On the Dispensability of New Transportation Technologies: Evidence from the Heterogeneous Impact of Railroads in Nigeria

Okoye, Dozie and Pongou, Roland and Yokossi, Tite

Dalhousie University, University of Ottawa, Massachusetts Institute of Technology

24 February 2017

On the Dispensability of New Transportation Technologies:
Evidence from the Heterogeneous Impact of Railroads in Nigeria[*]

Dozie Okoye
Dalhousie University

Roland Pongou
University of Ottawa

Tite Yokossi
MIT

February 24, 2017

Abstract

Exploring heterogeneity in the impact of a technology is a first step towards understanding conditions under which this technology is conducive to economic development. This article shows that colonial railroads in Nigeria have large long-lasting impacts on individual and local development in the North, but virtually no impact in the South neither in the short run nor in the long run. This heterogeneous impact of the railway can be accounted for by the distance to ports of export. We highlight the fact that the railway had no impact in areas that had access to ports of export, thanks to their proximity to the coast and to their use of waterways, and that those areas barely adopted the railway as it did not reduce their shipping costs. Our analyses rule out the possibility that the heterogeneous impacts are driven by cohort effects, presence of major roads, early cities, or missionary activity, or by crude oil production.

Keywords: Impact Heterogeneity, Colonial Investments, Railway, Africa, Long-run Effects, Development, Nigeria
JEL Classification: O15, O18, N30, N37.

1

1 Introduction

A number of recent empirical studies document the importance of transportation technologies for economic development in a variety of settings.[1] A question that naturally arises is whether these uncovered impacts of transportation technologies are homogeneous in connected areas, and if not, how much they vary based on pre-existing characteristics. In this article, we address these questions in the context of colonial railroads in Nigeria. Exploring heterogeneity in the impact of "new" transportation technologies is a first step towards understanding the conditions under which they lead to sustained economic growth and development. Identifying such conditions is important for the design of optimal policies on new transportation infrastructure.

In Nigeria, the railroads were primarily constructed to enhance export trade with Europe, but various parts of the country differed with respect to the availability of alternative transportation technologies and initial market access to Europe, as revealed by pre-railway trade volumes. The colonial railroads connected the interior of the country to coastal ports, but the South, by virtue of its proximity to the sea, already had viable alternatives, such as waterways and roads, which enabled trade with Europeans, and these alternatives co-existed with the railway. Such alternatives did not exist in the North, and the railroads were essential in opening this region to European trade and in shifting trade from the Sahara to the coast. These important baseline differences motivate our choice of Nigeria to explore the potential heterogeneity of the impact of railroads.

We proceed in a number of steps. First, we present a framework that enables us to causally estimate the long-run impacts of the railway on local economic development. Based on individual and household data from the 2008 Nigerian Demographic and Health Survey (DHS) and railway data from the Digital Chart of the World (DMA, 1992), the framework involves state fixed effects to explore differences between areas close to the railway and areas further away within the same state, and an instrumental variable approach involving the distance to straight lines between nodes[2] as an instrument for connection to the railway line. We use the framework to investigate the differential impacts of the railway in the North and the South of Nigeria, areas with significant differences in alternative transportation technologies. Our main finding is that the colonial railway has neither a short-run nor a long-run economic impact in Southern

---

[1]See Banerjee et al. (2012), Faber (2014), and Baum-Snow et al. (2012) on railways and roads in China; Donaldson (2016) on railways in India; and Donaldson and Hornbeck (2016) and Bleakley and Lin (2012) on the impact of the railway and portage sites in the United States. Jedwab et al. (2015), Jedwab and Moradi (2015), Fourie and Herranz-Loncan (2015), and Storeygard (2016) provide evidence on the positive economic impacts of roads and rail networks in Africa.

[2]Hypothetically, this is the shortest path between railway nodes.

Nigeria, but it has large positive impacts on local development in the North. This is true for broad indicators of economic development measured at the individual and household level. These measures include human capital, occupational characteristics, media access, household wealth, and urbanization.

Second, we analyze historical urbanization and city growth data from Jedwab and Moradi (2015) in order to establish that the heterogeneous impacts of the railway we document were also present (or absent) in colonial times. The non-impact of colonial railways in the South and its large positive impact in the North are stable over time and have persisted long after the railways became dysfunctional. These empirical findings are consistent with a theoretical model in which, in the North, railroads improved market access to Europe and encouraged the concentration of production factors in connected localities, inducing a spatial equilibrium that persisted in the long run, even after the demise of railroads. In the South, however, railroads did not significantly change the initial spatial equilibrium; they hardly had any short- or long-run local development impacts.

Third, we present evidence from historical sources that shed light on the potential underlying mechanisms. We document the fact that prior to the construction of the railway, Southern Nigeria was already involved in large and significant export trade with Europe, and earned a reputation as the "oil palm coast" well before the twentieth century (Law, 2002). Following the introduction of the railway, this trade did not grow significantly, compared to double digit growth rates in export of the main Northern crops.

Furthermore, railway adoption rates were significantly lower in the South, with less than 30% of Southern crops being shipped by rail compared to over 80% for most Northern crops. We find that the proportion of the main Southern crops that were railed to the coast significantly declined during the period of railway expansion, which indicates that railways did very little to stimulate economic activities in these areas. Our evidence on shipping volumes and costs suggests that the low adoption rates in the South are due to higher opportunity costs. While the railway decreased transportation costs in the North by more than 65% compared to roads, our calculations reveal that they were significantly more expensive than alternatives (roads and rivers) in the South.

In addition, we show that railroads have a persistent effect only in areas that were far away from ports of export. The North/South dichotomy in the impact of the railway can be entirely accounted for by the distance to ports of export. The railway had no impact in areas that were close to ports of export, no matter whether they are Northern or Southern areas. We also show that the heterogeneity of the railway impact cannot be

explained away by distances to early cities which were more prevalent in the South.

To uncover the long-term impacts of colonial railroads on local economic development, we exploit a number of empirical strategies, falsification exercises and robustness checks. Our first empirical strategy is to compare areas connected to the railway to areas unconnected to the railway within localities, or states, that were targeted. The assumption is not that states were exogenously connected, but that locations within a state were connected exogenously. In other words, the precise location of the railway, within a targeted state, was exogenous to local characteristics related to current development. For example, while the railway was intended to connect the areas surrounding Kano known for being suitable for groundnut production, the precise location of the line within the Kano area was plausibly exogenous to contemporary or future development. We claim that within a state, the railway was not systematically placed in the most developed localities or in localities that had the most potential for growth.[3]

To provide evidence that railway-connected localities and non-connected localities were similar, we compare railroad and non-railroad localities with respect to broad geographic and climatic determinants of development as well as presence (and size) of early cities and Christian mission stations. We do not find railroad and non-railroad locations to differ on these time-invariant geographic characteristics. Furthermore, we do not find that connected localities are more likely to have a Christian mission station, or to be connected to the road network or to have a river running through their local area.[4] Our estimates are robust to the control of the aforementioned factors in addition to other individual and household variables, including ethnicity and state fixed effects.

Nevertheless, in the absence of information on all the factors that contribute to local development, we are unable to completely rule out the claim that railroads were endogenously placed within states. In order to address further endogeneity concerns, we use an instrumental variable approach. We compute the distance to straight lines joining major nodes and use it as an instrument, for being connected to the railway line. This identification approach has been implemented in Banerjee et al. (2012), Jedwab et al. (2015), and Jedwab and Moradi (2015), among several others. Once again, within a state, the

---

[3]In fact, there are instances where the railway was located in less prosperous areas for a variety of geographic and other local idiosyncrasies. For example, the railway in "Lagos" began in neighboring Iddo because Lagos itself was an island. A second example is the line that terminated in the state of "Oil Rivers", which in fact ended in Port Harcourt, a city built from scratch, as opposed to more prosperous pre-colonial ports such as Bonny, Calabar, New Calabar, and Opobo.

[4]The local government is the smallest administrative unit in Nigeria with an average area of 1020 km^2, and median area of 705 km^2, and serves as the primary measure of the "local area" in which the individual lives. Individuals are identified by their DHS clusters which we refer to as localities, and we match localities to the local government area that they belong to.

4

straight line connecting nodes is the hypothetical line that, in theory, would have minimized construction costs, all else equal. Deviation from this line therefore might reveal the extent to which the actual rail trajectory might not have been chosen randomly by the colonial government.[5] The identification assumption in this empirical strategy is that straight lines between nodes affect economic development only through their correlation with actual lines. In implementing this method, we exclude observations in nodes, as these connected locations might have been endogenously chosen. We do not find the instrumental variable estimates to be very different from the estimates based on state fixed effects. We continue to find that the railway has a positive effect in the North and no effect in the South. Interestingly, the first stage of the instrumental variable results reveals that most geo-climatic and other local area variables are not significant determinants of connectedness to the railway lines.

We perform several identification checks. We show that our results are not driven by alternative transportation means. We also estimate the effect of placebo lines on our outcome variables. These lines are segments that were surveyed and proposed for railway construction, but were not constructed. They were abandoned for a variety of reasons unlikely to be related to short- or long-term economic development, such as the turnover of officials in charge of colonial railways and the conflicting interests of the colonial government (Jaekel, 1997). If the effects of the railway we identify using our instrument reflect the developmental impact of the railway on localities closer to a straight line connecting nodes, we would expect the placebo lines to have no impact, especially in areas where placebo lines do not coincide with roads or waterways. Indeed, we do not find the various placebo lines to have any economic effect, whether the effects are estimated for the whole country or separately for the North where the railway had a significant impact.

Additionally, we use localities close to these placebo lines as a control group to analyze the impact of the railway. Precisely, we estimate the effect of being within 20 *km* of a railway line relative to being within 20 *km* of a placebo line. We find a large economic effect of the railway in the North. In the South, the effect is close to zero and is not statistically significant. These findings provide further evidence that the impact of the railway in the North and its non-impact in the South are indeed causal, and not merely driven by being close to a "straight line" connecting early urban areas. This is especially true if the placebo lines were not constructed for idiosyncratic reasons, as Jedwab et al.

---

[5]This argument ignores other geographic and climatic conditions that might call for deviations from straight lines between nodes. Deviations from straight lines might not be endogenous to the local economic development of connected areas if they are motivated by technical or geo-climatic characteristics of the localities the railway passes through.

(2015) and Jedwab and Moradi (2015) argue.

Our results are robust to a variety of other confounders. In addition to the geo-climatic variables discussed above, our estimates are robust to controlling for the presence of mission stations within the local area and for distances to rivers and roads (possibly endogenous to railways). The estimated impacts of the railway lines are also found for migrants and non-migrants, and areas with and without mission stations. We detect a positive impact of the railway when we exclude local areas that are run through by railway tracks, nodes or stations. In addition, our results are robust to the exclusion of oil-producing areas of the South that might have altered the post-railway spatial equilibrium. Using a continuous measure of connectedness to the railway line (distance to a railway line) instead of the discrete measure in our main specification (being within 20 km of a railway line) yields quantitatively similar results. The estimated impacts of the railway lines are robust to various definitions of the control group: individuals outside 20 km, between 20 to 40 km, between 40 to 60 km, and so on. We further interpret these findings as evidence that there are no significant negative spillovers to adjoining localities.

A few recent studies have examined the long-run economic impact of infrastructure investments (Huillery, 2009; Banerjee et al., 2012; Baum-Snow et al., 2012; Donaldson and Hornbeck, 2016; Donaldson, 2016; Bleakley and Lin, 2012). In Africa, urbanization is found to have been sustainably impacted by colonial railroads (Jedwab et al., 2015; Jedwab and Moradi, 2015). Similar to the argument we provide to shed light on the long-run economic impact of colonial railroads in Northern Nigeria, Jedwab and Moradi (2015) explain that, in Ghana, colonial railways lowered trade costs and boosted the cultivation of cocoa in railroad locations, fostering the emergence of cities in these locations. This initial spatial equilibrium persisted because railroad locations facilitated the coordination of subsequent investments.[6]

Our study differs from these papers in a number of important respects. First, we are interested in the heterogeneity of the railway impact. In particular, we do not find that colonial railways have had any local economic impact in Southern Nigeria, and areas closer to the coast, in contrast to their positive impacts in the Northern regions of the country. In further contrast to the average impacts of the railway we estimate, and hence to the general conclusions reached by the extant literature on the long-run effect of colonial railways on urbanization in Africa, we find no evidence that colonial

---

[6]Also see Storeygard (2016) who underscores the importance of road networks and connection to coastal ports for local economic performance, and Fourie and Herranz-Loncan (2015) who document the importance of the railway in South Africa.

railways were the engine of urbanization in Southern Nigeria. In fact, most cities in the South do not lie along the railway line, while almost all large cities in the North are connected to the railway. This is important for understanding the policy implications of recent studies of the impact of transportation investments. Our results suggest that these investments are most worthwhile in areas where they significantly improve market access and stimulate new trade. The contrast between the low adoption rates and impacts of the railway in the South, and the high adoption rates and impacts in the North, is consistent with the views of Fogel (1964) that new transportation technologies have little impact if pre-existing technologies are viable or can readily be improved.

The results on the impacts of railways on individual-level developmental outcomes are also of independent interest to studies of African urbanization. Fay and Opal (2000) and Jedwab and Vollrath (2015) document the poor economic performance of several urban areas in developing countries, compared to historical examples from other regions. We find that, in areas without pre-existing viable transportation technologies, connection to the colonial railway increased urbanization, and that individuals living in these urban areas are more educated, more literate, more likely to work in professional occupations, less likely to work in agriculture, more likely to engage with mass media (TV, radio, newspapers), and live in wealthier households. This suggests that, while urban areas are not industrializing or growing as fast as one would expect from historical examples, urban areas connected to the railway are still generally better off than surrounding countrysides. Furthermore, the fact that, within rural areas, colonial railways have a positive impact on individual-level economic outcomes in the North of Nigeria implies that estimated impacts are not entirely driven by urbanization.

