Revolutionizing Transport: Modern Infrastructure, Agriculture and Development in Ghana*

Rémi Jedwab PSE and LSE

Alexander Moradi Sussex University

March 2012

Abstract

We study the impact of colonial investments in modern transportation infrastructure on agriculture and development in Ghana. Two railway lines were built between 1901 and 1923 to connect the coast to mining areas and the large hinterland city of Kumasi. This unintendedly opened vast expanses of tropical forest to cocoa cultivation, allowing Ghana to become the world's largest producer. Using data at a very fine spatial level, we find a strong effect of railroad connectivity on cocoa production in 1927, generating rents in the order of 4.5% of GDP. We show that the economic boom in cocoa-producing areas was associated with demographic growth and urbanization. We find no effect for lines that were not built yet, and lines that were planned but never built. Lastly, railway construction had a persistent impact: railway districts are more developed today despite a complete displacement of rail by other means of transport.

JEL classification codes: O18, R4, 013, F1, N17
Keywords: Transportation Infrastructure, Trade Costs, Agriculture, Africa

---

*Remi Jedwab, Paris School of Economics (e-mail: jedwab@pse.ens.fr) and STICERD, London School of Economics (e-mail: r.c.jedwab@lse.ac.uk). Alexander Moradi, University of Sussex (email: a.Moradi@sussex.ac.uk). We would like to thank Denis Cogneau, Gareth Austin, Robert Allen, Robert Eastwood, Francis Teal, James Fenske, Marcel Fafchamps, William Masters, Jean-Marie Baland, Jean-Philippe Platteau, Branko Milanovic, Elizabeth Sadoulet, Marcel Fafchamps, Christopher Adam, Paul Collier, Francois Bourguignon, Elise Huillery, Sonia Bhalotra and seminar audiences at Geneva (African Economic History Workshop), Sussex, LSE (STICERD and Economic History seminars), Oxford (CSAE conference, CSAE and Economic History seminars), Carlos Tercero and Namur for very helpful comments. We are grateful to Vincent Anchirinah, Francis Baah and Frederick Amon-Armah from CRIG for their help with data collection. We furthermore gratefully acknowledge the financial support of PSE and ESRC First Grant (RES-061-25-0456).

1 Introduction

Recent research has confirmed that trade has large positive effects on income (Acemoglu, Johnson and Robinson 2005; Feyrer 2009a,b; Donaldson 2010). Similarly, research has suggested that the lack of intercontinental and intracontinental trade integration is a determining factor of African underdevelopment (Rodrik 1998; Johnson, Ostry and Subramanian 2007; Buys, Deichmann and Wheeler 2010). In this regard, the literature often mentions the conjunction of bad geography and poor infrastructure as the main obstacle to trade expansion in Africa (Radelet and Sachs 1998; Limão and Venables 2001; Buys, Deichmann and Wheeler 2010). In 2005, sub-Saharan Africa had 0.002 km of railroad track per 1000 sq km and Ghana had 4 km, whereas China had 7 km, India 21 km, the United States 25 km and Europe 51 km per 1000 sq km (World Bank 2010). Africa's road infrastructure is in a similar state.[1] Clearly, African countries are underprovided with transport infrastructure and that is why international organizations are advocating massive investments there. The Trans-Highway network is a set of transcontinental road projects to develop road-based trade corridors in Africa (African Development Bank 2003). The African Union is also working out modalities on how to link all railways in the continent (African Union 2006). Lastly, Ghana, the country we study in this paper, has just signed a $6 billion contract to rehabilitate the existing network and build a railroad to the country's northern border. This will make it the largest rail investment in Africa since at least 50 years.

Despite this recent interest in transportation projects, little is known on their economic effects. A first strand of the literature looks at the macroeconomic impact of transport infrastructure on trade and development (Radelet and Sachs 1998; Limão and Venables 2001; Clark, Dollar and Micco 2004). Those studies find that better infrastructure diminish transport costs, especially for landlocked countries, with a positive impact on exports and economic growth. As infrastructure is endogenous to economic conditions, one cannot be sure such effects are causal. A second strand of the literature has focused on the impact of rural roads on

[1]Sub-Saharan Africa had 85 km of roads per 1000 sq km and Ghana had 239 km, while China had 201 km, India 1115 km, the United States 702 km and European countries 1377 km per 1000 sq km. Magnitudes are similar if the lengths of the rail and road networks are standardized by population instead of country area. Then, 53% of African countries have a railway.

neighboring communities (Jacoby 2000; Renkow, Hallstrom and Karanja 2004; Mu and van de Walle 2007; Dercon, Gilligan, Hoddinott and Woldehan 2008; Jacoby and Minten 2009). Those studies find that rural roads reduce poverty in connected villages, by integrating labor and goods markets, thus providing new economic opportunities to their inhabitants. For instance, farmers obtain higher profits thanks to cheaper agricultural inputs and higher farmgate prices for their crops. Yet, this literature also faces identification issues.[2] A last strand studies the impact of large transportation projects, whether highways (Akee 2006; Michaels 2008) or railroads (Banerjee, Duflo and Qian 2009; Atack and Margo 2009; Atack, Bateman, Haines and Margo 2010; Donaldson 2010; Burgess and Donaldson 2010). They show there are significant gains from market integration for connected areas.[3]

We investigate this issue by looking at the impact of modern transportation infrastructure on agriculture and development in colonial Ghana. We study the impact of two railroad lines that were built between 1901 and 1923 on the production and export of cocoa beans and development, which we capture by looking at population growth and urbanization. The first line was built between 1901 and 1903 to connect the port of Sekondi to gold mines and the hinterland city of Kumasi. The second line was built between 1908 and 1923 to connect the port and capital city of Accra to Kumasi. The construction of those railway lines unintendedly opened vast expanses of tropical forest to cocoa cultivation, allowing Ghana to become the world's largest exporter for most of the 20th century. Using data at a very fine spatial level, we find a strong effect of railroad connectivity on cocoa production. We explain that modern transportation infrastructure has considerably reduced transportation costs, as cocoa was before head-loaded or rolled in casks over medium distances. We show that the economic boom in producing areas is associated with demographic growth and urbanization. We find no spurious effect from lines that were not built yet, and lines that were planned but never built. We demonstrate that our results are robust to considering nearest

[2] A related literature focuses on the impact of roads on deforestation (Pfaff 1999; Pfaff *et al.* 2007). Roads then have positive development effects at the expense of the environment.

[3] Those studies also face identification issues, but the placement is more "instrumentable". Michaels (2008); Banerjee, Duflo and Qian (2009); Atack, Bateman, Haines and Margo (2010) use the fact that being on a straight line between two large cities makes it more likely to be connected to a highway or a railroad. Donaldson (2010) does not find any effect for railway lines that were approved but never built.

neighbor estimators. We provide historical evidence for high cocoa incomes, high migration rates in producing areas and urban sectoral specializations in line with consumption linkages stemming from the rising wealth of those areas. We then argue that the "social savings" methodology gives upper bound estimate in our context, directly comparing our results with previous attempts at measuring them in the U.S. (Fogel 1964) or Ghana (Chaves, Engerman and Robinson 2010). Our estimates are more accurate and indicate that the railroad caused 30% of cocoa production, representing 4.5% of GDP. Lastly, we examine the long-term effects of railways by showing that districts connected to the railway system are more urbanized today, have better infrastructure and larger manufacturing and service sectors, despite of a complete displacement of rail by other means of transport.

This paper contributes to the literature on the respective roles of geography and history in determining the location of economic activity. The locational fundamentals theory argues that geographical endowments have a long-term impact on economic activity and there are no multiple spatial equilibria (Gallup, Sachs and Mellinger 1999; Davis and Weinstein 2002, 2008; Holmes and Lee 2009; Bleakley and Lin 2010; Nunn and Qian 2011). The increasing returns theory explains that there is path dependence in the location of economic activity (Krugman 1991; Henderson 2005; Ellison, Glaeser and Kerr 2010; Bleakley and Lin 2010; Combes, Duranton, Gobillon and Roux 2011; Redding, Sturm and Wolf 2011). Thus, historical accidents or regional policy can have long-term consequences on regional inequality, this supporting the existence of multiple spatial equilibria. This paper shows how colonial infrastructure investments one century ago have shaped the economic geography of Ghana, as a result of cumulative agglomeration economics, supporting the second hypothesis. Cocoa production could have boomed anywhere in Southern Ghana, but railroads determined where it boomed first, thus giving a non-natural advantage to some locations. The effects are still visible today.

Hence, we also offer a new concrete channel for colonial legacy arguments. The existing literature relates economic outcomes today to the duration of colonization (Bertocchi and Canova 2002; Feyrer and Sacerdote 2009), the identity of the colonizer (Bertocchi and Canova 2002; Acemoglu and Johnson 2005), and the form of colonization (extraction versus settlement: Sokoloff and Engerman 2000; Acemoglu, Johnson and Robinson 2001, 2002 ; direct versus indirect rule: Mizuno and Okazawa 2009; Iyer 2010). While there is strong evidence that colonization has shaped institutions (Sokoloff and Engerman 2000; Acemoglu, Johnson and Robinson 2001; Acemoglu and Johnson 2005), channels of how colonial investment decisions mattered remain speculative (Huillery 2009; Chaves, Engerman and Robinson 2010; Burgess, Jedwab, Miguel and Morjaria 2011).

