or more		-0.168 (0.464)	0.061 (0.446)		
River KM in district		-0.279 (0.197)	-0.483 (0.418)	-0.018 (0.085)	-0.210 (0.237)
Ruggedness				0.330* (0.167)	0.002 (0.150)
Princely State			-0.119 (0.260)		-0.148 (0.172)
Log Real Income 1930			0.225 (0.161)		-0.011 (0.114)
Log Area under Cultivation 1930			0.001 (0.176)		0.370*** (0.103)
Constant	15.144*** (2.758)	11.725*** (4.009)	7.775** (3.174)	11.955*** (2.862)	8.856*** (2.690)
Number of Districts	491	294	166	303	165

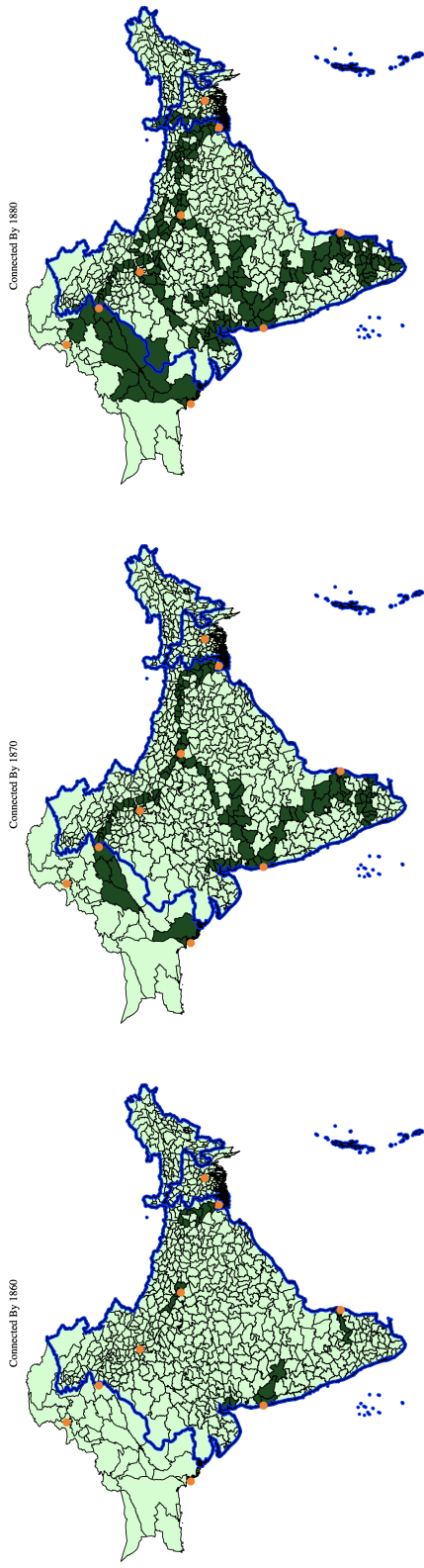
Notes: Dependent variable: $\log(1 + \text{sumlight}_{ds})$ in district d in state s (1992 in columns 1–3 and 2013 in columns 4 and 5). Sample restricted to present day districts in India. OLS regression results presented. Regressions include state fixed effects. State clustered standard errors presented in parenthesis. Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Figure B1: Classification of Districts based on type of rule in Colonial Period. India only.



Notes: Darker shaded districts denote under direct British Rule. Sample includes districts in present day India excluding those in Andaman and Nicobar Islands.

Figure B2: Districts Connected by Decade



Districts Connected by Decade (Continued)