The rest of the article is organized as follows. Section 2 depicts the historical context of the construction of the colonial railway in Nigeria. Section 3 presents our data and the various identification strategies and robustness checks that we use to assess the local impact of the railroads. The corresponding results are described in section 4. Section 5 analyzes the North-South differences in the long-run impact of the railway. Section 6 discusses the dynamics of the path of the impacts of the railway and compares short- and long-run effects. Section 7 documents the mechanisms underlying the heterogeneity of the impact of railroads. The final section concludes.

2   Historical Background

> So vast an area as Nigeria, comprising in all some 380,000 square miles...
> cannot be commercially developed except by railways — *p. 19 0f the Colonial*

Toward the end of the nineteenth century, after the area now known as Nigeria officially came under British control, the colonial government began to seek out ways of linking the interior of the country to its ports in order to facilitate export trade. The construction of the railway was seen as an effective means of moving goods and services from the interior of the country to the coast. Construction of the railway lines largely occurred between 1898 and 1930, with an additional extension completed after independence in 1964.[7] The railway was generally constructed to open up the country to export trade with Europe. Three specific reasons were given for the construction of the various lines: agricultural, mineral exploitation, and political or administrative reasons (Taaffe et al., 1963; Onyewuenyi, 1981).

Table 1 shows the dominant motivations for each of the lines constructed between 1898 and 1964. It establishes that the export of agricultural products was the main motive for the railway. Of all the segments shown in Table 1, only Zaria-Jos-Bukuru and Kaduna-Kafanchan were not constructed for agricultural exploitation reasons. In terms of spatial distribution, the colonial railroads were slightly more extensive in the Northern region which covers 4/5th of the country's area.[8]

2.1 Alternative Transportation Modes

Before the railways, transportation of goods was done through head portage, bicycles, animals, cart and inland waterways. In the North, there were caravan routes going through Timbuktu to major agglomerations such as Kano and Sokoto and on to North Africa (Cairo, Tripoli). One consequence of the railway, as we discuss later in the article, was to redirect Northern trade from Trans-Saharan routes to the coastal ports.

The most important transportation mode for goods before the advent of railways were inland waterways. Many rivers, their tributaries, and creeks traverse the coastal plains of the country. In the South, between the coastal ports of Lagos and Opobo for example, the abundant creeks allow transportation of produce and many ports were installed along the way: Epe, Sapele, Warri, Forcados, Burutu, Brass, Degema, and Port Harcourt (Onyewuenyi, 1981). These river networks, as well as direct access to roads using bicycles, exposed the South to trade with Europe long before the railway was constructed.

---

[7]A rail line joining Abuja to Kaduna was built between 2011 and 2014. Since it was constructed after the dates of our outcome measures, it should have no bearing on our results.

[8]There is an average of 3.1 rail length meter per area square kilometer in the South and of 3.4 in the North.

8

While the country's two main rivers, the Niger and the Benue, run through the North, the rivers are navigable only for part of the year and for a fraction of the distance they cover. They are heavily dependent on water levels in the rainy season, and the Niger itself is filled with dangerous rapids. As a result, the only available means of consistent transportation from the Northern parts of the country to the coast was through roads, which were not viable because of the enormous distance and other dangers of road transportation in pre-colonial Nigeria.[9] For example, Hodder (1959) in a study of tin-mining in Jos estimated that the road journey from the mines to the coast took 35 days by road, and while this was tolerable for mining, it was not conducive to agricultural exports. These problems with river and road transportation meant that most areas of the North were cut off from export trade prior to the construction of the railway.

2.2 Railway Construction

The railway construction was done in three main phases. The first phase consisted of initial penetration lines. The origination points were the ports of Lagos, Zungeru and Baro for the Western line, and Port Harcourt for the Eastern line. The Western line originated in Lagos in 1898 and reached the Niger river at Jebba in 1909. The construction of the Eastern line began in 1913 in Port Harcourt and reached Enugu by 1916. In the second phase of the railway development, more interior centers were linked to the ports with lines such as Baro-Kano and Enugu-Kaduna. By 1927, both main North-South links were established giving Northern areas access to the ports of Lagos and Port Harcourt. Building branch lines and extensions such as Zaria-Kaura Namoda or Kano-Nguru made up the third stage of railway development. At the end of this phase, in 1931, the railway was 3,067 *km* long.

New centers of economic activities quickly appeared along the newly constructed railroads. By the time the main lines were built, more than 200 buying and selling stations had emerged along the railway lines (Onyewuenyi, 1981). One of the fastest growing centers was the coastal town of Port Harcourt which was chosen as a terminal node of the Eastern line in 1913 before it even existed as a town. Because of its deeper harbor and direct access to the hinterland, Port Harcourt had developed, by the 1930s, as the second largest port of the country, at the expense of previously established non-railway ports within the region such as Bonny, Opobo, or Degema. Similarly, in the "Lagos" area, the railway did not begin in Lagos itself but in another town known as Iddo because Lagos is an island which would have made construction more expensive.

---

[9]On average, Northern populated areas and coastal ports are more than 600 km apart.

We exploit these local idiosyncrasies to motivate one of the empirical strategies we use to estimate the impact of the railway.

2.3 Growth of Export Agriculture Following the Railway Construction

> In the Northern Provinces, the history of export cotton production, like that of groundnut has been closely linked with the history of railway expansion, and it was not until the railway reached Kano in 1912 that the export cotton production attained any importance — *Lamb 1925, p. 19*.

The incentives to produce more than what was needed for consumption were weak in remote areas in the North of the country, especially in areas poorly connected to rivers. The advent of the railways dramatically changed the trade opportunities available to these areas. The railways were used almost exclusively for goods transportation as more than 90% of rolling stock units were devoted to goods service. Over the period 1901-1950, an average of 2/3 of these goods were agricultural products. According to the Colonial Reports of 1913, only a year after the first railways were built in the North of Nigeria, the value of Northern agricultural exports jumped by 150% (groundnuts by 666%, benniseed by 157%, gum arabic by 133%, cotton lint by 45%, hides and skins by 41%, and sheanut products by 20%). Acreage under cultivation increased at all station areas.

The railway stations allowed the concentration of markets along the railroads making possible the clustering of a traditionally scattered population and agricultural production. The development of export agriculture was initially limited to "the irradiation area of the railways, the inter-regional roads and auxiliary local roads" (Schätzl, 1973, p. 89). As our results indicate, the incredible boom in export agriculture that followed the railway construction had long-run consequences on the human and social development of people living in areas connected to the railways. This impact was concentrated in the North of the country, which was the main beneficiary of the introduction of railways.

Next, we describe the empirical strategies and data we use to estimate the long-term impact of the colonial railway in Nigeria, and how this impact differs according to initial market access.

3 Data and Empirical Strategy

3.1 Data

Data on colonial railroads in Nigeria come from the Digital Chart of the World (DMA, 1992). These data are combined with individual-level data from the 2008 Nigeria *Demographic and Health Survey* [DHS] (NBS and ICF International, 2008) to estimate the long-run impacts of railroads. The DHS uses a two-stage probabilistic sampling technique to select clusters in the first stage and households in the second stage. In general, DHS clusters are census enumeration zones, to which we will sometimes refer simply as localities. Using DHS-provided information on the geographical coordinates of each such locality, we match individuals to local areas[10] and rail networks.

The DHS provides information on each individual's characteristics including age, sex, migration status, religion, ethnicity, and area of residence. Individuals without specific information on ethnicity are dropped from the analysis. Information on each of our individual-level outcome variables - years of schooling, literacy, type of employment (professional or agricultural), and the frequency at which an individual reads newspapers, listens to the radio, and watches TV - is also available in the DHS. Household-level variables include size, the gender and age of the household head as well as a composite wealth index.[11]

We complement our outcome dataset with panel data on urban population density and city presence in 1900, 1960, and 2010, from Jedwab and Moradi (2015). This allows us to analyze urbanization outcomes and explore short- and long-run effects.

We collect detailed information on geographic, climatic, and soil conditions from the FAO GAEZ database (Fischer et al., 2008). Specifically, we gather information on average annual rainfall (in mm), average annual temperature (in degree Celsius), elevation (in meters), two important soil characteristics (nutrient retention capacity and workability), and suitability for the production of cocoa, cotton, groundnuts, and oil palm.[12]

---

[10]The local government area is the smallest administrative unit in Nigeria with an average area of 1020 km^2, and median area of 705 km^2, and serves as the primary measure of the "local area" in which the individual lives. Their boundaries are obtained from the GADM database of global administrative areas (UC Berkeley, 2014).

[11]A measure of households' cumulative living standard based on observables such as asset ownership (radio, TV, bicycles etc.), materials used in housing construction, water access, and sanitation. See http://www.dhsprogram.com/topics/wealth-index/Wealth-Index-Construction.cfm for more details on the construction of this index by DHS country teams.

[12]The soil characteristics are measured on a 4-point scale ranging from no or slight constraints (1) to very severe constraints (4). Crop suitability is the average estimated agro-climatically attainable yield in kg/ha for rain-fed agriculture, using medium or low inputs within the local area because that is the dominant form of agriculture in Nigeria. These measures are provided by the FAO for cells of 30×30 arcseconds

We also collect information on the presence of primary roads and major rivers in each local area. These data are supplied by DMA (1992) and are available in Hijmans et al. (2001). Lastly, we collect data on the presence of Christian mission stations in local areas in 1928 by combining maps published in Ayandele (1966) and Roome (1925). Altogether, we have information on over 30,000 individuals living in 22,798 households belonging to 845 clusters (localities) spread out over 550 local government areas in 37 states. These individuals belong to 30 major ethnic groups that make up over 90% of the country's population.

3.2 Identification Strategies

Since railway-connected and railway-unconnected areas might differ on many dimensions that are relevant to economic development, comparing them does not necessarily yield the long-term impact of the railway. We use a mix of strategies to deal with the possible endogeneity of railway placement. We first use a state fixed effects approach that compares connected and unconnected areas within states that were targeted or not to host railroads. We then exploit an instrumental variable strategy based on straight lines between major railway nodes. We complement these strategies with a host of falsification exercises and identification checks using placebo lines and various definitions of control groups.

3.2.1 State Fixed Effects

The railway was intended to connect large areas suitable for agricultural and mineral exploitation to the coast. In order to avoid comparing targeted areas with non-targeted areas, we include in our regressions state fixed-effects as states are the closest administrative level to capture these areas of interest.

Within states, we compare areas that are close to the railway lines to areas that are further away. A concern with this strategy is that even within targeted areas, the railway might be placed in areas that are systematically different on dimensions relevant to economic development. Historical accounts of railway placement tend to indicate otherwise. Engineering decisions that took into account elevation and other considerations based on the cost of the railway construction were central to determining the exact location of the railway within a targeted state. For instance, the railway line that terminates in Lagos state actually terminated in a small locality called Iddo because the city of Lagos was an island and having the terminus in such a city would have been very expensive.

(approximately 0.9 km^2 at the equator)(FAO, 2016).

12

A second example is the line that terminates in the "Oil Rivers" which is now the city of Port Harcourt. Despite the fact that the city did not even exist at the time, it was chosen instead of prosperous pre-colonial ports such as Bonny, Calabar, New Calabar, and Opobo, because of its deeper harbor and more direct access to the hinterlands.

Comparing observables within railway-connected and unconnected area lends further support to the claim that the exact placement of the railway within targeted areas does not appear to systematically be in areas more developed or more likely to be developed in the future. We refer to localities within 20 *km* of a rail track as connected areas. Table 2 presents summary statistics for various observable characteristics in connected and unconnected areas. Baseline observables such as crop suitability, geo-climatic factors and soil constraints do not exhibit significant differences in observable characteristics between connected and unconnected localities. Strikingly, connected individuals are not more likely to live in local areas with mission stations, nor in local areas with primary roads, nor in areas crossed by a major navigable river. At the introduction of the railway, there were no statistically significant differences in city presence and city population between connected and unconnected areas.

Nevertheless, we control for all these geographic, climatic, population-based and other pre-railway observables. For individual and household-level outcomes, we control for additional factors such as age, religion, ethnicity and household size in order to get more precise estimates of the railway effect.[13]

Specifically, our baseline specification for individual- and household-level outcomes is:

$$Y_{i,h,a,e,s} = \phi + \beta R_{i \leq 20} + \mathbf{X_a \Lambda} + \mathbf{X_i \Pi} + \mathbf{X_h \Gamma} + \alpha_\mathbf{s} + \gamma_\mathbf{e} + \epsilon_\mathbf{i,h,a,e,s} \qquad (1)$$

The parameter of interest, β, is the effect of living within 20 *km* of a railway track on outcome Y. The outcome is measured for each individual i in household h, who lives in local area a in state s, and belongs to ethnic group e. The outcome variables are education (years of schooling, literacy), occupation (professional or agricultural worker), media access (newspaper, radio, TV), the DHS-based household wealth index and the probability of living in an urban area. If the railway has a positive impact on local development today, then individuals who are closer to the railway should be more educated and more literate, and should have greater access to the media. Under this hypothesis, we would also expect individuals in railway areas to be more likely to live in urban areas

---

[13]While we use living within 20 *km* of a rail line as the measure of connectedness, we confirm that observables are also balanced using other measures of connectedness such as an indicator for having the railway pass through the local area. We also divide the country into 40 *km* × 40 *km* grid cells and show that observables are balanced between connected and unconnected grid cells. These results are available upon request.

and in wealthier households and to be non-farm workers.[14]

Individual, household, and local area observable characteristics are denoted by $\mathbf{X_i}$, $\mathbf{X_h}$, and $\mathbf{X_a}$, respectively. All regressions involving individual- or household-level outcomes include state fixed effects (α_s) and ethnic group fixed effects (γ_e). Individual controls ($\mathbf{X_i}$) are age, age-squared, and an indicator for being Christian. At the household level, we control for gender and age of the household head, as well as for the size of the household. At the local area level, we control for average rainfall, temperature, soil nutrient retention capacity and workability, elevation, suitability for key colonial area cash crops (oil palm, cocoa, cotton, groundnut), and presence of a mission station in the local area as at 1928. In the remainder of the article, these local area controls are referred to as baseline controls.