The paper is organized as follows. Section 2 presents the historical background of railway construction, cocoa and cities in Ghana and the data used. Section 3 explains the methodology, while section 4 displays the results. Section 5 discusses them and investigates the long-term effects of railways, and section 6 concludes.

2 Background and Data

We discuss some essential features of the Ghanaian colonial economy and the data we have collected to analyze how railway construction has contributed to booming cocoa production, population growth and urbanization. Appendix A contains more details on how we construct the data.

2.1 New Data on Ghana, 1901-1931

In order to analyze the effect of railway construction on agriculture and development in Ghana, we have constructed a new data set on 2091 grid cells of 0.1x0.1 degrees from 1901 to 1931. We choose a high resolution grid because we have precise GIS data on railways, cocoa production, population and urbanization. We obtain the layout of railway lines in GIS from *Digital Chart of the World*. We use various documents to recreate the history of railway construction. For each line, we know when it was started and finished, and when each station was opened. From the same sources, we know lines that were built but not planned. For each real or placebo line, we create cell dummies equal to one if the Euclidean distance of the grid cell centroid to the line is 0-20, 20-40, or 40-60 km. Our main analysis focuses on railway lines in 1918. We also create a dummy equal to one if the cell contains a railway station in 1918. We proceed similarly to construct a GIS database on roads and we create cell dummies for being crossed by a motor road.[4]

[4]There were two types of roads at that time: class II (*roads suitable for motor traffic but occasionally closed*) and class III (*roads suitable for motor traffic in dry season only*).

The data on cocoa land suitability was recreated from a map of cocoa soils in Ghana. A cell is defined as *suitable* if it contains cocoa soils. It is *highly suitable* if more than 50% of its area consists of forest ochrosols, the best soils for cocoa cultivation, and *poorly suitable* if more than 50% of its area consists of forest oxysols, which are poor cocoa soils. Production data was collected from a historical map and we use GIS to calculate the amount of cocoa production (tons) for each cell in 1927. Since production was almost nil around 1900, production in 1927 is the change over the period. We also have at our disposal producer price data in 1931-36 for 173 cells. We then use this data in GIS to spatially extrapolate the producer price for the 260 producing cells in our sample.[5] We then multiply cocoa production in volume by the producer price to obtain the cell value of cocoa production. Obviously, farmers in the interior faced a much lower farmgate price and made lower profits per ton due to higher transportation costs.

To obtain urbanization figures, we construct a GIS database of localities with more than 1,000 inhabitants using census gazetteers. We have data for 1901 and 1931. Using the 1,000 threshold, Ghana had respectively 127 and 428 cities in 1901 and 1931. Since our analysis is at the grid cell level, we use GIS to recreate urban population for each cell-year observation. While we have exhaustive urbanization data for all the country, we only have consistent population data for the South of Ghana. The 1901 census was exhaustively conducted and geospatialized only in the South. We have population data at a very fine spatial level from the 1931 census. We thus have population for each Southern cell (756 cells), and we reconstruct rural population by subtracting urban population from total population.

2.2 The Railway Age

2.2.1 Railroads Built

Improvements in transport infrastructure are typically endogenous, driven by the economic potential that would justify them. Hence, a simple comparison of regions with and without transport infrastructure is misleading, and likely to overstate the output created by it. The railway age in Ghana provides us with a quasi natural

[5]Spatial extrapolation was done by universal spatial kriking, which assumes a general linear trend model. We alternatively use a second order trend and a third order trend. We also use inverse distance weight interpolation.

experiment which we use to identify the effect of reduced transportation costs on cocoa production. This summary draws on Gould (1960), Tsey (1986) and Luntinen (1996). Figure 1 shows the geographic location of the mentioned lines.

After the defeat of the Ashanti Kingdom, the British consolidated their power and established the Gold Coast colony in the South in 1874. Taking a more active administrative role, improving transport infrastructure was on the agenda since trade was constrained by very high transport costs (Dickson 1968; Austin 2007; Chaves, Engerman and Robinson 2010). Draft animals could not be used due to the tse-tse fly transmitting trypanosomiasis, which is deadly for livestock. The Gold Coast also lacked navigable waterways.[6] Headloading was the main means of transport, except for cocoa beans and palm oil, which could be rolled in barrels along the few forest tracks. Owing to the thick primary forest in Southern Ghana, there were only a few well-cleared tracks and villages along those paths were in charge with maintaining them. Railroads were the transport technology of the time. Areas to be connected were not chosen randomly. The colonial administration insisted that the railroad had to be profitable. One important factor determining the location of railroad projects was population density. But the factors finally influencing the decision were exogenous to cocoa production.

The railroad was built by Europeans for Europeans. Strong interest groups of British capitalists lobbied to connect the gold fields in the hinterland. Mines needed heavy machinery, large quantities of firewood (or coal) and workers from other regions. Headloading made gold production prohibitively costly. The colonial administration gave in to the pressure, turning down alternative lines, for which railroad surveys attested a greater potential for agricultural exports and benefit to the country in general (see Appendix C). There were then military reasons to connect the Ashanti capital Kumasi. The British fought over four wars before they annexed the Ashanti Kingdom in 1896. An uprising in 1900 reinforced the perceived threat. The railroad was meant to allow to quickly dispatch troops.

The construction of the first railroad begun in 1898. The line started from Sekondi on the coast and reached the gold mines of Tarkwa and Obuasi in 1901 and 1902 respectively (see Figure 1). The line indeed represents the shortest distance

[6] The Volta River was the only noteworthy exception; shallow-draught boats could be used to ship cargo from Ada Foah to Akuse, about 105 km inland (Dickson 1968).

from the coast to the gold fields. As Obuasi was only 60 km from Kumasi, the line was further extended in 1903. Much of the railroad went through virgin forest. Cocoa did not play a role in the choice of location. In 1900, cocoa exports were tiny, ca. 1000 tons, coming from the area immediately north of Accra. The colonial administration was taken by surprise of the cash crop revolution that followed.

A railway in the eastern part of the colony had been planned from the very first. Accra, as the recently designated administrative capital of the Gold Coast, was chosen as terminus. The colonial administration followed a transportation paradigm of building short longitudinal lines to the interior (Tsey 1986, p. 56). In 1905 it was decided to construct a railway of 65 km from Accra northwards to Mangoasi.[7] Construction started in 1909 and the line opened in 1912. If this line was to link up with Kumasi, topographic conditions - the Akwapim ridge in the east and Atewa range in the west - predetermined the remaining route. The line was extended by 40 km via Koforidua and stopped in Tafo in 1916. Due to wartime shortages, Kumasi was connected only in 1923. An alternative route Accra-Oda-Kumasi around the Atewa range was proposed in 1898 but not built. The region around Koforidua became a centre of cocoa growing, also because of the highly suitable soil quality. The station attracted much traffic, transporting 16,000 tons of cocoa in 1915. The original justifications for the eastern line, however, did not mention at all the cultivation of cocoa; around 1900 the British were more interested in the trade possibilities offered by palm oil and kernel, timber and rubber; the opening of new areas to cocoa cultivation became ex-post an objective of railway construction. The railroad network was subsequently expanded. Other lines, however, opened too late to influence cocoa output in 1927.

2.2.2 Placebo lines

Qualitative accounts do not provide compelling evidence for reverse causality. Still, endogeneity remains a concern. The decision making process was complex and may have followed a different model than what we conclude from qualitative sources. Moreover, despite a comprehensive set of controls such as population density, soil quality and rainfall, there is always a risk of omitted variables. We address those

[7]Mangoasi also served recreational needs of European expatriates (Tsey 1986, p. 58).

concerns using six alternative railroad routes as a placebo check of our identification strategy. Placebo control mimics the tested treatment in all ways except the treatment. Appendix C gives a brief background of each placebo line.

Four alternative railroad routes were proposed before the first line was actually built. These lines were driven by the same model. The aim was to increase exports of gold, palm oil and kernels, and timber, making the lines profitable, ensuring military domination, and bringing civilization to the native people. Judged by observables, the proposed lines were influenced by population density, soil quality, altitude in a similar way as the actual lines built. But the proposed lines are not completely identical to the lines built. The fact that these lines were not built let us wonder as to whether they had a lower probability to be selected. There is not much evidence in favor of this. In fact, the appointment of Governor Maxwell after Governor Griffith's retirement in 1895 seems to have tipped the balance from the central trunk route Accra-Oda-Kumasi to the western railroad line Sekondi-Kumasi (Luntinen 1996). Nevertheless, to dispel the concern, we use two lines that were actually built, but too late to influence cocoa output in 1927.