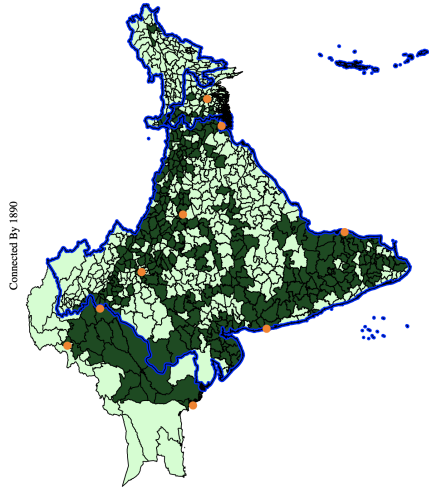
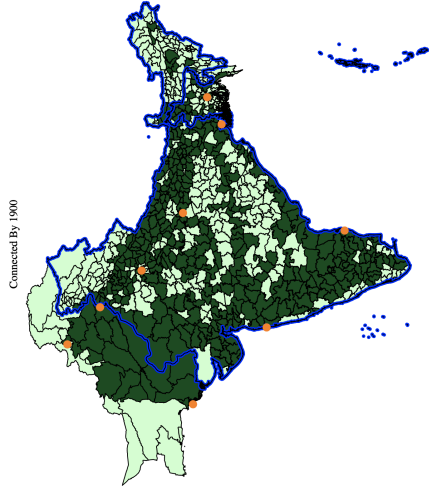
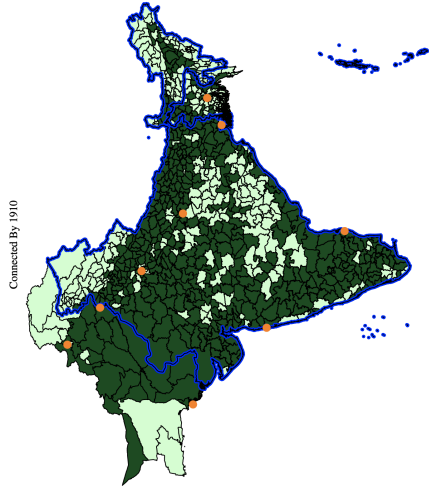
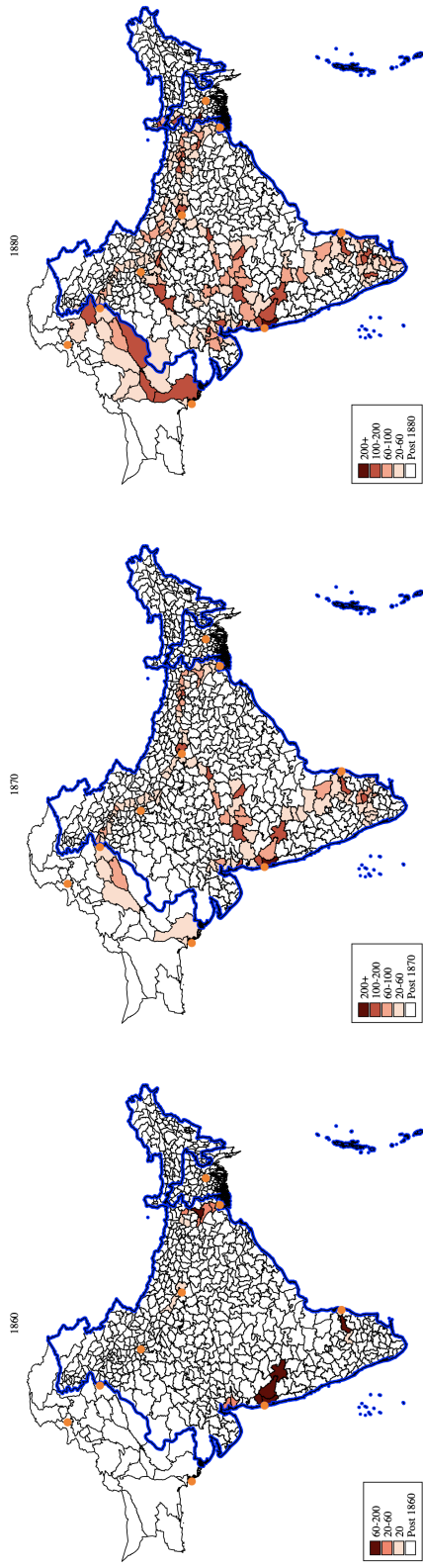
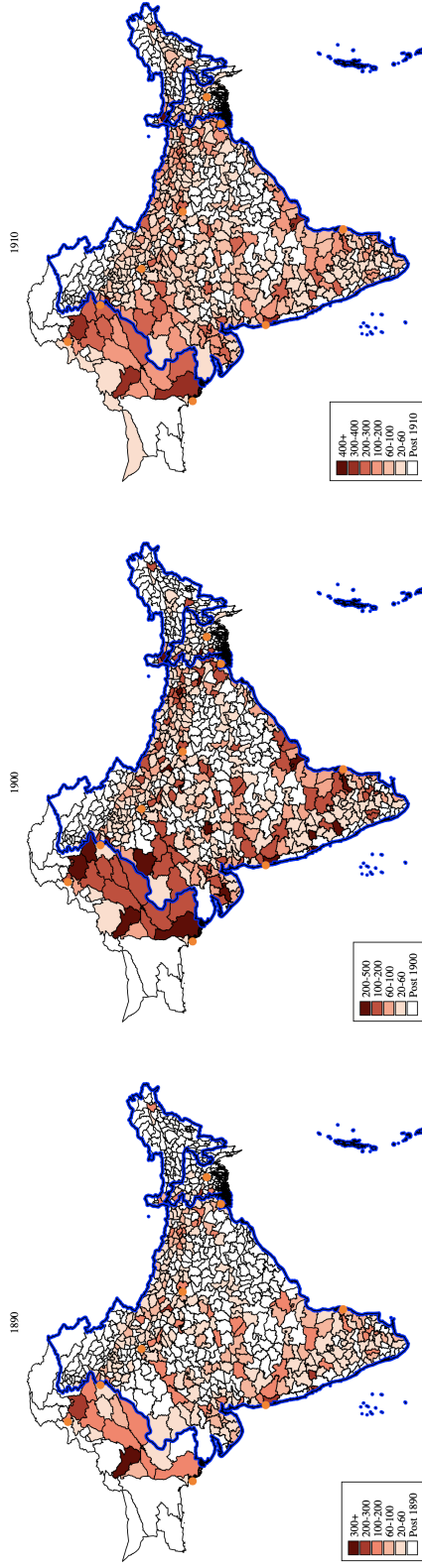