For urbanization outcomes (city presence and city growth), our baseline specification has the same structure as in (1) but the set of controls excludes individual and household-level controls:

$$Y_{a,s} = \phi + \beta R_{i \leq 20} + \mathbf{X_a \Lambda} + \alpha_\mathbf{s} + \epsilon_\mathbf{a,s} \tag{2}$$

3.2.2 Instrumental Variable Approach

To further address endogeneity concerns, we adopt an instrumental variable approach, similar to that used in Banerjee et al. (2012) and Jedwab and Moradi (2015). We exploit the distance to straight lines joining major nodes and use it as an instrument for being connected to the railway line. We also exclude individuals living in railway nodes from the sample. The identification assumption is that, besides its correlation with the railway line, a straight line connecting nodes is unrelated to economic development.

The major nodes are chosen to be major historical cities existing at the time of the introduction of the railway such as Lagos, Abeokuta, Ibadan, Kano, and junctions in the middle of the country such as Kafanchan. We connect these 12 major nodes in the spirit of the railway introduction, that is, by finding the minimal path to connect areas of interest to the coast for each railway line defined by periods of planning/construction (pre-1912, 1916-1930, and 1964). The result of this simple algorithm is shown in Figure 2.

The instrument for being within 20 *km* of a rail line is defined as the log of the dis-

---

[14]We categorize an individual as being literate if they can read some or parts of a sentence. We deem them as utilizing media resources (newspaper, radio, TV) if they use these resources at some point during a month. We adopt broad and inclusive definitions in order to provide conservative estimates, and deal with concerns about arbitrary cutoffs for inclusion into these categories. Our results are in fact stronger if literacy is limited to individuals who can only read whole sentences, or to individuals who use media at least once a week.

tance to node-joining straight lines. All observations within the same local government area as a node are dropped. Only "intermediate" observations are used to estimate the specification.

3.2.3 Identification Checks

We use a number of exercises to confirm the causality of the effects we uncover.

Alternative Transportation Technologies. A natural concern with our empirical strategies is that distances to straight lines connecting nodes could be correlated with roads and other transportation technologies and that the impacts we bring to light might be unrelated to the railway line. We address this concern by demonstrating that the estimated impacts of the railway line are robust to the inclusion of other transportation technologies.

Placebo Lines. Following Donaldson (2016), we use placebo lines to test whether the effects we measure have to do with the railway and not merely with being close to lines joining nodes that could have been linked by the railway or any other important transportation technology. Placebo lines are segments that were surveyed and proposed for railway construction but were never actually constructed. Given the prohibitive costs of railway surveys, these segments were seriously considered. They were ultimately abandoned for a host of plausibly exogenous reasons that have to do with the turnover of officials in charge of the Nigerian railways and the conflicting interests of the British colonial decision-makers (unlikely to be related to local economic development). The data on placebo lines comes from Jaekel (1997) who lists the lines that were extensively surveyed but ultimately not constructed.

As Figure 3 shows, the placebo network that we reconstruct covers an extensive part of the country. In the Southern part of the country, these lines were meant to connect already existing cities. Following independence, these very early cities were finally connected by roads. Thus, in this exercise, we control for the effects of alternative technologies (roads and rivers). We are aware of the fact that the placement of roads, presented in Figure 4, might be endogenous to the existing railways. Hence, the results of this identification check are only suggestive and should be interpreted with caution. Our hypothesis is that if the effects that we uncover are indeed causal, one would expect them to disappear or reverse once we replace actual railway lines with placebo lines, especially in areas where placebo lines do not coincide with roads or waterways.

15

Actual Lines versus Placebo Lines. As explained above, it is plausible to assume that surveyed localities were de facto exogenously assigned to two groups: the group of localities connected to the railway, and the group of localities that were surveyed but not connected (placebo lines). This makes the use of areas close to placebo lines a powerful control group to check our identification. We implement this by re-estimating specification (1) on the clusters that are within 20 *km* of rail or placebo lines. This analysis would effectively yield a causal effect of the railway if, as argued above, placebo lines were not constructed for exogenous reasons (or reasons not related to long-run economic development).

Varying Control Groups. Another concern with our identification strategies is that we might be merely identifying differences between localities within 20 *km* of a railway and those very far away, such as clusters over 80 *km* away from the railway line, which may not be good control clusters. This is because clusters that are very far away are more likely to be different in dimensions not captured by our control variables. To address this concern, we break down the control group into various distances: clusters within 20-40 *km*, 40-60 *km*, 60-80 *km*, and farther than 80 *km* of the railway line.[15] We then re-estimate equation (1) using the disaggregated distances as different control groups, and exclude the indicator for individuals living beyond 80 *km* of the railway from the regression. The coefficient on living within 20 *km* of a railway line now represents the impact of the railway relative to individuals living beyond 80 *km* of a railway. This strategy allows us to compare individual outcomes across multiple distances and to account for potential spatial spillovers beyond 20 *km*. A similar exercise consists of estimating equation (1) on samples limited to areas within 40 *km*, 60 *km* and 80 *km* of the railway. Both exercises yield similar results.

Other Identification and Robustness Checks. We complement our empirical strategies with the following identification and robustness checks. We test the robustness of our results to other measures of connectedness such as continuous measures based on the distance to the rail lines or the distance to the closest railway station or an indicator for being in the same local government area as the rail line.

We also test the robustness of our results to: using Conley standard errors to correct for spatial autocorrelation, limiting the sample to rural areas only, limiting the sample to migrant individuals or to non-migrants, excluding areas with mission stations, excluding

---

[15]In our dataset, we find that 32% of individuals live within 20 *km* of a railway track, 12.3% within 20-40 *km*, 13.57% within 40-60 *km*, 11.42% within 60-80 *km*, and about 30.61% farther than 80 *km* of a rail line.

local government areas with rail tracks or rail stations, and even excluding areas within 20 *km* of a railway station.

4   Average Effect of the Railway: Countrywide Estimates

4.1   State Fixed Effects Results

Table 3 presents results of the estimation of our main specification (1). Standard errors are clustered at the local government area level in order to deal with arbitrary correlation between localities (clusters) within local areas.[16] As reported in Column 1, living within 20 *km* of a rail line increases schooling attainment by 1.37 years on average. This is associated with a 14% increase in the probability of being literate, a 1.8% increase in the probability of working in a professional wage job, and a 7.1% decline in the probability of being an agricultural worker (Columns 2-4). Furthermore, being connected to the rail line is positively associated with media access. Individuals in connected areas are more likely to read newspapers, listen to the radio and watch TV (Columns 5-7). Finally, being connected to the rail line is associated with living in a wealthier household, and it increases the probability of living in an urban area by 18% (Columns 8-9). These results are all consistent with a strong positive impact of proximity to the rail line on individual and household development outcomes.

It is interesting that none of the geographic and climatic variables have a consistent impact on the outcomes. Similarly, we do not find that areas that are more suitable for oil palm, cocoa, cotton, or groundnut are more developed in the present. Importantly, this evidence supports our identification assumption that, within a state, geographic characteristics, and any pre-colonial advantages they might have conferred, are largely unrelated to contemporary development. However, and in accordance with previous studies, we find that missionary activity is strongly associated with development at the local level (Gallego and Woodberry, 2010; Nunn, 2014; Okoye and Pongou, 2014; Cagé and Rueda, 2016; Wantchekon et al., 2015). Local missionary activity has a positive impact on years of schooling, literacy, occupational choices, media access, household wealth and urbanization. The impact of the railway and missionary activity, and the non-impact of geographic and climatic variables, are remarkable and speak to the importance of historical circumstances for development at the local level.

---

[16]We show in a Section 4.4 that our results are robust to using Conley standard errors to correct for spatial autocorrelation. We find that the Conley standard errors are similar to the cluster-robust standard errors used throughout the article (See Table A2).

4.2   Instrumental Variable Estimates

Before turning to instrumental variable results, we first note that the first-stage estimates (presented in Table 4) indicate that a doubling (100% increase) of the distance to a line joining nodes lowers the probability of being connected to the rail line by about 29.4%. Also important is the finding that missionary activity and most of the geo-climatic variables are uncorrelated with rail presence. An exception is elevation which is negatively correlated with probability of connectedness, a result consistent with a higher cost of building in elevated areas.

The 2SLS estimates of the impact of being connected to the railway are shown in Table 5. They are generally not statistically different from OLS estimates. An exception is the estimated impact of being within 20 *km* of a rail line on the probability of being an agricultural worker, which falls in magnitude from -7.1% (OLS) to -4.0% (IV). The IV approach, along with the robustness checks results reported below, lend support to our causal interpretation of the impact of the railway.

4.3   Identification Checks Results

Alternative Transportation Technologies.   As shown in the top panel of Table 6, we control for the presence of other transportation technologies, because if, for instance, roads were built close or along some of the railway lines, their impacts could be picked up by our estimates. Indeed, when roads and rivers are controlled for,[17] we find a robust long-term impact of the railroads. The fact that the results are not driven by the correlation of the railway with other transportation infrastructure is important as, otherwise, this might call into question the identification assumption behind the IV strategy. We find that increased distance to roads is negatively correlated with living in an urban area, household wealth, and other measures of local development. Rivers, on the other hand, are not positively related to development, consistent with our findings on other geographic variables. However, we do not take the estimated impacts of the road network as causal because they might have been constructed in response to the rail network. These results suggest nevertheless that the estimated impact of the railway cannot be explained away by the road network.

Placebo Lines.   We conduct two different exercises involving placebo lines. First, we replace actual railways by placebo rail lines (the surveyed lines that were eventually not

---

[17]Data on the presence of major roads and navigable rivers are obtained from DMA (1992) and are available in Hijmans et al. (2001).

constructed), and exclude railway-connected areas. As a result, most estimates decrease dramatically and lose their statistical significance. As shown in the middle panel of Table 6, the estimates of the coefficients on outcomes such as schooling, literacy, professional occupation, reading newspapers and being an urban resident are significantly smaller and not statistically different from zero.

In the bottom panel, we control for the presence of other transportation technologies, because transportation technologies were built on some of the placebo segments, and their impacts could be picked up by our estimates. Indeed, when roads and rivers are controlled for, the results for the placebo effect, presented in the bottom panel of Table 6, paint a clear picture. Most of the coefficient estimates are not significant at the 5% level.

Second, we use localities close to these placebo lines as a control group to analyze the impact of the railway. Precisely, we estimate the effect of being within 20 *km* of a railway line relative to being within 20 *km* of a placebo line. As shown in the top panel of Table 7, we find a large economic effect of the railway compared to placebo areas. These results suggest that localities closer to the railway lines have better development outcomes today than localities close to other straight lines joining major cities and initially proposed to be connected by rail.

By providing evidence that straight lines connecting pre-existing cities are not correlated with local development outside of localities connected to the railway line, these placebo results reinforce our causal interpretation of the effect of railroads.

Varying Control Groups.   As shown in the bottom panel of Table 7, the impact of the railway on individuals living within 20 *km* of the rail line is robust to the use of different control groups. The impact of being connected to the rail line does not change significantly when we compare individuals living within 20 *km* of the rail line to those living within 20-40 *km* or to those living farther away. In fact, the estimated impact of the railway is stronger when individuals living within 20 *km* of a rail line are compared to those living a further 20 *km* away at most. For example, relative to individuals living beyond 80 *km* of a railway line, being connected is associated with an additional 1.23 years of schooling. However, when compared to those living just "outside", within 20-40 *km*, the impact of being connected is an increase of about 1.6 years of schooling $(1.231 + .362)$ although this difference is statistically insignificant. Thus, extending the control group to all distance groups actually produces a conservative estimate of the impact of the railway as there is no evidence of a positive spillover beyond 20 *km* of a rail line.

Overall, our results strongly suggest that being connected to the historical railway

line has a positive impact on development in today Nigeria, even though the railway has deteriorated after the country's independence.

4.4 Additional Identification Checks Results

We also confirm that the estimated impact of the railway is not sensitive to the measure of connectedness. Although the point estimates are not directly comparable, the impact of the railway is robust to the use of a continuous measure of closeness to the rail line.[18] The corresponding results are presented in the top panel of Table A1. Our results are also robust to using other measures of connectedness to the railway such as proximity to railway stations (second panel) or an indicator for being in the same local government area as the rail line (bottom panel of Table A1).

In Table A2, we show that our results are robust to using Conley standard errors to correct for spatial autocorrelation.[19] We find that Conley standard errors are not much different from the cluster-robust standard errors used throughout the article.

We continue to find an impact of the railway when we exclude individuals living in urban areas (top panel of Table A3). This important result indicates that our results are not merely driven by urbanization. We also estimate the differential impacts of railways by migration status (second and third panels of Table A3). While the impacts are larger for migrants, the estimates are not statistically different from the estimated impact on non-migrants.[20] This suggests that the long-term effects of railroads are not driven primarily by migrants who might have higher ability, education or skills.

The top panel of Table A4 shows that our results are not driven by missionary activity, since they are robust to excluding areas without mission stations. If anything, the impacts of the railway is stronger in localities without mission stations. In areas with missionary activity, missions had a positive impact on schooling, literacy, and media access, possibly attenuating the impact of the railway on these outcomes.

The second panel of Table A4 indicates that the impact of the railway is sustained when we exclude local areas that contain rail tracks. This is further evidence that the effect that we find is driven by connectedness to the railway line, and not merely by the presence of a railway line or station in the local area. Finally, an important part of the connectedness to the railway line is the proximity to railway stations. The bottom panel

---

[18]It is defined as $-log(1 + clusterdistance)$ where *clusterdistance* is the survey cluster's distance to the railway network.

[19]We use the methodology described in Conley (1999), and follow the implementation by Rappaport (2007), with a cutoff of 80 *km*.

[20]Non-migrants are defined as individuals who indicate they have never lived anywhere else beside their current place of residence.

of Table A4 suggests that the impact of the railway is attenuated when we exclude areas within 20 *km* of railway stations.

4.5 Urbanization Outcomes Results

In addition to individual and household-level outcomes, we explore urbanization outcomes, namely urban population density and city presence, using the methodology developed in Jedwab and Moradi (2015). We analyze the long-run effect on city presence and urban population (measured in 2010) of the presence of rail tracks within 20 *km* of a grid cell, controlling for the 1900 population density Z-score, missionary presence, and state fixed effects. Using the standardized score (Z-score) ensures that we measure changes relative to the mean, and controlling for 1900 Z-score ensures that we capture relative city growth. The results are presented in the first column of Table 8. We find that the presence of a rail track within 20 *km* of a cluster has a positive effect on both outcomes. Furthermore, the estimates indicate that controlling for 1960 rates of urbanization (column 4), the railway had no further impact on urbanization, suggesting that the impacts of the railway on urbanization largely occurred before independence in 1960.