2.2.3 Reduction in Transport Costs

Railroads permitted a massive decrease in transportation costs. While the freight rate per ton mile was 30-60d for headloading, 23d for cask rolling and 20d for lorries (before 1914), it was only 11.3d for rail. Railway charges decreased even further in 1922, when 7.5d, 6d, 4.5d and 3d per ton mile were charged for the first, second, third and fourth 50 miles (Austin 2007). Did railroads make cocoa cultivation profitable beyond a narrow coastal strip? The producer costs for one ton of cocoa beans were about 8s6d and the opportunity wage of unskilled labor was 1s3d (Cardinal 1931). Using the Accra port price in 1930 (15.5s per ton), we arrive at a maximal Euclidean distance where cocoa cultivation is profitable of 22-30 km for headloading, 59 km for cask rolling, and 120 km for rail, which increases to 451 km for prices charged after 1922. Without rail cocoa cultivation would have been indeed limited to the coast. However, cocoa prices were volatile. It was almost three times higher in 1926-27 (43s), the maximal distance for cask rolling being then 722 km. Our calculation is an upper bound estimate as we consider the Euclidean distance and do not include transaction costs not related to transportation: fees of cocoa brokers, losses due to thief or rotting during transportation, etc.[8] From 1912 on, the share of cocoa transported by rail was around 80% (see Figure 2). According to Luntinen (1996) (p.107) "The very existence of the transport network encouraged the production of surplus for the market. It was cocoa that made the Gold Coast the richest colony in Africa. The farmers seized the opportunity as soon as the railway reached them, so eagerly that foodstuffs had to be imported." The main effect of railroads was thus to make profitable cocoa cultivation in areas where it was not.[9]

Roads were first complementary to the railway system as they were feeders to it. Roads were of poor quality until 1924 when the government started the "Tarmet Program" which made roads suitable to motor traffic throughout the year (Gould 1960). Roads became serious competitor for the railway. Roads also opened new areas to cocoa cultivation. As a result, the share of cocoa transported by rail decreased to 60% after 1924 (see Figure 2). Even if no railway had been built, roads would have permitted the cocoa boom. But our goal is not to compare the respective impacts of railways and roads. We focus on the railway age in 1901-1931 because it provides us with a natural experiment which we use to identify the effect of reduced transport costs on development. The first lorry was imported in 1903, but there were only two lorries in the colony in 1914 (Luntinen 1996).

2.3 The Cash Crop Revolution

Cocoa has been the main motor of Ghana's economic development (Teal 2002; Austin 2008), and this made it a leader of the African "cash crop revolution" (Tosh 1980). Yet, as cocoa is produced by consuming the forest, this success has been a major factor of deforestation. Cocoa farmers go to a patch of virgin forest and replace forest trees with cocoa trees. Pod production starts after 5 years,

[8] Rail distance from Kumasi to Sekondi is 273.1 km, while it is 285.5 km from Kumasi to Accra. Using 3d as the freight rate per ton mile for rail, this meant that farmers in Kumasi received 3.4-3.6d less per ton that farmers close to Sekondi or Accra. Using price data at the grid cell level for 1931-1936, we find that farmers around Kumasi were indeed receiving 3.5d less per ton that farmers around Accra.

[9] Although gold mining was the main original reason for railway construction, the railway lines were mostly transporting manganese ore (52.3% of the total weight of goods being transported), cocoa beans (17.5%) and firewood (12.7%). Coal, gins and spirits, imported food staples, native produce, machinery and exported timber only accounted for 5.4% altogether.

peaks after 25 years, and declines thereafter. When cocoa trees are too old, cocoa farmers start a new cycle in a new forest (Jedwab 2011).[10]

Cocoa was introduced by missionaries in 1859, but it took 30 years before cocoa was widely grown, making Ghana the world's largest exporter as soon as 1911. Figure 2 shows aggregate cocoa production and the export share of cocoa from 1900 to 1931. Figure 3 shows grid cells that are suitable to cocoa cultivation and cocoa production in 1927. Cocoa production originally spread out in the Eastern province from Aburi Botanical Gardens, where the British sold cocoa seedlings at very low price (Hill 1963, p.173-176).[11] As Ghanaian farmers realized how much profit they could make out of cocoa, more and more people specialized in the crop. The Accra-Tafo railway line was instrumental in opening new land to cocoa cultivation. Production also boomed around Kumasi when the railway line was opened in 1904. Why did production boom in Ashanti and not in the South-West, around Sekondi? Transportation costs were certainly lower there as it was closer to the coast. But the South-West of Ghana is characterized by very poor cocoa soils and too much rainfall.[12] The railway line to Kumasi thus made the Ashanti more competitive than the South-West for the development of cocoa cultivation.

2.4 Demographic and Urban Growth in Colonial Ghana

Ghana has experienced dramatic population growth after 1901, due to rising livings of standard, international migrations and large-scale health campaigns organized by the British colonizer. Its population increased from 1,9 millions in 1901 to 2,2 in 1911, 2,5 in 1921 and 3,2 millions in 1931 (Austin 2008). 31.9% of this growth took place in the Gold Coast Colony, 10.6% in Ashanti, 21.6% in Northern Territories and 9.0% in British Togoland (see Figure 1). Figures 4 and 5 display rural population in 1901 and 1931 for Southern cells. The comparison with Figure 3 suggests that rural population increased in areas where cocoa production boomed.

[10] Removing forest trees alters the original environmental conditions and replanted cocoa trees die or are much less productive (Ruf 1995).

[11] The British established the Government Botanical Gardens in Aburi in 1890, because of its health climate and its proximity to the capital city, Accra.

[12] Most of the South-West area consists of oxysols or intergrades, which are poor cocoa soils. The lack of soil minerals leads to low yields and premature tree aging. Annual rainfall often exceeds 2,000mm, with a dry season that is very wet, which favors cocoa diseases.

Austin (2008) describes how the labor-land ratio increased as a result of migration related to cocoa and how some areas in the forest zone were already approaching the "critical population density" for foodcrop growing (p.8-14).

Then, while Ghana was almost unurbanized at the turn of the 20th century, it is now one of the most urbanized countries in Africa. It has also started its urban transition earlier than most African countries, due to the boom in cocoa exports (Jedwab 2011). Defining as urban any locality superior to 5,000 inhabitants, Ghana's urbanization rate increased from 5.0% in 1901 to 7.7% in 1931. But it moved from 22.4% in 1901 to 29.6% in 1931 if we define as urban any locality superior to 1,000 inhabitants. Our period of study thus only captures the beginning of the urban transition in Ghana. The two largest cities were then Accra, the national capital, and Kumasi, the hinterland capital. Altogether, they have accounted for 9.5% of total urban growth in 1901-1931. 41.4% of it has come from the Gold Coast Colony, 21.5% from the Ashanti, 24% from Northern Territories and 3.6% from British Togoland. Thus, around 66.5% of urban growth has come from areas suitable to cocoa cultivation. Figure 6 shows cities in 1901 and 1931, using the 1,000 population threshold. Before the 20th century, towns were state capitals or trading centres (see Dickson 1968, p.70-71). Most of the latter were on the coast, where European merchants would meet local merchants from the interior. But there were also trading centres in the North, which benefitted from their location on historical trade routes. In the early 20th century, most of urban growth took place in the forest zone, with the development of mining, modern transportation and cocoa production (see Dickson 1968, p.246-261). In particular, many towns grew because they were cocoa buying centres, the homes of wealthy cocoa farmers or market towns where cocoa farmers would spend their income.

3 Empirical Strategy

Having data on railway connectivity, cocoa production and population and urban growths at the grid cell level in 1901 and 1931, we test if connected cells experience a boom in cocoa production and whether this drives demographic growth and urbanization. We explain the strategies we implement to obtain causal effects.

3.1 Main Econometric Specification

The main hypothesis we test is whether railway connectivity drives cash crop production, which then drives demographic and urban growths, focusing on 1901-1931 Ghana. We first run the following model for cells c:

$$\triangle Cocoa_{c,1927} = \alpha + \beta Rail_{c,1918} + \theta Urban_{c,1901} + u_c \tag{1}$$

where our dependent variable is cocoa production (tons) of cell c in 1927. Given aggregate production was almost nil in 1901, we study the change in cocoa production. Our variables of interest $Rail_{c,1918}$ are cell dummies which capture railway connectivity: for being 0-20, 20-40 and 40-60 km away from a railway line in 1918 and having a railway station in 1918. We use the railway network in 1918 as the railway lines built later (1923 and 1927) were unlikely to affect cocoa production given it takes 5 years to start. We then run a second model for cells c:

$$\triangle Pop_{c,1931} = \alpha' + \gamma Rail_{c,1918} + \theta' Pop_{c,1901} + \delta Cocoa_{c,1927} + v_c \tag{2}$$

where our dependent variable is population of cell c in 1931, simultaneously controlling for cell population in 1901 ($Pop_{c,1901}$). We actually include just urban population in 1901 and rural population in 1901, as we are then better able to control for initial conditions. We expect railway connectivity to have a positive and significant effect on cocoa production ($\beta > 0$), and population ($\gamma > 0$) when we do not include $Cocoa_{c,1927}$. Once we include $Cocoa_{c,1927}$, we expect it to capture the effect of railway connectivity ($\gamma = 0$, $\delta > 0$). This means the effect of railway connectivity on population goes through more cocoa production. Railway connectivity might also have an independent effect on mining, other tropical products and the trade of consumer goods, but this effect might not be significantly different from 0 once we control for cocoa. We can look at urban growth and rural growth to understand the nature of population growth.