Figure B3: Track Length in District by Decade



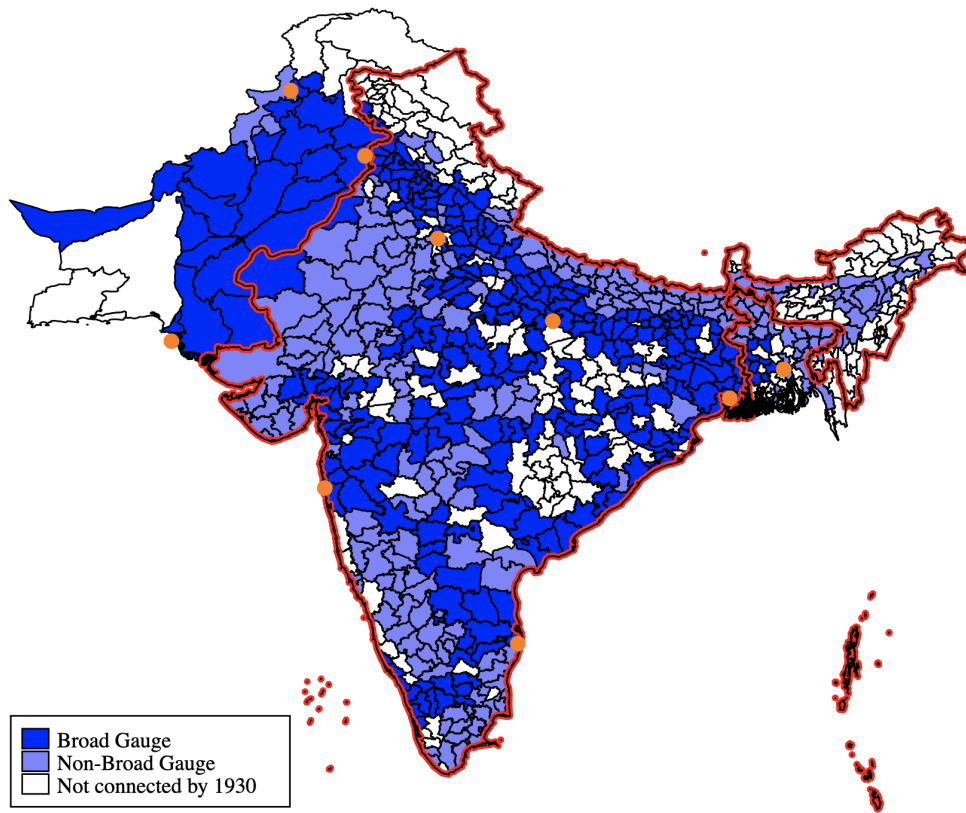
Notes: Sample includes all districts in present day Bangladesh, India and Pakistan.

Track Length in District by Decade (Continued)



Notes: Sample includes all districts in present day Bangladesh, India and Pakistan.

Figure B4: Distribution of Gauge type at Introduction



Notes: Sample includes all districts in present day India, Bangladesh and Pakistan.