5 North-South Differences in the Impact of the Rail Line

The North and the South of Nigeria had very different situations at the advent of the railway. The South had access to ports of export, thanks to its proximity to the coast and to its use of waterways, while there was no viable transportation alternative from the Northern areas to access the coast. In addition to having alternative access routes to the European market, the South had established and operated trade routes prior to the construction of the railway, thanks to centuries of pre-railway European trade (Anene, 1966; Crowder, 1980; Falola and Heaton, 2008).[21] These pre-railway differences motivate our exploration of the impact of the railway in each of these regions, separately.

5.1 Estimated Impact of the Railway in the North and in the South

Table 9 presents estimates of the impact of connectedness to the railway line on contemporary individuals living in the North and in the South of Nigeria. The top panel shows the state fixed effects results for both the North and the South, and the bottom panel

---

[21]This is reflected in trade statistics. Between 1900-1904, the South was already exporting an annual average of 176,511 tons of palm produce to Europe, while the North was exporting a modest 475 tons of its main crop: groundnuts.

shows the instrumental variable results estimated by two-stage least squares (2SLS). Both panels suggest that the local impact of the railway in the North is larger than the country-wide average impact. In the North, living within 20 *km* of a rail line increases schooling attainment by almost 2 years on average. This is associated with a 19% increase in the probability of being literate, a 2% increase in the probability of working in a professional wage job and a lower probability of being an agricultural worker. Furthermore, being connected to the rail line in the North is positively associated with media access, higher household wealth, and a 24% greater probability of residing in an urban area. The railway has virtually no impact on contemporary development outcomes in the South. There is a significant impact on household wealth and on the probability of watching TV when the model is estimated using the IV strategy, but this result is not robust. All coefficients are economically small and insignificantly different from zero in the South when estimated using state fixed effects.

Turning to urbanization outcomes, we find the same pattern. Although overall the railways have had a sustained economic impact in the country, their effects are only visible in the North, as shown in columns 2 and 3 of Table 8. Indeed, in the North, living within 20 *km* of a rail line increased the Z-score of city presence by 0.153 in 2010, an estimate significant at the 1% level. Similarly, living within 20 *km* of a rail line increased the Z-score of urban population by 0.128 in 2010. The equivalent estimates are not significantly different from zero in the South.

5.2 Differential Impact in the North and South: Robustness Checks

We carry out several robustness checks to confirm our findings on the differential impacts of the railroads in the North and in the South of Nigeria. First, we find that the placebo results obtained as an identification check for the country-wide analysis are robust to separate estimations for the North and the South. In the top panel of Table A5, we show that the strong and positive effect of the railway in the North completely disappears once actual lines are replaced by placebo lines.

We also analyze the impact of the railway using placebo lines as the control group for the South and the North. For each region, we estimate the effect of being within 20 *km* of a railway line relative to being within 20 *km* of a placebo line. The results are presented in Table A6. The effect of the railway on each outcome is large and statistically significant in the North, but is close to zero in the South. Overall, these results indicate that the positive impact of the railroads in the North and their non-impact in the South are causal.

A possible reason for the non-impact of the railway in the Southern local areas might be that, because of the discovery of oil in the South, oil cities might have eclipsed railway cities. To test for this possibility, we restrict the sample to non-oil-producing areas of the South.[22] Table A7 presents the results of this exercise. As was found with the full sample, there is no impact of connectedness to the railway on local economic development in the South when oil-producing areas are excluded. There are no significant differences between Southern connected and unconnected areas in schooling attainment, literacy, occupational choices, media access, household wealth, and urbanization. This is true regardless of the estimation strategy, which provides evidence that our results in the South are not driven by oil cities.

Next, we examine the dynamics of the railway impact both in the North and in the South of Nigeria.

6   Dynamics and Persistence

A non-existent long-term effect of the railway in the South of Nigeria could mask a short-run effect of the railway that dwindled over time - especially after the demise of railroads - or a non-existent short-run effect that remained stable over time.

This raises the question of the stability of the impact of the railway and the comparison of its short- and long-run effects. To explore the path of the railway impacts, we need to examine economic outcomes for which information is available at different points in time. The individual and household outcomes we have examined so far do not meet this criterion but they allow us to look at how the impact of the railway varies across different cohorts. We focus on individuals who have never changed their place of residence (non-migrants) to ensure that the estimates we find reflect conditions in the locality at time of birth. We look at individuals born during the peak of the railway (1948-1975), those born between 1975 and 1984, and the youngest individuals in our dataset, born between 1985-1993. These cohorts were chosen to ensure that observations are roughly of equal number within each category. The results are shown in Table 10.

We first observe that in both parts of the country, the younger cohort (1985-1993) is generally better educated, and has higher media access than older cohorts. This is consistent with the general upward trend in education in Nigeria, with the Southern education trend having a steeper slope (Csapo, 1981; Ajayi et al., 1996; Dev et al., 2016). We find no support for the hypothesis that the impact of the railway was stronger for the

---

[22]We exclude localities in the oil-producing states which are Abia, Akwa, Ibom, Bayelsa, Cross River, Delta, Edo, Imo, Ondo, and Rivers.

older cohort, be it in the North or in the South. If anything, in the North, the impact of being connected to the railway line on schooling appears larger for the youngest cohort.

For urbanization outcomes, we use panel data from Jedwab and Moradi (2015) on city presence and urban population density in 1900, 1960, and 2010 to analyze the impact of the railway on these outcomes. This allows us to more clearly separate short-term and long-term impacts.

In a first exploration of the path of the railway effect, we look at the effect of the railway on urbanization outcomes in 2010 after controlling for the 1960 population density Z-score. The results are presented in the three rightmost columns of Table 8. Controlling for the 1960 city presence or urban population Z-score, when estimating the effects of the railway on urbanization, renders the effect either not significant or much smaller than before. We view this result as evidence that the effect of the railroad has hardly changed since independence in 1960.

We then explore shorter-run effects of the railway looking at urbanization outcomes in 1960 and find that the effect of the railway is remarkably stable over time. As shown in Table 11, living within 20 *km* of a rail line increased the Z-score of city presence by 0.105 in 1960 (column 1). Similarly, living within 20 *km* of a rail line increased the Z-score of urban population by 0.175 in 1960 (column 1, bottom panel).

Although overall, the railways have had a sustained economic impact in the country, their effects are only visible in the North, which is consistent with what we have found for individual-level outcomes. Indeed, in the North, living within 20 *km* of a rail line increased the Z-score of city presence by 0.127 in 1960, as shown in the top panel of column 2, with these estimates being statistically significant at the 1% level. Similarly, living within 20 *km* of a rail line increased the Z-score of urban population by 0.124 in 1960 (column 2, bottom panel). It follows that the stability of the railway effect in the overall sample carries through to the North sub-sample. The estimated impact in the North is robust to the exclusion of all nodes (second to last column). In the South, railways did not have any statistically significant effect on urbanization outcomes, neither in the short run nor in the long run.

The empirical evidence regarding the North of the country is consistent with a model in which, in the North, railroads encouraged the concentration of production factors in connected areas and induced a spatial equilibrium that persisted in the long run.[23] This equilibrium persistence, even after railroads had become dysfunctional, can be rationalized by railroad locations serving as a coordination mechanism for factor investments in each subsequent period.

---

[23]See Bleakley and Lin (2012) for more details on theoretical underpinnings of such path dependence.

7 Mechanisms

The North and the South of Nigeria differed in many respects before the introduction of the railway. While we cannot rule out the role of all other possible differences between the two regions, key baseline differences as far as the impact of railroads is concerned include the availability of viable transportation alternatives connecting Southern areas to the ports of export at the coast and, related to it, the early existence of a more urban spatial equilibrium in the South involving ancient cities and major trading centers. We present evidence suggesting that railroads were dispensable in the South where viable pre-existing transportation technologies enabled trade with Europe, whereas railroads played a key role in connecting the North to the coast and the European market.

7.1 Adoption Rates and Benefit of Rail by Key Regional Crops

> The Zungeru-Barijuko (Kaduna) to Baro would "traverse the greatest trade route in Nigeria, and render possible the export of cotton and other produce grown in the Nupe province and in Southern Zaria [Northern Nigeria]. **Without it cotton cannot ... be profitably exported from those districts**." — *p. 58 of Colonial Report, Northern Nigeria, 1902, as quoted in (Onyewuenyi, 1981, p.66).*

> The increase in the export of this valuable product (hides) is most gratifying, and as communication with Northern Nigeria is facilitated, **it is expected to divert the greater proportion of this trade, which at present is said to go across to Tripoli**. — *p. 22 of Colonial Report, Southern Nigeria, Lagos, 1905 as quoted in (Onyewuenyi, 1981, p.66).*[24]

The need for railroads for agricultural purposes was much greater in the North of Nigeria than in the South. The South had already been trading with Europe for centuries, through the slave trade and the palm oil trade which replaced it, while trade in most of Northern Nigeria was directed towards North Africa. The railway was crucial in diverting the trans-Saharan caravan trade to the Nigerian coast by lowering transportation costs. To see this, we compute the net benefit of shipping agricultural goods by railway relative to other means such as waterways and roads, which were the two other modes of transportation available in colonial Nigeria. The calculation is done over the period 1945-1949 for each of the key regional crops - groundnuts and cotton in the North and palm oil and cocoa in the South. The results are shown in Table 12. Data on prices, quantities, and distances come from World Bank (1955) and Onyewuenyi (1981).

---

[24]The emphasis highlighted in bold is ours.

The cost of river shipments from the North is estimated as the cost of railing to Baro and then shipping by river to the Delta ports.[25] We use the railing distance as the shipping distance for rivers in the South, although this might be an overestimate given the proximity of the South to several rivers which lead to the coast.[26] While an overestimate of shipping distance through rivers, it helps to illustrate the fact that the railway could not compete with pre-existing means of transportation in the South even with implicit and explicit government subsidies.[27]

We estimate cost reduction from shipping groundnuts and cotton by railway rather than by river to be 1.4% and 51%, respectively.[28] The equivalent cost reduction from railing these goods instead of shipping them by road was 65% and 75%, respectively. These estimates are similar to the estimated reduction in Hodder (1959), who finds that the railway reduced the cost of shipping from the Jos mines by about 70% relative to road transportation. In comparison, however, railing palm oil and cocoa instead of shipping them by river would have *increased* their cost by 119% and 16%, respectively. Similarly, railing these crops versus shipping them by road would have *increased* their cost by 58% and 60%, respectively.

As regards to shipping goods to the coast, railroads were cheaper than alternative transportation modes in the North, whereas the latter were cheaper in the South. This rationalizes the high adoption rate of railroads in the North and its low adoption rate in the South. Figure 5 presents percentages of Northern and Southern goods shipped through the railway over the 1931-1949 period.[29] On average, 96% of groundnuts and 81% of cotton were railed to the coast from the North, compared to only 18% of palm kernel, 31% of palm oil, and 26% of cocoa, the three main Southern crops. We also observe that the fraction of goods shipped to the coast through rail increased over time for Northern goods, while it declined for Southern goods as subsidies for the railways declined (except palm oil).[30]

---

[25]Hence, the rail prices in parentheses in Table 12.

[26]Using average straight line distances from survey clusters in the DHS to various transportation nodes, we calculate that Southern populated areas are four times closer to rivers than Northern populated areas (23 km vs. 90 km). This discrepancy is worsened if we take into account the navigability of rivers which is much better in the South.

[27]See discussion in Onyewuenyi 1981, p. 89–93.

[28]These figures underestimate the cost reduction brought about by railways in the North as we are not taking account the navigability of the rivers in the North with was poor and impossible for several months in a year.

[29]Figure 5 plots the ratio of quantities of goods railed to the quantities of goods exported. The outlier above 100% is due to crop spoliage.

[30]The cocoa example is striking since the percentage shipped through rail shrank from a third of the production to virtually zero over the period 1931-1950.

Consistent with this, the volume of groundnut exports (the main export in the North) grew at an annual rate of 13.8%, from 475 metric tons in 1900-1904 to 268,409 metric tons in 1945-1949. Over the 1900-1949 period, during which the railway expanded, exports of palm produce (the main export in the South) grew at an average annual rate of 1.9% only. A non-trivial volume of palm produce was already being exported in the early twentieth century, which illustrates initial access to export markets.

The evidence presented here illustrates two important facts. First, producers in the South had viable transportation alternatives to the railway, and the railway did not substantially lower transportation costs in the region (it actually increased transportation costs, on average). Second, and as a result of the first fact above, adoption rates of the railway were substantially lower in the South. In the North, however, the railway substantially lowered transportation costs and increased market access and producers adopted it as the primary means of exporting commodities.

7.2 Key Factor of Heterogeneity: Distance to Ports of Export

The highlighted heterogeneity of the impact of the railway line, presented so far as a North/South dichotomy can be accounted for by distances to ports of exports. Southern areas happened to be closer to ports of export and on these short distances, waterways were a viable transportation means to ship goods to the coast. The railway did little to decrease these shipping costs and actually increased them in most areas closer to ports of export. Areas further away from the coast experienced a big drop in their shipping costs to the coast at the advent of the railway. This effectively allowed the export of goods from areas further away from ports of export to flourish.

To confirm that distance to ports is a key factor of heterogeneity, we measure the impact of the railway by distance to ports of export. For each individual, we compute the distance of her survey cluster to the closest port. We then split the sample into two subsamples: individuals with above-median distance to ports and individuals with below-median distance to ports. We estimate the impact of railroads on the outcome variables for each subsample, separately. The results are presented in Table 13.[31]

The effect of the railway is generally larger in local areas that were further away from ports of export. The effect of the railway for individuals in local areas that had higher pre-railway European market access is economically small and mostly not statistically

---

[31]We also divided the sample by mean distance to ports, and into various groups, within 100 *km*, 100-300 *km*, and other distances that included a balanced mix of districts in the South and the North. We consistently find that the impacts of the railway begin to emerge beyond 300 *km* from the coast. These results are omitted for brevity but are available upon request.

different from zero. Importantly, state fixed effects and ethnicity fixed effects are included in the regressions presented in Table 13. Hence, these findings are not driven by specific characteristics of a given state or by ancestral exposure to railways.