We have a cross-section of 2091 cells. Our main analysis is performed on the restricted sample of suitable cells. We run the risk of just comparing the Southern and Northern parts of Ghana if we use the full sample. We drop observations that are a node of the railway lines: Accra, Kumasi and Sekondi. Our urban analysis is focused on other cities. We also restrict our sample to those cells for which we have data on the cell population distribution in 1901 as this data is only available for a set of Southern cells that year. We end up with 611 observations and we believe such restriction gives us conservative estimates. We will show later that our results are robust to removing such restrictions. We privilege OLS regressions. The issue is whether railway construction was endogenous to cocoa production and demographic and urban growths. We argue in section 2.1 that both the Western and Eastern lines were not planned with having cocoa cultivation or urbanization in mind. We now describe the tests we perform to ensure these effects are causal.

3.2 Controls

We include controls at the cell level to account for potentially contaminating factors. First, we control for demography in 1901, by including urban population, rural populations, and the number of large towns, towns (> 500 inhabitants), head-chief towns, large villages ($100-500$ inhabitants) and villages (< 100 inhabitants). Second, we control for economic geography by having Euclidean distances (km) to Accra, Kumasi, Aburi and the coast and dummies for bordering another country or the sea.[13] Third, we add physical geography variables such as the shares of soils which are ochrosol class 1, class 2, class 3 and unsuitable, oxysols or intergrades to directly control for land suitability to cocoa cultivation, or the mean and standard deviation of altitude (m) and average annual rainfall (mm) in 1900-1960. Lastly, we have a dummy equal to one if the cell has a mine.[14]

3.3 Placebo Regressions

As explained in section 2.2.2 and shown in Figure 1, four major railway lines were planned but never built: Cape Coast-Prasu-Kumasi (1873), Saltpond-Oda-kumasi (1893), Apam-Oda-Kumasi (1898) and Accra-Oda-Kumasi (1899). Two lines were built after 1918: Tafo-Kumasi (1923) and Huni Valley-Kade (1927). The expansion of cocoa cultivation was mentioned as one of the objectives of the

[13]Those dummies also capture the fact that border cells are by construction not entirely contained in the territory of Ghana.

[14]There were five mines in 1931: three gold mines, one diamond field and one manganese mine. Mineral exports amounted to 24.2% of exports in 1930 and the number of Africans engaged in mining was 12,048. Cocoa and mineral products thus accounted for 94.5% of exports in 1930.

two latter lines, but they were unlikely to have any impact on cocoa production in 1927 and urbanization in 1931 given production takes 5 years to start. For each *planned but not built* or *not built yet* line, we create a placebo treatment dummy equal to one if the cell is more than X km from the 1918 lines (not connected) but less than X km from the placebo line (counterfactually connected). We expect no effect on both cocoa production and demographic growth for cells that would have gained access if the placebo had additionally been built before 1918.[15]

3.4 Nearest Neighbor Estimators

We can compare connected cells to non-connected cells, or connected cells to cells that would have been connected if the placebo lines had been built before 1918. This guarantees that treatment and control cells are similar in terms of both observables and inobservables, in the spirit of nearest neighbor estimators. The only difference between the two just comes from the fact that connected cells were effectively treated. We can alternatively consider cells counterfactually connected to planned but not built lines or lines that were not built yet in 1918 or both. The lines that were planned but not built mostly followed historical trade routes and had more population in 1901 along their potential layout, as indicated by the comparison of figures 1 and 5. We then regress each available control on a dummy equal to one if the cell is than 20 km from a 1918 railway line. We alternatively consider as a control group all the suitable cells (see col. (1) of Table 1), placebo cells (see col. (2)) and placebo cells using Cape Coast-Kumasi only (see col. (3)). Placebo cells are similarly defined as cells that are less than 20 km from a placebo line. Treated cells have a larger rural population, lower cocoa soil quality (less ochrosols) and are closer to large cities and the coast than control cells (see col.(1)). Yet, they are farther from large cities and the coast if we compare them to placebo cells (see col.(2)) and do not have a large population if we compare them to placebo cells along the Cape Coast-Kumasi line (see col.(3)). We will thus test various sets of placebo lines to see if differences in observables really matter.

[15]We cannot directly test if being less than X km from a placebo line has any counterfactual effect as a large number of placebo cells are currently connected to the 1918 railway lines. We can only perform a counterfactual analysis for potentially newly connected cells.

4 Results

In this section, we display the main results, examine their robustness and investigate the mechanisms behind the potential effect of railway connectivity on cocoa production, and the effect of cocoa production on population and urban growths.

4.1 Main Results

Table 2 contains our main results for cocoa production and population growth. Column 1 reports the results for regression (1), while columns (2)-(5) display the results for regression (2). All regressions include controls, but unconditional results are discussed in Table 7. We find a strong effect of railway connectivity on cocoa production, but this effect decreases as we move further away from the line and is nil after 40 km (see col. (1)). There is a strong effect of railway connectivity on population growth but only for those cells less than 20 km away from the line (see col. (2)). This indicates that people tend to live in the vicinity of the line, although there is some production beyond 20 km. Interestingly, the rail effect is lower when we include cocoa production, which has a strong effect on population growth (see col. (3)). This means the railway lines have a strong effect on population growth, and that this growth is coming from opportunities in the cocoa sector, and other sectors if there are intersectoral linkages. The remaining rail effect is then picked up by the cell dummy for having a railway station in 1918 (see col. (4)). But the latter effect also becomes lower and non-significant when we include the amount brought to the station in 1923 (see col. (5)). Railway connectivity has thus a strong effect on population growth, but this effect goes through more cocoa production.[16]

Table 3 displays the results for urban growth, which is population growth in localities superior to 1,000 inhabitants. The effects are more or less half those for population growth, which means that the effects are not significantly different if we consider rural growth. Indeed, we obtain the estimates for rural growth by

[16]Including urban population in 1931 in regression (1) does not change the effect of railway connectivity on cocoa production in 1927. This confirms that the relationship is not from railways to population and then to cocoa production. Besides, as discussed in Jedwab (2011), cocoa production is traditional, requiring very few inputs except forested land, axes, machetes, hoes, cocoa beans labor. It is thus very unlikely that population drives cocoa production.

subtracting the urban effect (Table 3) from the total population effect (Table 2).[17] This implies that railway connectivity, by permitting cocoa cultivation, attracts people to work in the countryside, where cocoa cultivation takes place, but also people to work in the cities. Column (3) of Tables 2 and 3 shows that the effect of cocoa production on demographic growth is relatively higher for the countryside than for the cities.[18] But columns (4) and (5) of Tables 2 and 3 indicate that the effect of cocoa production going through larger rail stations only impact urban growth, as the effects for urban growth are the same as for population growth.

Table 4 investigates whether the effect of cocoa production goes through more cocoa beans being produced or higher incomes thanks to the cocoa sector. We run regression (2), except we add cocoa production in value. Whether we consider total population growth (col.(1)-(2)), urban growth (col.(3)-(4)), or rural growth (col.(5)-(6)), the effect of cocoa production is absorbed by the cocoa income variable.[19] This signifies that the cocoa sector impacts demographic growth because it provides people with high incomes. The cocoa income effect is actually much larger for urban growth than for rural growth (see col.(4) versus col.(6)), which confirms that there are urban consumption linkages from the cocoa sector: people are attracted to cocoa-producing areas because of the high income generated by cocoa cultivation, but if they spend a large share of their income in city, this also drives urban growth. Further evidence on consumption linkages is provided in section 4.4, as we now examine the robustness of those results.

4.2 Placebo Regressions and Nearest Neighbor Estimators

Tables 5 and 6 report results for placebo regressions and confirm that there are no spurious effects from lines that were planned but not built (col. (2)-(5)) or lines that were not built yet in 1918 (col (6)-(7)). The only issue is for Huni Valley-Kade (1927), for which we find an effect on cocoa production between 20 and 40 km. This is due to the south-side of the line having some production (see Figures 1

[17] If we use the cell urbanization rate as an outcome, we find that railway connectivity increases the rate by 5.8%, but this effect is only significant at 15%.

[18] The effect of cocoa production on rural growth is equal to 1.5 - 0.5 = 1, while it is equal to 0.5 for urban growth. The effect is thus twice higher for villages.

[19] The effects for cocoa production in volume become negative, but the overall effects of cocoa production _ volume plus value _ remain positive.

and 3), but we argue in the next section that it was permitted by good coastal roads. We then test the robustness of our results to removing controls (see col.(2) and (6) of Table 7), considering placebo cells as the control group (see col.(3) and (7) of Table 7) and using the full sample (see col.(4) and (8) of Table 7). We also alternatively use as a control group counterfactually connected cells for each placebo line, but results are similar to using all placebo cells altogether (results not shown but available upon request). Results are actually higher when we drop controls, use placebo cells as a control group or the full sample. This confirms that the estimates we are privileging in our main analysis are conservative estimates.

4.3 Results on Alternative Transportation Networks

Did railways cause trade diversion? Figure 7 shows historical trade routes, which were forest tracks used by merchants and slave traders (Dickson 1968). The cells amongst those routes were the most likely to lose from the introduction of a modern transportation network. We run regressions (1) and (2), except we now also include a dummy equal to one if the cell is less than 20 km from a historical trade route. If there is trade diversion, we expect a negative effect of being on a trade route on population growth. Results reported in column (5) of Table 8 indicate that those cells were not losing against other control cells. Then, as argued in section 2.2.1, roads only became a competitive transportation technology in the late 1920s. Figure 8 displays roads in 1922, distinguishing class 2 and class 3 roads. Most of the non-coastal roads served as feeders to the railway. Then, many coastal roads were class 2 roads ("roads suitable for motor traffic but occasionally closed") that were used to transport cocoa beans from coastal producing areas to the coast. We run regressions (1) and (2), except we now also include cell dummies for containing a class 2 road or a class 3 road, and their interactions with being less than 40 km from the railway line.[20] Results reported in column (3) of Table 8 confirm that class 2 roads have an independent effect on cocoa cultivation, which is logical since most of them are along the coast. Being on a class 3 road less than 40 km from a railway line also has a strong impact on cocoa cultivation. This stems from those roads being feeders to the railway. We obtain similar results if we consider population

[20] We use only one 0-40 km dummy for railway connectivity to make results easily interpretable.

growth instead (see col.(6)). There is a high but not significant interacted effect of being on a class 3 road and less than 40 km from the railway lines. But the effect of being on a class 3 road on population growth is high and significant if we add the individual class 3 road effect and its interacted effect with the railway (the effect is significant at 5% if we do the F test using the two coefficients).

4.4 Discussion on Channels

The possibility to cultivate cocoa along railway lines attracted migrants to producing areas. Those migrants established their own cocoa farm or worked as laborers for cocoa landowners.[21] Then, cocoa farmers obtained a surplus which they spent on urban goods and services. This created labor opportunities in the urban manufacturing and service sectors, and labor demand was matched by local inhabitants or migrants. Having no data at the cell level for those channels, we use evidence collected by historians and data available at more aggregate spatial levels.

Hill (1963) magisterially describes the origins of the forest colonization process. In the late 1890s, farmers from Akwapim (around Aburi, see Figure 3) migrated to the west to buy forested land and establish cocoa farms, initiating a scramble for land. Those forests were scarcely inhabited and belonged to the inhabitants of the few surrounding towns. She writes (p.15): "Their chiefs were glad to seize the opportunity of selling land outright to the enterprising Akwapim." and "The cash received for the land seemed like a windfall to the vendor chiefs." The proceeds of cocoa farming were invested in buying new land (p.183-186), but there were other forms of investments (p.190-192): house-building in the home town; buying urban properties; sending their children to school; cocoa produce-buying and transportation companies.[22] Migrants first relied on family labor, but large farmers quickly employed laborers from other regions. Production spread to Ashanti, where Asantes were establishing their own farms or leasing land to Akwapim farmers (Austin 2008, p.16). Migration was widespread, with a clear social stratification between late and recent migrants. Cardinal (1931) writes (p.84): "An influx of strangers drawn here as it were to El Dorado has opened up the country to an

[21] This suggests they got a higher income in cocoa-producing areas than in their region of birth.

[22] Unfortunately, she did not discuss how they were consuming their available income, once investments were realized.

extent that no man could have foreseen as possible within so short a period." and "The industrious planter has been forced to hire labor in order to cope with the fruits of his industry and is gradually ceasing to be a working farmer with the inevitable result that in course of time he will be a non-working landlord."

We use data from the 1931 census and other sources to support the historical evidence. First, we learn that 47.6% of adult males were directly involved in cocoa cultivation. Second, those farmers were very wealthy. Back-of-the-envelop calculations indicate that each male cocoa farmer was 15.0% wealthier than unskilled laborers in low price years, but 14.5 times wealthier in high price years. Third, producing regions had a higher rate of migrants, which we capture by the share of people being born in another province or country. Excluding the West of the country where mines also attracted migrants, 14.0% of people were born elsewhere in producing regions versus 10.4% for non-producing areas. Then, the ratio of males to females was 1.06 in producing areas versus 0.94 in non-producing areas. Fourth, we learn that 48.5% of the urban male workforce works in agriculture.[23] The census does not distinguish cocoa farmers and other farmers, but this figure indicates that many farmers use towns as their main residence. 9.5% of urban male workers are then engaged in industry, and 42% in other sectors, mostly services and light manufacturing. The fifteen main non-agricultural occupations are: traders (8.0%), domestic servants (6.6%), commerce (4.5%), carpenters (4.5%), car drivers (3.2%), hawkers (2.7%), tailors (2.4%), government civil servants (2.3%), masons (2.0%), goldsmiths (1.3%), policemen (1.3%), cocoa brokers (1.2%), teachers (1.1%), motor mechanics (1.0%) and washermen (1.0%). Altogether, they account for 82.3% of the non-agricultural sector. The same positions account for 84.2% of the urban male workforce in producing provinces.[24] This means that cities in producing areas and non-producing areas are not significantly different, but producing areas have more cities. A large share of urban employment is thus linked to trade, construction, consumer goods and services and public administration. Overall, those results are consistent with producing areas having a large economic surplus which

[23]The criterion to define cities differs from the 1,000 threshold we use. The total urban population we find for Ghana is three times larger than the one reported in the census.

[24]Considering the whole country, 29.8% of the urban female workforce work in the agricultural sector. Other occupations are: market traders (38.8%), hawker (16.2%), domestic servants (7.7%), bakers (5.4%), dressmakers (1.6%), teachers (0.2%) and government civil servants (0.2%).

is spent on urban goods and services, thus driving urban growth.

5 Discussion/Extensions

We now provide two extensions to our results. First, we compare them to the 'social savings" methodology. Second, we investigate the long-term effects of railroads.

5.1 Welfare Effects

The railway proofed to be profitable. The *private* rate of return was 1% in 1903, steadily increased to 9% in 1909, and varied in the period 1910-1928 thereafter around 7%.[25] Chaves, Engerman and Robinson (2010) have estimated the *social savings* of railroads in the Gold Coast and arrived at a value of 9% of GDP in 1925. The social savings approach is based on Fogel (1964) and is defined as the cost difference to the next-best transportation alternative in the absence of railroads:

$$\text{Social savings} = (c_a - c_r)R \tag{3}$$

where c_r and c_a represent marginal costs of the railroad and the next-best alternative transport respectively; R is the total volume transported by rail. Methodologically, it is best to think of a simple supply demand framework of transportation where cocoa producers represent transport consumers: Under perfect competition (where price p equals c) and completely inelastic demand for transport (a horizontal supply and vertical demand curve), social savings are identical to the change in consumer surplus brought about by the railroad.[26] Where those conditions do not hold, the social savings estimate will exceed the actual increase in consumer surplus (Fogel 1979; Leunig 2010). In particular, the bias increases with price elasticity and $\frac{p_a}{p_r}$. Hence, for Ghana, the two measures may disagree quite considerably. Demand was highly elastic: our results suggest that railroad and modern road infrastructure triggered significant cocoa production inland. Secondly, the railroad was much more efficient than available alternatives in 1900; $\frac{p_a}{p_r}$ ranged between

[25] The interest rate on the loan amounted to 3.5% till the early 1920s, before increasing to 6% (Luntinen 1996).

[26] Jara-Diaz (1986) shows that given those assumptions consumer surplus in transportation exactly equals the net economic benefits created in the products' markets.

4 and 10 (see Table 9). In sum, assuming R under c_a implies a deadweight loss that the social savings approach counts as a gain.[27] In the following, we calculate the social rate of return based on cocoa income alone using: (i) the social savings approach, and (ii) a more direct approach estimating producer rents.

5.1.1 Social Savings Approach

Results of the calculation are presented in Table 9. Details and underlying sources are provided in Appendix C. We consider cask-rolling and head porterage as alternatives to railroads.[28] In areas with suitable roads, particularly north of Accra and around Kumasi, cocoa was rolled in casks of about 500 kg. For head porterage, we consider two cost estimates. Chaves, Engerman and Robinson (2010) assumed costs of 3s per ton mile, which is derived from a daily wage of 9d that was paid to forced labour and in poorer regions of West Africa. Our preferred estimate of costs is 5s per ton mile, which is more in line with what contemporaries reported and the relatively high wage for free labour in the Gold Coast colony. The railroad transported a total of ca. 15 million ton miles of cocoa in 1927. Under zero profits freight revenues equal costs. Available evidence suggests that supernormal profits are indeed largely negligible. Lorry transport was a strong competitor, to which the Gold Coast railway responded with price reductions (Gould 1960). Moreover, the private rate of return was only slightly higher than the interest on capital. Nevertheless, relaxing this assumption of perfect competition, allowing for positive profits lowering railroad costs, would not change results, because in our case social savings are overwhelmingly determined by the very high costs of the alternative transport method. We then calculate the hypothetical costs, if the same volume of cocoa was moved by these other transport methods. The cost difference are the social savings, which we find to range between 3% and 9% of GDP.[29]

[27] We thank Robert Allen for this point.