In the bottom panel of Table 13, we interact our main independent variable with the log of distance to coastal ports. The local effects of railways are unequivocally stronger, the farther individuals are from ports. These findings provide a consistent explanation of the non-impact of the railway in the South, based on its geographical proximity to ports, existing waterways connecting it to the different ports and a low change in its market access.

We also show that these results cannot be explained away by proximity to early cities. If the heterogeneity of the effect of the railway that we uncover is driven by the fact that the South had more early cities, one would expect that once we interact proximity to the railway with proximity to early cities, the effect on areas close to both the railway and early cities would be null. Table 14 shows that this is not the case. For most development outcomes, the effect on areas close to both the railway and early cities is significantly positive. This suggests that the heterogeneity we highlight is not driven by the stronger presence of early cities in Southern parts of Nigeria, or by the fact that Southern Nigeria was more urbanized. This finding, along with the other pieces of evidence shown above, suggests that opportunity costs are important to the transformative power of railroads. Unlike in the South, railroads were vital in the North of Nigeria to enable export trade with Europe and unlike in the South, they had a tremendous effect in the region.

8 Concluding Remarks

Studying heterogeneity in the impact of a technology is a first step towards understanding conditions under which this technology is important for economic growth. In this article, we show that colonial railroads did not have a homogeneous impact in all the areas they connected. Railroads had very little economic impact in the South of Nigeria which, thanks to its proximity to the coast and to its waterways, already had viable alternatives for the transportation of goods to exporting ports. The North, however, was lacking viable transportation technologies to export goods. The railways were essential in linking this region to ports of export. We find that connected areas in the North were transformed by the railway, not only in the short run but also over time and until today.

Our analysis shows that, in the North, individuals in areas that were connected to the railways are more likely to go to school, to be literate, to have media access, to work in higher-paying professions and to live in wealthier households than individuals in

unconnected areas. Connected areas are also more urbanized than unconnected areas. We do not find any of these effects in the South, neither in the short run nor in the long run.

Exploring potential mechanisms, we find that areas that had better access to ports of export thanks to their proximity to the coast and to their use of waterways barely adopted the railway as it did not reduce their shipping costs. The converse is true for areas that were effectively connected to exporting ports thanks to the railway. For these areas, we find that the railway significantly reduced the shipping costs for their main agricultural goods to much lower levels than with roads and waterways.

Our findings also indicate strong path dependence in the positive effect of the railway in the North of Nigeria and in its non-effect in the South. They are consistent with the theoretical argument that, in the North, the railway led to the concentration of production factors in connected areas. These initially advantaged areas helped to coordinate factor investments in each subsequent period. This implied a persistence of the initial spatial equilibrium induced by the railway, even after it became obsolete. In the South, the railway did not even change the existing spatial equilibrium in the short run. It is perhaps not surprising that its non-impact in this region also persisted over time.

References

Ajayi, J.F.A., L.K.H. Goma, and A.G. Johnson, *The African Experience with Higher Education*, Association of African Universities, 1996.

Anene, J.C., *Southern Nigeria in Transition, 1885-1906: Theory and Practice in a Colonial Protectorate*, Cambridge University Press, 1966.

Ayandele, E.A., *The Missionary Impact on Modern Nigeria, 1842-1914: A Political and Social Analysis* Ibadan History Series, Longmans, 1966.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian, "On the Road: Access to Transportation Infrastructure and Economic Growth in China," NBER Working Papers 17897, National Bureau of Economic Research, Inc Mar 2012.

Baum-Snow, Nathaniel, Loren Brandt, J Vernon Henderson, Matthew A Turner, and Qinghua Zhang, "Roads, Railroads and Decentralization of Chinese cities," Technical Report, Working paper 2012.

Bleakley, Hoyt and Jeffrey Lin, "Portage and Path Dependence," *The Quarterly Journal of Economics*, 2012, *127* (2), 587.

Cagé, Julia and Valeria Rueda, "The Long-Term Effects of the Printing Press in Sub-Saharan Africa," *American Economic Journal: Applied Economics*, July 2016, *8* (3), 69–99.

Conley, T.G., "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

Crowder, M., *The Story of Nigeria* Faber paperbacks, Faber & Faber, 1980.

Csapo, Marg, "Religious, Social and Economic Factors Hindering the Education of Girls in Northern Nigeria," *Comparative Education*, 1981, *17* (3), 311–319.

Dev, Pritha, Blessing U. Mberu, and Roland Pongou, "Ethnic Inequality: Theory and Evidence from Formal Education in Nigeria," *Economic Development and Cultural Change*, 2016, *64* (4), 603–660.

DMA, Defense Mapping Agency, *Digital Chart of the World*, Fairfax, Virginia: Defense Mapping Agency, 1992. Available at http://www.diva-gis.org/gdata.

Donaldson, Dave, "Railroads of the Raj: Estimating the impact of transportation infrastructure," *American Economic Review*, 2016, *Forthcoming.*

$\rule{1cm}{0.4pt}$ **and Richard Hornbeck**, "Railroads and American Economic Growth: A Market Access Approach," *The Quarterly Journal of Economics*, 2016.

Faber, Benjamin, "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *Review of Economic Studies*, 2014, *81* (3), 1046–1070.

Falola, T. and M.M. Heaton, *A History of Nigeria*, Cambridge University Press, 2008.

FAO, "Nigeria At a Glance," 2016. Available at http://www.fao.org/nigeria/fao-in-nigeria/nigeria-at-a-glance/en/, Accessed: February, 2016.

Fay, Marianne and Charlotte Opal, *Urbanization Without Growth: A Not So Uncommon Phenomenon* Policy Research Working Papers, World Bank, Policy Research Dissemination Center, 2000.

Fischer, G., F. Nachtergaele, S. Prieler, H.T. van Velthuizen, L. Verelst, and D. Wiberg, *Global Agro-Ecological Zones Assessment for Agriculture (GAEZ 2008)*, IIASA, Laxenburg, Austria and FAO, Rome, Italy, 2008.

Fogel, R.W., *Railroads and American Economic Growth: Essays in Econometric History*, Baltimore: Johns Hopkins Press, 1964.

Fourie, Johan and Alfonso Herranz-Loncan, "Growth (and Segregation) by Rail : How the Railways Shaped Colonial South Africa," 2015.

Gallego, Francisco A. and Robert Woodberry, "Christian Missionaries and Education in Former African Colonies: How Competition Mattered," *Journal of African Economies*, 2010, *19* (3), 294–329.

Hijmans, R. J., M. Cruz, E. Rojas, and L. Guarino, "DIVA-GIS, version 1.4. A Geographic Information System for the Management and Analysis of Genetic Resources Data.," http://www.diva-gis.org/gdata 2001.

Hijmans, Robert J., Susan E. Cameron, Juan L. Parra, Peter G. Jones, and Andy Jarvis, "Very High Resolution Interpolated Climate Surfaces for Global Land Areas," *International Journal of Climatology*, 2005, *25* (15), 1965–1978.

Hodder, B. W., "Tin Mining on the Jos Plateau of Nigeria," *Economic Geography*, 1959, *35* (2), 109–122.

Huillery, Elise, "History Matters: The Long-Term Impact of Colonial Public Investments in French West Africa," *American Economic Journal: Applied Economics*, 2009, pp. 176–215.

Jaekel, Francis, *The History of the Nigerian Railway: Opening the Nation to Sea, Air and Road Transportation*, Vol. 1, Spectrum Books, 1997.

Jedwab, Remi and Alexander Moradi, "The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa," CEH Discussion Papers 031, Centre for Economic History, Research School of Economics, Australian National University Jan 2015.

_ **and Dietrich Vollrath**, "Urbanization Without Growth in Historical Perspective," *Explorations in Economic History*, 2015, *58*, 1–21.

___ , **Edward Kerby, and Alexander Moradi**, "History, Path Dependence and Development: Evidence from Colonial Railroads, Settlers and Cities in Kenya," *The Economic Journal*, 2015, p. Forthcoming.

Lamb, P. H., "Past, Present and Future of Cotton-Growing in Nigeria," *Empire Cotton-Growing Review*, 1925, pp. 18–23.

Law, Robin, *From Slave Trade to 'Legitimate' Commerce: The Commercial Transition in Nineteenth-Century West Africa* African Studies, Cambridge, U.K.: Cambridge University Press, 2002.

NBS and ICF International, *2008 Nigeria Demographic and Health Survey [Data File]*, Vol. NGIR52FL.DTA and NGMR52FL.DTA, Calverton, Maryland: ICF International: Nigeria Bureau of Statistics and ICF International, 2008. Available at `http://dhsprogram.com/data/dataset/Nigeria_Standard-DHS_2008.cfm?flag=1`.

Nunn, Nathan, "Gender and Missionary Influence in Colonial Africa," in Emmanuel Akyeampong, Robert H. Bates, Nathan Nunn, and James Robinson, eds., *African Development in Historical Perspective*, Cambridge: Cambridge University Press, May 2014.

Okoye, Dozie and Roland Pongou, "Historical Missionary Activity, Schooling, and the Reversal of Fortunes: Evidence from Nigeria," MPRA Paper 58052, University Library of Munich, Germany Aug 2014.

Onyewuenyi, Remy N, *Railway Development and the Growth of Export Agriculture in Nigeria during the 1900-1950 Period.*, University of Ottawa (Canada)., 1981.

Rappaport, Jordan, "Moving to Nice Weather," *Regional Science and Urban Economics*, 2007, *37* (3), 375–398.

Roome, William R.M., *Ethnographic Survey of Africa Showing the Tribes and Languages: Also the Stations of Missionary Societies*, Edward Stanford Ltd, 1925.

Schätzl, Ludwig, *Industrialization in Nigeria: A Spatial Analysis*, Vol. 81, Weltforum Verlag München, 1973.

Storeygard, Adam, "Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa," *The Review of Economic Studies*, 2016, *83* (3), 1263–1295.

Taaffe, Edward J, Richard L Morrill, and Peter R Gould, "Transport Expansion in Underdeveloped Countries: A Comparative Analysis," *Geographical review*, 1963, *53* (4), 503–529.

UC Berkeley, Museum of Vertebrate Zoology, *GADM Database of Global Administrative Areas, Version 2.7*, University of California, Berkeley, 2014. Available at `www.gadm.org`.

Wantchekon, Leonard, Marko Klašnja, and Natalija Novta, "Education and Human Capital Externalities: Evidence from Colonial Benin," *The Quarterly Journal of Economics*, 2015, *forthcoming*.

World Bank, "The Economic Development of Nigeria," Technical Report 11151, The World Bank, Baltimore, MD 1955.