[28] With the expansion of the road network from the 1920s on, the next-best alternative and biggest competitor to railroads was lorry transport, with costs of 2s3d per ton mile in 1922 and falling. In this exercise, however, we do not consider motor transport as counterfactual, but rather as part of modern transport technology introduced in the Gold Coast colony.

[29] Cocoa accounted for 30% of revenues. Applying the costs of 5s per ton mile to all goods would give social savings of about 27% of GDP. This estimate is considerably larger than the 9% from Chaves, Engerman and Robinson (2010). Besides the low wage rate for carriers, Chaves, Engerman and Robinson (2010) also used an inflated figure for GDP. This demonstrates how

5.1.2 Cocoa Income and Producer Rent

Our analysis allows us to directly estimate cocoa farmers' income and rents that were created by railroad and modern infrastructure. Table 8 tells us how much of cocoa production was caused by railroads (Table 8, col. (1)) and feeder roads interacting with railroads (Table 8, col. (3)). The counterfactual is "no railroad" and is obtained by setting the railroad dummies equal to 0. The effect on quantity is relatively large. Without railroads cocoa production would have been ca. 30% lower. Transport costs caused cocoa prices to significantly vary across space. We take this into account by using farmgate prices of cocoa cooperatives of the same grid. We find that in 1927 cocoa income created by railroads was £2,750,000, or 7.2% of GDP. With £48.5 per ton, the cocoa price was exceptionally high in 1927.[30] Taking the 1920-29 average price as a basis instead (£31.3 per ton) yields £1,750,000, or 4.5% of GDP. Knowing the cocoa farmers' opportunity costs, we find that the producer rent is £1,374,000 for the same period, or 3.5% of GDP.

We can also contrast the benefits to the costs of railroad construction. Cocoa represented only 30% of freight revenues though transportation would require the same fixed capital. In 1927, the railroad's capital outlay was £8,432,831 and the social rate of return would be in the order of 20% to 30%. Clearly, including other goods would increase social returns. Nevertheless, the social rate of return from cocoa is impressive. For comparison, the social rate of return to primary education is below 20% (Psacharopoulos and Patrinos 2004).[31] Both approaches also ignore externalities. Chaves, Engerman and Robinson (2010) pointed to the ending of slavery as one positive externality. In this paper, we present evidence for another externality, increasing population density and the growth of urban areas.

5.2 The Long-Term Effects of Railways

In 1931, the railways were transporting 760,000 tons of goods and 1,340,000 passengers. By the end of the 1960s, the railways were transporting 3,500,000 tons of

sensitive conclusions are to measurement problems.

[30]Because of the life cycle of a cocoa tree, farmers cannot immediately respond to price peaks (Hattink, Heerink and Thijssen 1998).

[31]We thank Francis Teal for this comment. Teal contends that the rate of return for education is actually much smaller.

goods and 6,000,000 passengers. Yet, massive investments in road infrastructure and underinvestments and management issues in the railway sector caused a significant decline of the latter. In 1984, the railways were only transporting 374,000 tons of goods – the same amount as in 1921 – and 2,180,000 passengers. Although recent investments have permitted the railways to be competitive again, they still transport much less than what they could do at independence.

We now study whether districts that are connected to the railways are relatively more developed today despite thirty years of marked decline in rail transportation. In other words, has railway construction at the beginning of the 20th century durably transformed the economic geography of Ghana? We use the 2000 *Population and Housing Census* and the 2000 *Facility Census* to see if districts connected to the railways today are now more urbanized, have better infrastructure and larger manufacturing and service sectors. We have data for 110 districts but we just want to compare connected districts to their nearest neighbors. We thus restrict our analysis to those Southern districts that are suitable to cocoa cultivation. We run the following model for districts d:

$$Dvt_{d,2000} = c + \lambda Rail_{d,2000} + \zeta X_{d,2000} + w_c \tag{4}$$

where our dependent variable $Dvt_{d,2000}$ is a development outcome of district d in 2000: urban density, access to infrastructure and sectoral shares.[32] $Rail_{d,2000}$ is a dummy equal to one if the centroid of the district is less than 20 km from a railway line in 2000.[33] $X_{d,2000}$ is then a set of district-level controls which we can include to account for potentially contaminating factors, whether political economy or economic geography: dummies for containing a national city or a regional capital, being a coastal district, and Euclidean distances to the coast, Accra and Kumasi.

Figure 10 displays urban density in 2000 and cities in 1931, and reports which districts are "connected" according to our definition. Urbanization in 2000 appears highly correlated with urbanization in 1931. When regressing urban density in 2000 on urban density in 1931, we find that the latter explains around 65% of

[32]Urban density is total population in localities with more than 1,000 inhabitants per sq km.

[33]There were very few lines built after 1931. The layout of those lines that were built after 1931 was strongly influenced by the layout of the railway network in 1931. We use the network in 2000, although we could have been using the network in 1931 or 1918.

the former (results not shown but available upon request). Regression results for model (3) are then reported in Table 9. We find that railway districts are relatively more urbanized (+136.7 inhabitants per sq km), have a better access to infrastructure – especially for non-universal public goods such as senior secondary schools (+12.1%), hospitals (+20.4%) or electricity (+12.7%) – and larger manufacturing (+4.7%) and service (+8.7%) sectors.

6 Conclusion

Using railroad construction in colonial Ghana as natural experiment, we have analyzed the impact of modern transportation infrastructure on agriculture and development. We find a strong effect of railroad connectivity on cocoa production. The cash crop boom is then associated with demographic growth and urbanization. While the possibility of producing cocoa attracts migrants to the countryside, consumption linkages seem to account for urbanization. We demonstrate that there are no spurious effect from lines that were not built yet, and lines that were planned but never built. Our results are robust to considering nearest neighbor estimators. Our study has shown that transportation infrastructure investments had large welfare effects and exhibit path dependence, as districts connected to the railway system are now more developed, despite thirty years of marked decline in rail transportation. The channels and pattern that we discussed for Ghana are paralleled in other African countries, e.g. cotton in Burkina Faso, groundnuts in Senegal or coffee and tea in Kenya. Whether the new railway line to Northern Ghana will have similar effects is difficult to say, as this is a very different geographical environment and there is already a widespread road network. But our study has shown that transportation infrastructure investments could have large welfare externalities, as they promote trade integration and economic growth.

References

Acemoglu, D. and Johnson, S. (2005). Unbundling institutions. *Journal of Political Economy*, **113** (5), 949–995.

—, — and Robinson, J. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, **91** (5), 1369–1401.

—, — and — (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics*, **117** (4), 1231–1294.

—, — and — (2005). The rise of europe: Atlantic trade, institutional change, and economic growth. *American Economic Review*, **95** (3), 546–579.

African Development Bank (2003). *Review of the Implementation Status of the Trans African Highways and the Missing Links.* African Development Bank: Tunis, and UN Economic Commission for Africa: Addis Ababa.

African Union (2006). *Rail Development in Africa: Stakes and Prospects, Objectives and Mis- sions of The African Rail Union (ARU).* Document prepared by the ARU General Secretariat for the Conference of African Ministers In Charge of Rail Transport 10-14 April 2006, Brazzaville (Congo).

Akee, R. (2006). *The Babeldaob Road: The Impact of Road Construction on Rural Labor Force Outcomes in the Republic of Palau.* IZA Discussion Papers 2452, Institute for the Study of Labor (IZA).

Atack, J., Bateman, F., Haines, M. and Margo, R. A. (2010). Did railroads induce or follow economic growth? urbanization and population growth in the american midwest, 1850-60. *Social Science History*, **34** (2), 171–197.

— and Margo, R. A. (2009). *Agricultural Improvements and Access to Rail Transportation: The American Midwest as a Test Case, 1850-1860.* NBER Working Papers 15520, National Bureau of Economic Research, Inc.

Austin, G. (2007). Labour and land in ghana, 1874-1939: A shifting ratio and an institutional revolution. *Australian Economic History Review*, **47** (1), 95–120.

— (2008). Resources, techniques, and strategies south of the sahara: revising the factor endowments perspective on african economic development, 1500-2000. *Economic History Review*, **61** (3), 587–624.

Banerjee, A., Duflo, E. and Qian, N. (2009). On the road: Access to transportation infrastructure and economic growth in china, unpublished manuscript, Department of Economics, Massachussets Institute of Technology.

Bertocchi, G. and Canova, F. (2002). Did colonization matter for growth?: An empirical exploration into the historical causes of africa's underdevelopment. *European Economic Review*, **46** (10), 1851–1871.

Bleakley, H. and Lin, J. (2010). *Portage: Path Dependence and Increasing Returns in U.S. History*. NBER Working Papers 16314, National Bureau of Economic Research, Inc.

Burgess, R. and Donaldson, D. (2010). Can openness mitigate the effects of weather shocks? evidence from india's famine era. *American Economic Review*, **100** (2), 449–53.

—, Jedwab, R., Miguel, T. and Morjaria, A. (2011). Ethnic favoritism, unpublished manuscript, Department of Economics, London School of Economics.

Buys, P., Deichmann, U. and Wheeler, D. (2010). Road network upgrading and overland trade expansion in sub-saharan africa. *Journal of African Economies*, **19** (3), 399–432.

Cardinal, A. W. (1931). *The Gold Coast*. Government Printer.

Chaves, I., Engerman, S. L. and Robinson, J. A. (2010). Reinventing the wheel: The economic benefits of wheeled transportation in early colonial british west africa, unpublished manuscript, Department of Political Science, Stanford University.

Clark, X., Dollar, D. and Micco, A. (2004). Port efficiency, maritime transport costs, and bilateral trade. *Journal of Development Economics*, **75** (2), 417–450.

Combes, P.-P., Duranton, G., Gobillon, L. and Roux, S. (2011). Estimating agglomeration economies with history, geology, and worker effects. In E. L. Glaeser (ed.), *The Economics of Agglomeration*, NBER and University of Chicago Press.

Davis, D. R. and Weinstein, D. E. (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review*, **92** (5), 1269–1289.

— and — (2008). A search for multiple equilibria in urban industrial structure. *Journal of Regional Science*, **48** (1), 29–65.

Dercon, S., Gilligan, D. O., Hoddinott, J. and Woldehan, T. (2008). *The impact of agricultural extension and roads on poverty and consumption growth in fifteen Ethiopian villages:*. Tech. rep.

Dickson, K. (1968). *A Historical Geography of Ghana*. Cambridge: Cambridge University Press.

Donaldson, D. (2010). *Railroads of the Raj: Estimating the Impact of Transportation Infrastructure*. NBER Working Papers 16487, National Bureau of Economic Research, Inc.

Ellison, G., Glaeser, E. L. and Kerr, W. R. (2010). What causes industry agglomeration? evidence from coagglomeration patterns. *American Economic Review*, **100** (3), 1195–1213.

Feyrer, J. (2009a). *Distance, Trade, and Income: The 1967 to 1975 Closing of the Suez Canal as a Natural Experiment*. NBER Working Papers 15557, National Bureau of Economic Research, Inc.

— (2009b). *Trade and Income – Exploiting Time Series in Geography*. NBER Working Papers 14910, National Bureau of Economic Research, Inc.

— and Sacerdote, B. (2009). Colonialism and modern income: Islands as natural experiments. *The Review of Economics and Statistics*, **91** (2), 245–262.

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

Fogel, R. W. (1979). Notes on the social saving controversy. *The Journal of Economic History*, **39** (01), 1–54.

Gallup, J. L., Sachs, J. D. and Mellinger, A. (1999). *Geography and Economic Development*. CID Working Papers 1, Center for International Development at Harvard University.

Gould, P. R. (1960). *The Development of the Transportation Pattern in Ghana*. Chicago: Northwestern University Press.

Hattink, W., Heerink, N. and Thijssen, G. (1998). Supply response of cocoa in ghana: A farm-level profit function analysis. *Journal of African Economies*, **7** (3), 424–44.

Henderson, J. V. (2005). Urbanization and growth. In P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, *Handbook of Economic Growth*, vol. 1, *24*, Elsevier, pp. 1543–1591.

Hill, P. (1963). *The Migrant Cocoa-Farmers of Southern Ghana: A Study in Rural Capitalism*. Cambridge: Cambridge University Press.

Holmes, T. J. and Lee, S. (2009). *Economies of Density versus Natural Advantage: Crop Choice on the Back Forty*. NBER Working Papers 14704, National Bureau of Economic Research, Inc.

Huillery, E. (2009). History matters: The long-term impact of colonial public investments in french west africa. *American Economic Journal: Applied Economics*, **1** (2), 176–215.

Iyer, L. (2010). Direct versus indirect colonial rule in india: Long-term consequences. *The Review of Economics and Statistics*, **92** (4), 693–713.

Jacoby, H. C. (2000). Access to markets and the benefits of rural roads. *Economic Journal*, **110** (465), 713–37.

Jacoby, H. G. and Minten, B. (2009). On measuring the benefits of lower transport costs. *Journal of Development Economics*, **89** (1), 28–38.

Jara-Diaz, S. R. (1986). On the relation between users' benefits and the economic effects of transportation activities. *Journal of Regional Science*, **26** (2), 379–391.

Jedwab, R. (2011). Why is africa urbanized but poor? evidence from ressource booms in ghana and ivory coast, unpublished manuscript, Department of Economics, Paris School of Economics.

Johnson, S., Ostry, J. D. and Subramanian, A. (2007). *The Prospects for Sustained Growth in Africa: Benchmarking the Constraints*. IMF Working Papers 07/52, International Monetary Fund.

Krugman, P. (1991). History and industry location: The case of the manufacturing belt. *American Economic Review*, **81** (2), 80–83.

Leunig, T. (2010). Social savings. *Journal of Economic Surveys*, **24** (5), 775–800.

Limão, N. and Venables, A. J. (2001). Infrastructure, geographical disadvantage, transport costs, and trade. *The World Bank Economic Review*, **15** (3), 451–479.

Luntinen, P. (1996). *Railway on the Gold Coast: A meeting of two cultures : a colonial history*. Helsinki: Suomalainen Tiedeakatemia.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics*, **90** (4), 683–701.

Mizuno, N. and Okazawa, R. (2009). Colonial experience and postcolonial underdevelopment in africa. *Public Choice*, **141** (3), 405–419.

Mu, R. and van de Walle, D. (2007). *Rural roads and poor area development in Vietnam*. Policy Research Working Paper Series 4340, The World Bank.

Nunn, N. and Qian, N. (2011). The potato's contribution to population and urbanization: Evidence from an historical experiment. *The Quaterly Journal of Economics*, **126** (2).

Pfaff, A., Robalino, J., Walker, R., Aldrich, S., Caldas, M., Reis, E., Perz, S., Bohrer, C., Arima, E., Laurance, W. and Kirby, K. (2007). Road investments, spatial spillovers, and deforestation in the brazilian amazon. *Journal of Regional Science*, **47** (1), 109–123.

PFAFF, A. S. (1999). What drives deforestation in the brazilian amazon?: Evidence from satellite and socioeconomic data. *Journal of Environmental Economics and Management*, **37** (1), 26 – 43.

PSACHAROPOULOS, G. and PATRINOS, H. A. (2004). Returns to investment in education: a further update. *Education Economics*, **12** (2), 111–134.

RADELET, S. and SACHS, J. (1998). Shipping costs, manufactured exports, and economic growth, presented at the American Economics Association annual meeting.

REDDING, S. J., STURM, D. M. and WOLF, N. (2011). History and industry location: Evidence from german airports. *Review of Economics and Statistics*, **93** (3), 814–831.

RENKOW, M., HALLSTROM, D. G. and KARANJA, D. D. (2004). Rural infrastructure, transactions costs and market participation in kenya. *Journal of Development Economics*, **73** (1), 349–367.

RODRIK, D. (1998). *Trade Policy and Economic Performance in Sub-Saharan Africa*. NBER Working Papers 6562, National Bureau of Economic Research, Inc.

RUF, F. (1995). *Booms et crises du cacao: Les vertiges de l'or brun*. Paris: Ministère de la Coopéeration, CIRAD-SAR and Karthala.

SOKOLOFF, K. L. and ENGERMAN, S. L. (2000). Institutions, factor endowments, and paths of development in the new world. *Journal of Economic Perspectives*, **14** (3), 217–232.