Figures and Tables

Figure 1: Rail Lines Across Clusters and Local Areas
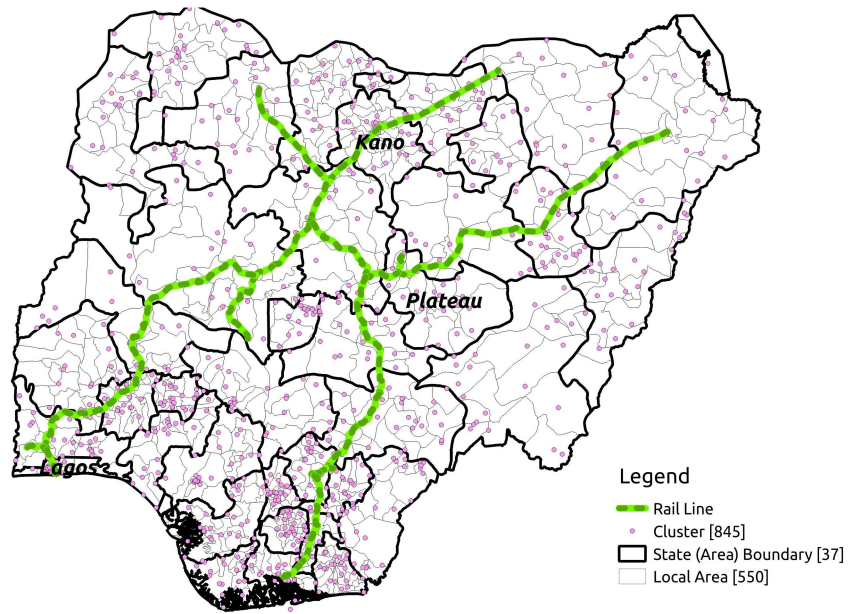
Figure 2: Straight Lines Between Nodes
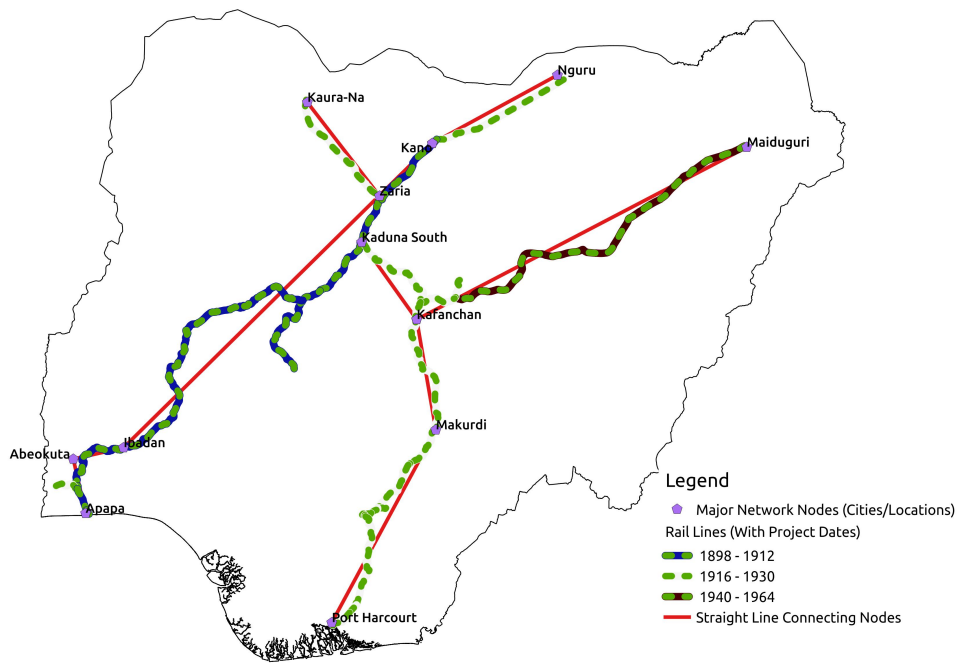
Figure 3: Rail Lines, Ports and Placebo Lines
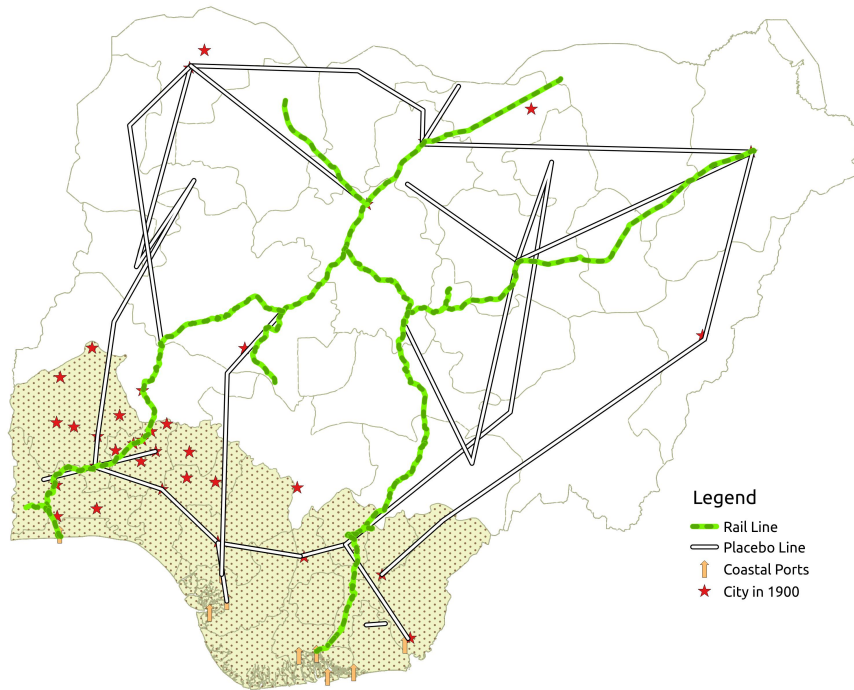
Figure 4: Rail Lines, Roads and Placebo Lines
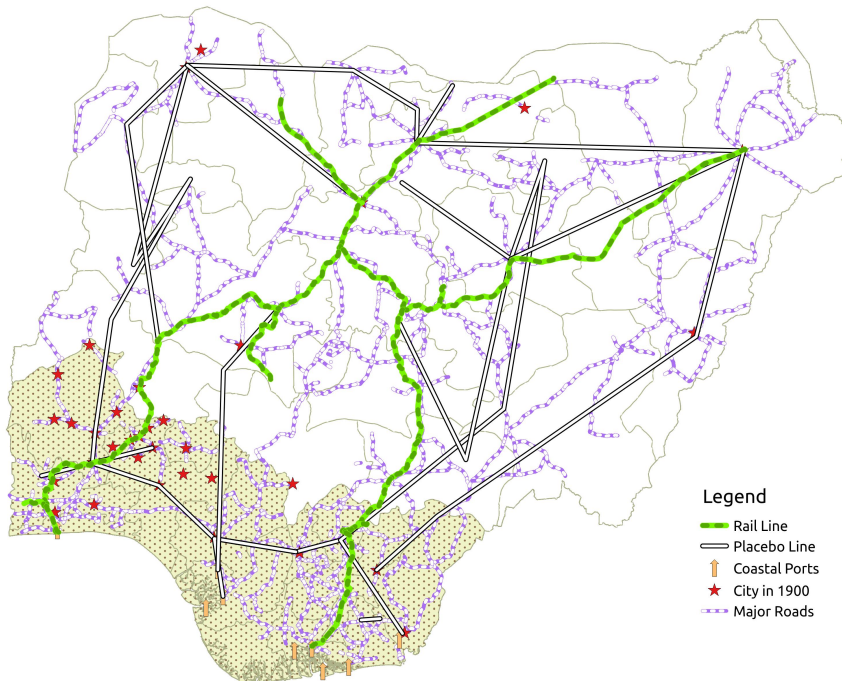
Figure 5: Railway Adoption for Northern and Southern Main Exports
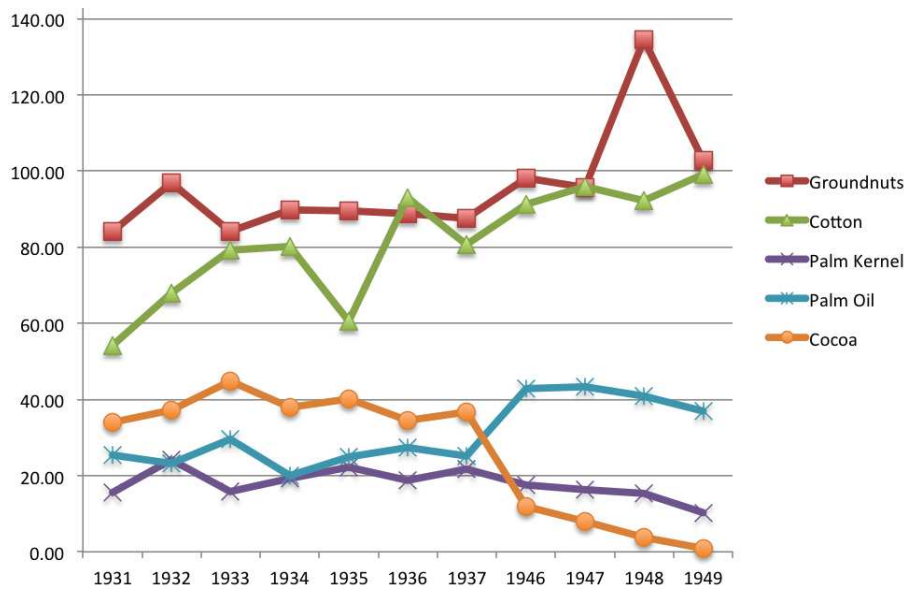
Note: Figure 5 plots the ratio of quantities of goods railed to the quantities of goods exported. The outlier above 100% is due to crop spoilage. During the 1931-1949 period, groundnuts and cotton were the main Northern exports and palm Kernel, palm oil and cocoa were the main Southern exports.

Table 1: History of Railway Construction in Nigeria

Link	Date	Length(km)	Motivation
Lagos - Otta	1898	32	Administrative & Agricultural
Otta - Abeokuta - Ibadan	1901	165	Administrative & Agricultural
Ibadan - Ilorin	1908	201	Administrative & Agricultural
Ilorin - Jebba	1909	96	Administrative & Agricultural
Zaria - Jos - Bukuru	1911	227	Mineral
Jebba - Zungeru - Minna	1912	233	Administrative & Agricultural
Baro - Kano	1912	573	Administrative & Agricultural
Port Harcourt - Enugu	1916	243	Agricultural & Mineral
Enugu - Makurdi - Jos	1927	596	Agricultural & Mineral
Kaduna - Kafanchan	1927	201	Administrative & Mineral
Zaria - Gusau - Kaura Namoda	1929	232	Agricultural
Kano - Nguru	1930	229	Agricultural
Ifo - Ilaro - Idogo	1930	39	Agricultural
Jos - Maiduguri	1964	645	Agricultural

Note: The motivations for the railway construction are classified in three categories : administrative (political or military), agricultural exploitation, and mineral exploitation.

Source: Onyewuenyi 1981, p. 39.

Table 2: Observables in Areas Within and Outside 20 *km* of Railway Tracks

	Within 20 *km* of Rail				Outside 20 *km* of Rail				Difference	
	Mean	SD	North	South	Mean	SD	North	South	Mean	SD
Rainfall (m)	1.47	.59	.91	1.79	1.5	.71	.94	2.02	-.03	.08
Temperature	26.36	.62	26.03	26.54	26.38	.84	26.4	26.35	-.02	.08
Nutrient Retention	1.53	.69	1.61	1.48	1.38	.64	1.45	1.31	.15	.1
Workability	1.99	.86	1.49	2.28	1.75	.77	1.44	2.04	.24	.11
Altitude (km)	.22	.21	.45	.08	.25	.19	.38	.13	-.03	.03
Oil Palm Suit.	2.86	2.05	.35	4.29	2.44	2.08	.5	4.2	.42	.24
Cocoa Suit.	.67	.47	.09	1	.59	.49	.14	1	.08	.06
Cotton Suit.	.35	.48	.97	0	.43	.49	.85	.04	-.08	.06
Groundnut Suit.	1.82	.39	2	1.72	1.8	.4	2	1.62	.02	.05
Mission Station	.25	.43	.11	.32	.21	.4	.1	.3	.04	.06
Road in Area	.37	.48	.73	.17	.44	.5	.51	.38	-.06	.06
River in Area	.37	.48	.27	.43	.33	.47	.26	.39	.05	.06
City Presence 1900	.07	.26	.08	.07	.07	.26	.02	.12	0	.03
City Pop. 1900 (K)	4.26	19.23	2.07	5.51	1.8	7.74	.27	3.18	2.46	2.08
Male Head	.84	.36	.91	.8	.84	.36	.93	.77	0	.01
Age of Head	44.2	13.44	44.16	44.21	44.14	14.26	42.42	45.71	.05	.4
Household Size	5.83	3.8	7.35	4.96	5.91	3.5	6.68	5.22	-.09	.19
Age	29.65	10.16	29.48	29.75	29.72	10.43	29.79	29.66	-.07	.18
Christian	.52	.5	.13	.75	.54	.5	.19	.86	-.02	.05
Number of Observations	10386	10386	4617	5769	28980	28980	15312	13668	39366	39366
Number of Clusters	208	208	94	114	637	637	343	294	845	845
Number of Local Areas	134	134	66	69	436	436	243	193	550	550

Note: Summary statistics of local government areas for clusters within (columns 1-4) and outside (columns 5-8) 20 *km* of a rail track. The two rightmost columns show differences in means between connected and unconnected areas and the associated standard errors clustered at the local government area level. For categorical variables, 1 = *Yes* and 0 = *No*, so the mean represents the proportion in each area. Soil suitability variables (for oil palm, cocoa, cotton and groundnuts) are measured in kg/ha.

Table 3: Effect of Proximity to Railway on Contemporary Outcomes (State Fixed Effects)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	1.371***	0.136***	0.018***	-0.071***	0.082***	0.068***	0.165***	0.649***	0.185***
	[0.259]	[0.023]	[0.006]	[0.016]	[0.016]	[0.014]	[0.028]	[0.101]	[0.050]
Rainfall (m)	-0.418	-0.000	-0.008	-0.009	-0.024	-0.004	-0.010	-0.098	0.064
	[0.439]	[0.037]	[0.010]	[0.037]	[0.041]	[0.031]	[0.048]	[0.207]	[0.097]
Temperature	-0.673**	-0.028	-0.005	0.051**	-0.035*	-0.005	-0.032	-0.227*	-0.000
	[0.298]	[0.027]	[0.008]	[0.022]	[0.020]	[0.024]	[0.033]	[0.130]	[0.065]
Nutrient Retention	-0.278*	-0.022*	-0.002	0.013	-0.011	-0.003	-0.019	-0.115*	-0.064**
	[0.148]	[0.012]	[0.004]	[0.012]	[0.011]	[0.010]	[0.015]	[0.063]	[0.031]
Workability	0.298**	0.027**	0.003	-0.008	0.025**	0.010	0.021	0.041	0.034
	[0.135]	[0.012]	[0.003]	[0.010]	[0.010]	[0.009]	[0.014]	[0.056]	[0.028]
Altitude (km)	-0.043	0.151	0.052	0.124	0.048	0.306**	0.172	0.022	0.320
	[1.709]	[0.156]	[0.045]	[0.128]	[0.120]	[0.131]	[0.184]	[0.721]	[0.384]
Oil Palm Suit.	0.325	0.030*	0.006	-0.040*	0.031**	0.028**	0.041**	0.130*	-0.016
	[0.225]	[0.017]	[0.006]	[0.021]	[0.015]	[0.011]	[0.017]	[0.074]	[0.036]
Cocoa Suit.	-1.061	-0.090	-0.022	0.126*	-0.092	-0.060	-0.142**	-0.467*	-0.113
	[0.798]	[0.063]	[0.024]	[0.066]	[0.061]	[0.043]	[0.061]	[0.250]	[0.149]
Cotton Suit.	-0.814	-0.108**	-0.011	0.090**	-0.013	-0.039	-0.042	-0.288*	-0.116
	[0.526]	[0.046]	[0.017]	[0.046]	[0.036]	[0.032]	[0.053]	[0.171]	[0.121]
Groundnut Suit.	-0.392	-0.031	-0.001	0.062*	-0.055	-0.023	-0.078**	-0.260*	-0.226**
	[0.299]	[0.027]	[0.010]	[0.034]	[0.034]	[0.024]	[0.032]	[0.157]	[0.088]
Mission Station	1.079***	0.075***	0.023***	-0.098***	0.085***	0.038***	0.116***	0.472***	0.217***
	[0.189]	[0.015]	[0.006]	[0.017]	[0.016]	[0.011]	[0.020]	[0.082]	[0.044]
Observations	39028	38706	38882	38882	38859	38962	38959	39063	39063
Adjusted R-sq.	0.47	0.39	0.05	0.20	0.20	0.15	0.34	0.49	0.29
Control Means	6.014	0.547	0.045	0.240	0.266	0.737	0.505	2.868	0.276

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. Table shows estimates of the impact of being within 20 *km* of a railway line on various individual outcomes. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network are computed using DHS data and GIS information on the rail network. Climatic and geographic controls are measured as the average within the local government area (county). Data on Christian mission stations comes from historical maps, as described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table 4: First-Stage Estimates based on Distance to Line Joining Major Nodes

	(1) Rail Within 20 *km*
ln(Dist. Straight Lines)	-0.294***
	[0.021]
Rainfall (m)	0.034
	[0.066]
Temperature	-0.043
	[0.039]
Nutrient Retention	0.000
	[0.023]
Workability	0.018
	[0.019]
Altitude (km)	-0.439*
	[0.253]
Oil Palm Suit.	0.002
	[0.021]
Cocoa Suit.	0.103
	[0.111]
Cotton Suit.	0.114
	[0.079]
Groundnut Suit.	0.014
	[0.028]
Mission Station	0.019
	[0.026]
Observations	37353

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. The instrument is the log of distance of DHS clusters to straight lines joining railway nodes of historical importance. All nodes of the railway line, regardless of historical importance, are dropped from regression samples. The first stage regression includes ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network are computed using DHS data and GIS information on the rail network. Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations comes from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table 5: Effect of Proximity to Railway on Contemporary Outcomes (2SLS)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	1.508***	0.142***	0.014*	-0.040*	0.088***	0.072***	0.181***	0.642***	0.115
	[0.393]	[0.034]	[0.008]	[0.024]	[0.026]	[0.023]	[0.042]	[0.159]	[0.078]
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	37319	37018	37180	37180	37157	37259	37250	37353	37353
Centered R-sq.	0.47	0.39	0.06	0.20	0.20	0.15	0.34	0.49	0.29
Control Means	6.014	0.547	0.045	0.240	0.266	0.737	0.505	2.868	0.276

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. The instrument is the log of distance of DHS clusters to straight lines joining railway nodes of historical importance. All nodes of the railway line, regardless of historical importance, are dropped from regression samples. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network are computed using DHS data and GIS information on the rail network. Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations comes from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table 6: Falsification Exercises: Other Transportation Means and Placebo Lines

Panel A: Other Modes of Transportation

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	0.863*** [0.222]	0.093*** [0.020]	0.011* [0.006]	-0.034** [0.015]	0.046*** [0.014]	0.040*** [0.013]	0.106*** [0.024]	0.391*** [0.084]	0.093** [0.047]
ln(Dist. River)	-0.052 [0.070]	-0.006 [0.006]	-0.002 [0.002]	0.005 [0.006]	-0.007 [0.006]	-0.004 [0.005]	-0.019** [0.007]	-0.036 [0.027]	-0.002 [0.019]
ln(Dist. Road)	-0.745*** [0.076]	-0.062*** [0.007]	-0.010*** [0.002]	0.055*** [0.006]	-0.051*** [0.006]	-0.041*** [0.005]	-0.083*** [0.009]	-0.376*** [0.030]	-0.136*** [0.017]
Observations	39028	38706	38882	38882	38859	38962	38959	39063	39063

Panel B: Placebo Lines

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Placebo Within 20 *km*	0.269 [0.179]	0.020 [0.015]	0.009 [0.006]	-0.029* [0.015]	0.022 [0.015]	0.035*** [0.013]	0.039* [0.021]	0.240*** [0.085]	0.061 [0.041]
Observations	28663	28454	28568	28568	28543	28606	28598	28687	28687

Panel C: Placebo Lines and Other Modes of Transportation

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Placebo Within 20 *km*	0.110 [0.173]	0.008 [0.014]	0.006 [0.005]	-0.015 [0.015]	0.010 [0.014]	0.025** [0.012]	0.021 [0.019]	0.151** [0.075]	0.030 [0.039]
ln(Dist. River)	0.001 [0.075]	-0.003 [0.006]	-0.000 [0.002]	-0.001 [0.007]	-0.002 [0.007]	-0.001 [0.005]	-0.019** [0.008]	-0.031 [0.030]	0.012 [0.021]
ln(Dist. Road)	-0.580*** [0.076]	-0.045*** [0.006]	-0.009*** [0.002]	0.050*** [0.007]	-0.044*** [0.007]	-0.035*** [0.005]	-0.064*** [0.009]	-0.321*** [0.033]	-0.115*** [0.019]
Observations	28663	28454	28568	28568	28543	28606	28598	28687	28687

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network, rivers and roads are computed using DHS data and GIS information on rail, river and road networks. Distances to placebo lines are computed using GIS information on locations of placebo lines (Jaekel, 1997). Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations comes from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table 7: Robustness Checks : Various Control Groups

Panel A: Placebo Lines as Control Group

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	1.036***	0.135***	0.007	-0.045*	0.065***	0.045**	0.138***	0.510***	0.142**
	[0.329]	[0.029]	[0.011]	[0.024]	[0.022]	[0.018]	[0.041]	[0.143]	[0.070]
Observations	19644	19462	19561	19561	19547	19611	19620	19659	19659

Panel B: Across Multiple Distances

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	1.231***	0.125***	0.022***	-0.062***	0.091***	0.076***	0.179***	0.687***	0.187***
	[0.314]	[0.028]	[0.008]	[0.021]	[0.021]	[0.020]	[0.034]	[0.128]	[0.062]
Rail Within 20-40 *km*	-0.362	-0.029	0.003	0.018	0.010	0.003	0.003	0.008	-0.068
	[0.245]	[0.021]	[0.007]	[0.021]	[0.019]	[0.019]	[0.027]	[0.107]	[0.053]
Rail Within 40-60 *km*	0.008	-0.001	0.006	-0.009	0.009	0.014	0.049*	0.120	0.064
	[0.232]	[0.019]	[0.007]	[0.022]	[0.019]	[0.017]	[0.029]	[0.108]	[0.057]
Rail Within 60-80 *km*	0.020	0.007	0.009	0.027	0.016	0.018	-0.005	0.008	0.069
	[0.218]	[0.016]	[0.009]	[0.019]	[0.017]	[0.014]	[0.021]	[0.092]	[0.051]
Observations	39028	38706	38882	38882	38859	38962	38959	39063	39063

Note: *$p < .1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the local government area level in brackets. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network and to placebo lines are computed using DHS data and GIS information on rail network and locations of placebo lines (Jaekel, 1997). Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations comes from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table 8: Long-Run Effects of Railway on Urbanization Outcomes

Panel A: Dependent Variable: Z-score of City Presence in 2010

	(1) All	(2) North	(3) South	(4) All	(5) North	(6) South
Rail Within 20 *km*	0.108** [0.051]	0.153*** [0.053]	-0.056 [0.139]	0.060 [0.043]	0.094** [0.045]	-0.061 [0.109]
1900 City Z-score	0.181*** [0.013]	0.201*** [0.016]	0.171*** [0.019]			
1960 City Z-score				0.454*** [0.009]	0.471*** [0.009]	0.430*** [0.015]
Observations	7510	5985	1525	7510	5985	1525

Panel B: Dependent Variable: Z-score of Urban Population in 2010

	(1) All	(2) North	(3) South	(4) All	(5) North	(6) South
Rail Within 20 *km*	0.180*** [0.067]	0.128** [0.054]	0.136 [0.189]	0.016 [0.020]	0.013 [0.013]	0.001 [0.086]
1900 Pop. Z-score	0.404*** [0.121]	1.304*** [0.400]	0.244*** [0.052]			
1960 Pop. Z-score				0.907*** [0.055]	0.945*** [0.070]	0.832*** [0.117]
Observations	7510	5985	1525	7510	5985	1525

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. Table estimates the impact of proximity to the railway line on urbanization in 2010 (measured as city presence and urban population) within $10km \times 10km$ local grid cells. The Z-score is the standardized score of the variable of interest, computed as the difference from the mean divided by the standard deviation. We control for the presence of mission stations within the grid cell and update the measure of rail connectedness to reflect the line completed after 1960. All regressions include state of residence fixed effects as well as baseline controls. Climatic and geographic controls are measured as the average within the grid cell.

Table 9: Effect of Proximity to Railway in North and South of Nigeria

Panel A: North (Estimation Strategy : State Fixed Effects)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 km	1.928***	0.192***	0.023***	-0.084***	0.107***	0.106***	0.224***	0.840***	0.236***
	[0.349]	[0.033]	[0.006]	[0.019]	[0.020]	[0.020]	[0.037]	[0.126]	[0.059]
Observations	19683	19514	19617	19617	19572	19652	19641	19713	19713

Panel B: South (Estimation Strategy : State Fixed Effects)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 km	0.230	0.033	0.005	-0.029	0.017	0.000	0.049	0.233	0.068
	[0.338]	[0.023]	[0.013]	[0.030]	[0.027]	[0.016]	[0.041]	[0.158]	[0.084]
Observations	19344	19191	19264	19264	19286	19309	19317	19349	19349

Panel C: North (Estimation Strategy : IV)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 km	2.267***	0.210***	0.013	-0.050*	0.111***	0.116***	0.245***	0.904***	0.159*
	[0.561]	[0.050]	[0.009]	[0.029]	[0.034]	[0.032]	[0.056]	[0.207]	[0.094]
Observations	18741	18582	18682	18682	18637	18717	18699	18771	18771

Panel D: South (Estimation Strategy : IV)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 km	0.591	0.062*	0.012	-0.044	0.061	0.021	0.114*	0.370*	0.111
	[0.459]	[0.034]	[0.017]	[0.044]	[0.041]	[0.028]	[0.059]	[0.221]	[0.121]
Observations	18578	18436	18498	18498	18520	18542	18551	18582	18582

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. The instrument is the log of distance of DHS clusters to straight lines joining railway nodes of historical importance. All nodes of the railway line, regardless of historical importance, are dropped from regression samples. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network are computed using DHS data and GIS information on the rail network. Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations comes from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table 10: Differential Impact by Cohort (Non-Migrants Only)

Panel A: North (Estimation Strategy : State Fixed Effects)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 km	1.267*** [0.388]	0.124*** [0.035]	0.016 [0.014]	-0.111*** [0.037]	0.077*** [0.026]	0.091*** [0.023]	0.131*** [0.038]	0.629*** [0.123]	0.202*** [0.061]
X 1975-1984	0.572 [0.374]	0.042 [0.033]	0.015 [0.015]	0.013 [0.031]	0.015 [0.027]	-0.050* [0.027]	0.056** [0.027]	0.038 [0.053]	0.014 [0.024]
X 1985-1993	0.936*** [0.358]	0.117*** [0.034]	-0.011 [0.016]	0.061** [0.030]	0.045 [0.032]	0.029 [0.029]	0.087*** [0.030]	0.143* [0.083]	0.037 [0.029]
Born 1975-1984	0.573** [0.223]	0.023 [0.025]	-0.002 [0.008]	0.039* [0.021]	0.075*** [0.021]	0.021 [0.020]	0.052** [0.021]	0.145*** [0.052]	0.030 [0.019]
Born 1985-1993	1.714*** [0.389]	0.107** [0.044]	0.009 [0.017]	0.018 [0.034]	0.154*** [0.040]	0.042 [0.035]	0.137*** [0.034]	0.243** [0.100]	0.060 [0.040]
Observations	9391	9308	9352	9352	9347	9373	9370	9406	9406

Panel B: South (Estimation Strategy : State Fixed Effects)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 km	-0.263 [0.492]	0.026 [0.043]	-0.021 [0.021]	-0.004 [0.054]	0.004 [0.038]	-0.008 [0.029]	0.000 [0.057]	0.040 [0.167]	0.019 [0.098]
X 1975-1984	0.621* [0.342]	0.012 [0.038]	0.024 [0.017]	0.000 [0.037]	0.023 [0.037]	0.020 [0.024]	0.038 [0.038]	0.192*** [0.069]	0.064** [0.030]
X 1985-1993	0.325 [0.362]	-0.035 [0.041]	0.019 [0.017]	0.029 [0.040]	0.044 [0.035]	0.003 [0.027]	0.017 [0.038]	0.158*** [0.056]	0.048* [0.026]
Born 1975-1984	2.010*** [0.254]	0.097*** [0.023]	0.011 [0.014]	-0.037 [0.024]	0.180*** [0.028]	0.061*** [0.022]	0.072*** [0.025]	0.099* [0.052]	0.013 [0.021]
Born 1985-1993	3.063*** [0.327]	0.156*** [0.033]	-0.002 [0.020]	-0.044 [0.032]	0.237*** [0.048]	0.116*** [0.035]	0.099*** [0.038]	0.174** [0.074]	0.044 [0.034]
Observations	7375	7328	7354	7354	7357	7363	7366	7377	7377

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. The table estimates, by cohort, the impact of distance to the railway on Northern (West, East, Central) and Southern (West, East, South) Nigeria. The omitted cohort is made of older individuals born between 1948-1974. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail and road networks are computed using DHS data and GIS information on rail and road networks. Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations come from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table 11: Short-Run Effects of the Railway on Urbanization Outcomes

Panel A: Dependent Variable: Z-score of City Presence in 1960

				All Nodes Excluded		
	(1) All	(2) North	(3) South	(4) All	(5) North	(6) South
Rail Within 20 *km*	0.105** [0.048]	0.127*** [0.043]	0.010 [0.159]	0.066 [0.046]	0.085** [0.040]	-0.015 [0.155]
1900 City Z-score	0.400*** [0.025]	0.412*** [0.044]	0.389*** [0.031]	0.397*** [0.026]	0.407*** [0.052]	0.385*** [0.032]
Observations	7510	5985	1525	7487	5971	1516

Panel B: Dependent Variable: Z-score of Urban Population in 1960

				All Nodes Excluded		
	(1) All	(2) North	(3) South	(4) All	(5) North	(6) South
Rail Within 20 *km*	0.175*** [0.065]	0.124** [0.053]	0.151 [0.159]	0.119* [0.062]	0.092* [0.052]	0.046 [0.112]
1900 Pop. Z-score	0.507*** [0.140]	1.346*** [0.508]	0.355*** [0.080]	0.489*** [0.142]	1.434** [0.621]	0.353*** [0.080]
Observations	7510	5985	1525	7487	5971	1516

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. Table estimates the impact of proximity to the railway line on urbanization in 1960 (measured as city presence and urban population) within $10km \times 10km$ local grid cells. The Z-score is the standardized score of the variable of interest, computed as the difference from the mean divided by the standard deviation. We control for the presence of mission stations within the grid cell. All regressions include state of residence fixed effects as well as baseline controls. Climatic and geographic controls are measured as the average within the grid cell.

Table 12: Benefits of Shipping by Rail For Key Regional Crops

	Northern Crops		Southern Crops	
	Groundnuts	**Cotton**	**Palm Oil**	**Cocoa**
Shipping Prices				
Rail Price (pence per ton km)	1.95	1.37	3.95	2.08
River Price (pence per ton km)	.9 (+ 3.1 rail)	2.5 (+ 3.1 rail)	1.8	1.8
Road Price (pence per ton km)	5.6	5.6	2.5	1.3
Shipping Distances				
Distance Rail (km)	1127	1159	61	193
Distance River (km)	575 river (552 rail)	575 river (584 rail)	61	193
Cost Reduction from Rail				
As % of River Cost	-1.4	-51.1	119.4	15.6
As % of Road Cost	-65.2	-75.5	58	60

Note: Table calculates the benefit of the railway over the period 1945-1949. For river shipments in the North, the cost is estimated as the cost of railing to Baro and then shipping by river to the Delta ports, hence the rail prices and distances in parentheses. We use the railing distance as the shipping distance for rivers in the South, although this might be an overestimate given the proximity of the South to several rivers which lead to the coast.

Table 13: Effect of Proximity to Railway By Distance to Coastal Port

Panel A: Above Median Distance to Port (192 *km*)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	1.807*** [0.326]	0.177*** [0.030]	0.022*** [0.005]	-0.073*** [0.019]	0.103*** [0.019]	0.104*** [0.019]	0.209*** [0.035]	0.779*** [0.120]	0.207*** [0.057]
Observations	22164	21987	22088	22088	22052	22129	22118	22194	22194

Panel B: Below Median Distance to Port (192 *km*)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	-0.080 [0.346]	0.009 [0.023]	-0.007 [0.015]	-0.026 [0.031]	-0.012 [0.028]	-0.017 [0.016]	0.019 [0.044]	0.161 [0.171]	0.131 [0.089]
Observations	16863	16718	16793	16793	16806	16832	16840	16868	16868

Panel C: Interaction with log(Distance to Ports)

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	-1.504* [0.885]	-0.194*** [0.067]	0.001 [0.024]	0.020 [0.067]	-0.086 [0.069]	-0.154*** [0.053]	-0.160* [0.090]	-0.604 [0.367]	0.065 [0.221]
X ln(Dist. Port)	0.494*** [0.166]	0.058*** [0.014]	0.003 [0.004]	-0.014 [0.011]	0.028** [0.012]	0.039*** [0.010]	0.056*** [0.017]	0.214*** [0.066]	0.019 [0.038]
Observations	39028	38706	38882	38882	38859	38962	38959	39063	39063

Note: Standard errors clustered at the local government area level in brackets. Table estimates the impact of distance to the railway by distance to coastal ports (Bonny, Burutu, Calabar, Degema, Lagos, Opobo, Port Harcourt, Sapele, Warri). All railway nodes are dropped from the regressions. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distance to rail network is computed using DHS data and GIS information on the rail network. Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations comes from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table 14: Effect of Railway By Proximity to Early Cities

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	1.198***	0.116***	0.019***	-0.067***	0.078***	0.067***	0.149***	0.611***	0.160***
	[0.260]	[0.024]	[0.005]	[0.017]	[0.017]	[0.014]	[0.030]	[0.107]	[0.052]
X 1900 City Within 20 *km*	-0.501	-0.021	-0.028*	0.058**	-0.066**	-0.048**	-0.067	-0.393**	-0.116
	[0.458]	[0.038]	[0.016]	[0.028]	[0.030]	[0.023]	[0.049]	[0.169]	[0.092]
1900 City Within 20 *km*	1.755***	0.138***	0.040***	-0.120***	0.128***	0.087***	0.198***	0.859***	0.329***
	[0.252]	[0.022]	[0.012]	[0.024]	[0.022]	[0.016]	[0.032]	[0.117]	[0.061]
Observations	39028	38706	38882	38882	38859	38962	38959	39063	39063

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. Table estimates the impact of the railway, by proximity to a city in 1900. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network are computed using DHS data and GIS information on the rail network. Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations comes from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Appendix: Other Results and Robustness Checks

Table A1: Effect of Railway: Robustness to other Measures of Connectedness

Panel A: Closeness to Railway Lines

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Closeness to Rail	0.593*** [0.099]	0.056*** [0.009]	0.008*** [0.002]	-0.029*** [0.006]	0.038*** [0.006]	0.029*** [0.006]	0.072*** [0.010]	0.281*** [0.037]	0.074*** [0.019]
Observations	39028	38706	38882	38882	38859	38962	38959	39063	39063

Panel B: Proximity to Railway Station

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Station Within 20 *km*	1.655*** [0.314]	0.155*** [0.028]	0.021*** [0.007]	-0.097*** [0.016]	0.100*** [0.018]	0.057*** [0.016]	0.185*** [0.033]	0.722*** [0.113]	0.232*** [0.055]
Observations	39028	38706	38882	38882	38859	38962	38959	39063	39063

Panel C: Presence of Rail Tracks in Local Area

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail in Local Area	0.696*** [0.266]	0.072*** [0.023]	0.013** [0.005]	-0.025 [0.016]	0.047*** [0.017]	0.047*** [0.014]	0.096*** [0.029]	0.383*** [0.103]	0.090* [0.049]
Observations	39028	38706	38882	38882	38859	38962	38959	39063	39063

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. Closeness to the railway line is defined as the log of the inverse of 1 plus the distance of the individual's cluster to the railway line. Rail in Local Area takes the value 1 if a railway line crosses the individual's local government area. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network are computed using DHS data and GIS information on the rail network. Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations comes from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table A2: Conley Standard Errors

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	1.371***	0.136***	0.018***	-0.071***	0.082***	0.068***	0.165***	0.649***	0.185***
	[0.283]	[0.025]	[0.005]	[0.019]	[0.017]	[0.016]	[0.029]	[0.112]	[0.050]
Observations	39028	38706	38882	38882	38859	38962	38959	39063	39063

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Table estimates the impact of proximity to the railway adjusting for spatial correlation with Conley standard errors (in brackets). Conley standard errors are computed with a cutoff of 80 *km*. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to the rail network are computed using DHS data and information on the rail network. Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations comes from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table A3: Effect of Railway: Robustness to Various Sub-samples

Panel A: Rural Areas

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 km	0.835*** [0.298]	0.085*** [0.027]	0.011** [0.005]	-0.044** [0.018]	0.052*** [0.019]	0.053*** [0.016]	0.105*** [0.032]	0.467*** [0.117]	0.000 [.]
Observations	25508	25272	25410	25410	25387	25462	25452	25523	25523

Panel B: Migrants Only

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 km	1.478*** [0.274]	0.148*** [0.025]	0.022*** [0.007]	-0.071*** [0.017]	0.082*** [0.019]	0.082*** [0.017]	0.198*** [0.030]	0.754*** [0.107]	0.202*** [0.053]
Observations	22168	21977	22082	22082	22063	22132	22129	22186	22186

Panel C: Non-Migrants Only

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 km	1.225*** [0.295]	0.123*** [0.025]	0.010 [0.007]	-0.061*** [0.022]	0.082*** [0.018]	0.055*** [0.016]	0.126*** [0.030]	0.534*** [0.103]	0.162*** [0.054]
Observations	16767	16637	16707	16707	16705	16737	16737	16784	16784

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. Table shows estimates of the impact of being within 20 km of a railway line on various individual outcomes for different sub-samples. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network are computed using DHS data and GIS information on the rail network. Climatic and geographic controls are measured as the average within the local government area (county). Data on Christian mission stations comes from historical maps, as described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table A4: Effect of Railway: Robustness to Various Sub-samples

Panel A: No Mission Stations in Local Area

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	1.457*** [0.292]	0.142*** [0.026]	0.022*** [0.006]	-0.067*** [0.017]	0.082*** [0.018]	0.063*** [0.016]	0.173*** [0.033]	0.695*** [0.118]	0.196*** [0.055]
Observations	30706	30436	30588	30588	30562	30653	30643	30733	30733

Panel B: No Rail Tracks in Local Area

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	0.961** [0.410]	0.093** [0.039]	0.018* [0.011]	-0.076*** [0.022]	0.065** [0.028]	0.026 [0.022]	0.143*** [0.042]	0.614*** [0.161]	0.223** [0.092]
Observations	30179	29958	30069	30069	30051	30126	30117	30205	30205

Panel C: No Rail Station in Local Area

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	0.545* [0.279]	0.062*** [0.023]	0.013** [0.006]	-0.018 [0.021]	0.037* [0.019]	0.059*** [0.017]	0.086*** [0.031]	0.345*** [0.115]	0.074 [0.066]
Observations	32940	32696	32827	32827	32802	32880	32877	32971	32971

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. Table shows estimates of the impact of being within 20 *km* of a railway line on various individual outcomes for different sub-samples. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network are computed using DHS data and GIS information on the rail network. Climatic and geographic controls are measured as the average within the local government area (county). Data on Christian mission stations comes from historical maps, as described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table A5: Falsification Exercise: Placebo Lines Estimates in North and South

Panel A: Placebo Lines in the North

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Placebo Within 20 *km*	-0.056 [0.230]	-0.001 [0.021]	-0.002 [0.006]	-0.003 [0.017]	-0.016 [0.015]	0.020 [0.020]	0.023 [0.028]	0.143 [0.103]	0.023 [0.050]
ln(Dist. River)	-0.047 [0.123]	-0.001 [0.011]	-0.007** [0.003]	-0.001 [0.008]	-0.013* [0.008]	0.005 [0.008]	-0.019* [0.012]	-0.042 [0.044]	-0.017 [0.027]
ln(Dist. Road)	-0.338*** [0.087]	-0.031*** [0.008]	-0.005* [0.003]	0.030*** [0.007]	-0.016*** [0.006]	-0.030*** [0.007]	-0.043*** [0.011]	-0.208*** [0.037]	-0.094*** [0.022]
Observations	15099	14978	15053	15053	15022	15072	15057	15120	15120

Panel B: Placebo Lines in the South

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Placebo Within 20 *km*	0.464* [0.237]	0.035** [0.018]	0.012 [0.009]	-0.041* [0.025]	0.040* [0.024]	0.033** [0.016]	0.039 [0.027]	0.233** [0.113]	0.071 [0.064]
ln(Dist. River)	0.031 [0.093]	-0.006 [0.007]	0.007** [0.003]	0.002 [0.011]	0.008 [0.010]	-0.010 [0.007]	-0.014 [0.010]	-0.013 [0.044]	0.043 [0.030]
ln(Dist. Road)	-0.922*** [0.106]	-0.062*** [0.008]	-0.015*** [0.004]	0.078*** [0.012]	-0.088*** [0.010]	-0.038*** [0.007]	-0.086*** [0.012]	-0.471*** [0.048]	-0.144*** [0.033]
Observations	13563	13475	13514	13514	13520	13533	13540	13566	13566

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. The table estimates the impact of being within 20 *km* to a placebo line (surveyed lines that were not constructed). All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distances to rail network, rivers and roads are computed using DHS data and GIS information on rail, river and road networks. Climatic and geographic controls are measured as the average within the local government area. Data on Christian mission stations comes from maps described in the text. All other variables are taken from the 2008 Nigeria DHS.

Table A6: Placebo Lines as Control Group in North and in South

Panel A: North

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	1.603***	0.179***	0.025***	-0.031	0.089***	0.074***	0.183***	0.644***	0.131
	[0.384]	[0.037]	[0.008]	[0.022]	[0.023]	[0.022]	[0.048]	[0.159]	[0.085]
Observations	8617	8530	8582	8582	8554	8602	8605	8628	8628

Panel B: South

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	-0.348	0.024	-0.021	-0.009	-0.002	-0.008	0.014	0.116	0.074
	[0.432]	[0.031]	[0.025]	[0.038]	[0.036]	[0.021]	[0.064]	[0.236]	[0.118]
Observations	11024	10929	10976	10976	10990	11006	11012	11028	11028

Note: *$p < .1$, ** $p < 0.05$, ***$p < 0.01$. Standard errors clustered at the local government area level in brackets. The table estimates separately for Southern and Northern Nigeria the impact of being within 20 *km* of a railway line relative to being within 20 *km* of a placebo line. All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Climatic and geographic controls are from Fischer et al. (2008) and Hijmans et al. (2005), and are measured as the average within the local government area. Data on Christian mission stations comes from maps published by Ayandele (1966) and Roome (1925). All other variables are taken from the 2008 Nigeria DHS.

Table A7: Robustness of No Effect in South Excluding Crude Oil Producers

	(1) Schooling	(2) Literacy	(3) Professional	(4) Agricultural	(5) Read Paper	(6) Listen to Radio	(7) Watch TV	(8) Wealth Index	(9) Urban
Rail Within 20 *km*	0.030 [0.392]	0.025 [0.027]	-0.004 [0.018]	-0.008 [0.033]	0.004 [0.031]	-0.008 [0.019]	0.031 [0.056]	0.178 [0.203]	0.006 [0.098]
Observations	9329	9239	9292	9292	9301	9317	9320	9333	9333

Note: *p < .1, ** p < 0.05, ***p < 0.01. Standard errors clustered at the local government area level in brackets. Table estimates the impact of proximity to the railway in Southern Nigeria (West, East, South) excluding oil-producing areas. Oil-producing areas are the historically oil-producing states (Abia, Akwa Ibom, Bayelsa, Cross River, Delta, Edo, Imo, Ondo, Rivers). All regressions include ethnicity and state of residence fixed effects as well as all baseline controls. Distance to rail network is computed using data from the from the 2008 Nigeria DHS and information on rail network taken from DMA (1992). Climatic and geographic controls are from Fischer et al. (2008) and Hijmans et al. (2005), and are measured as the average within the local government area. Data on Christian mission stations comes from maps published by Ayandele (1966) and Roome (1925). All other variables are taken from the 2008 Nigeria DHS.