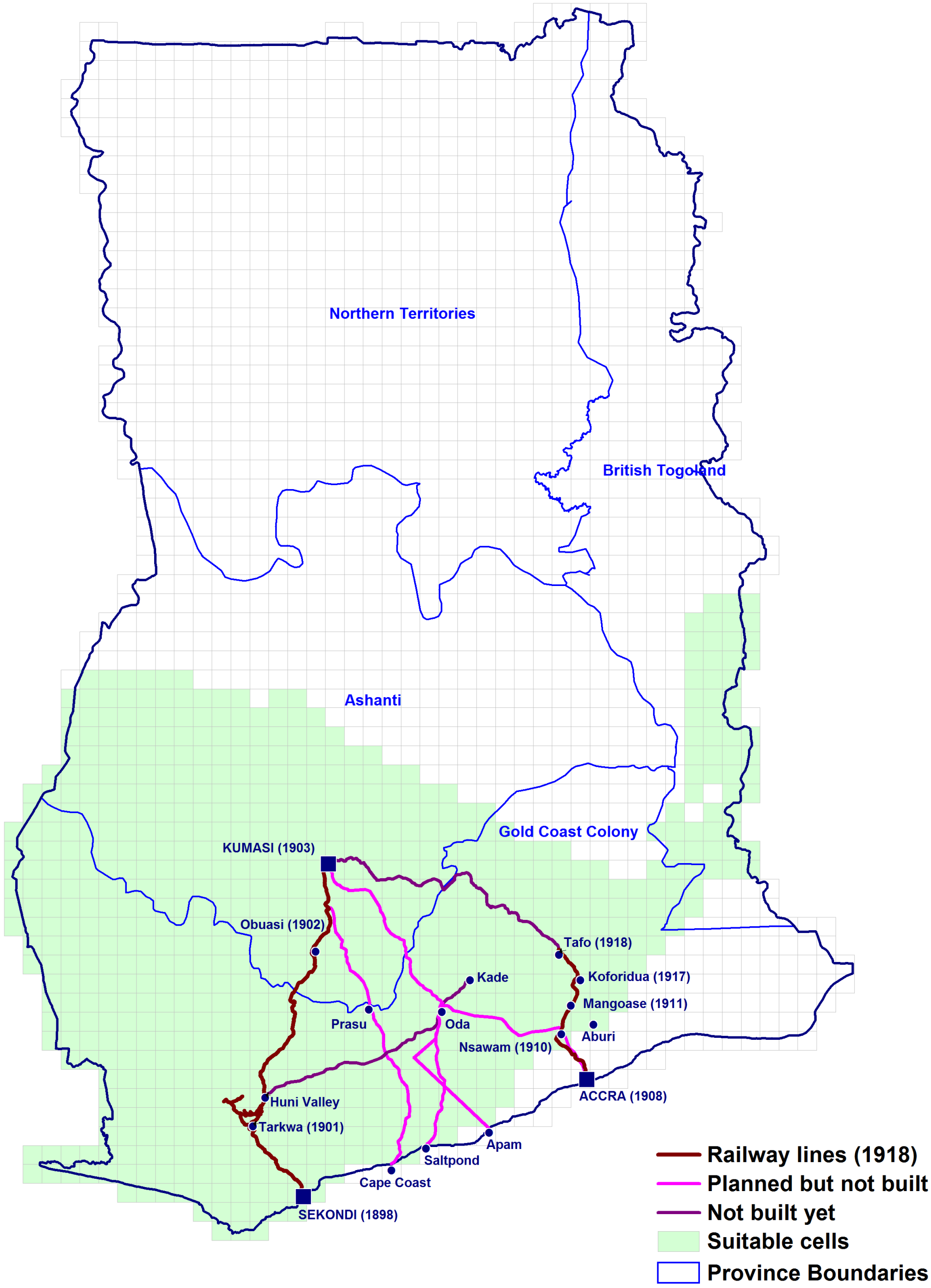
TEAL, F. J. (2002). Export growth and trade policy in ghana in the twentieth century. *The World Economy*, **25** (9), 1319–1337.

TOSH, J. (1980). The cash crop revolution in tropical africa: An agricultural reappraisal. *African Affairs*, **79** (314), 79–94.

TSEY, K. (1986). Gold coast railways: the making of a colonial economy, 1879-1929, phD Dissertation, University of Glasgow.

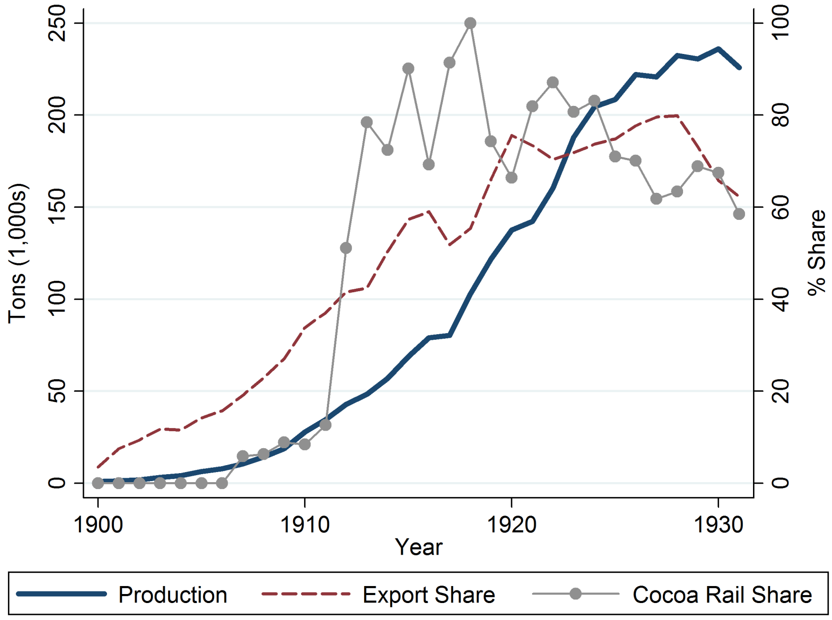
WORLD BANK (2010). *World Development Indicators*. Washington DC: World Bank Data.

Figure 1: Cocoa Land Suitability and Railway Lines in 1918.



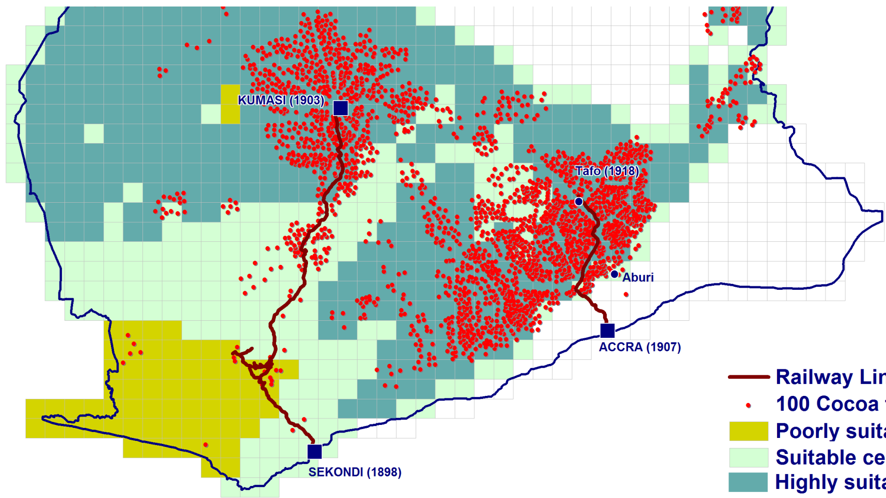
Note: A cell is defined as *suitable* if it contains cocoa soils. Tafo-Kumasi was built in 1923 and Huni Valley-Kade in 1927. Cape Coast-Prasu-Kumasi was planned in 1873, Saltpond-Oda-Kumasi in 1893, Apam-Oda-Kumasi in 1898 and Accra-Oda-Kumasi in 1899. Those projects were later abandoned. Province boundaries date from 1916. See Data Appendix A for sources.

Figure 2: Cocoa Production, Exports and Transportation, 1900-1931.



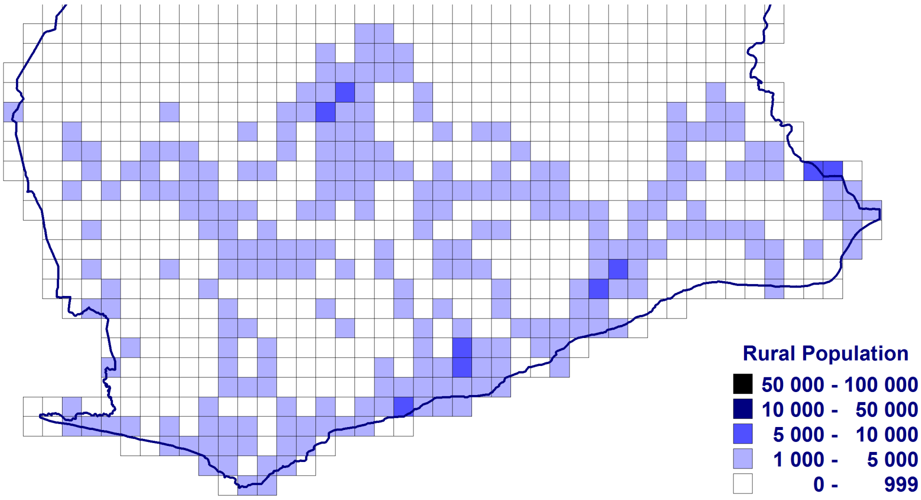
Note: The figure displays three-year moving averages for cocoa production ("Production") and the share of cocoa beans exports in total exports ("Export Share"), and the share of cocoa beans transported by rail to the port ("Cocoa Rail Share"). See Data Appendix A for sources.

Figure 3: Railway Lines in 1918 and Cocoa Production 1927.



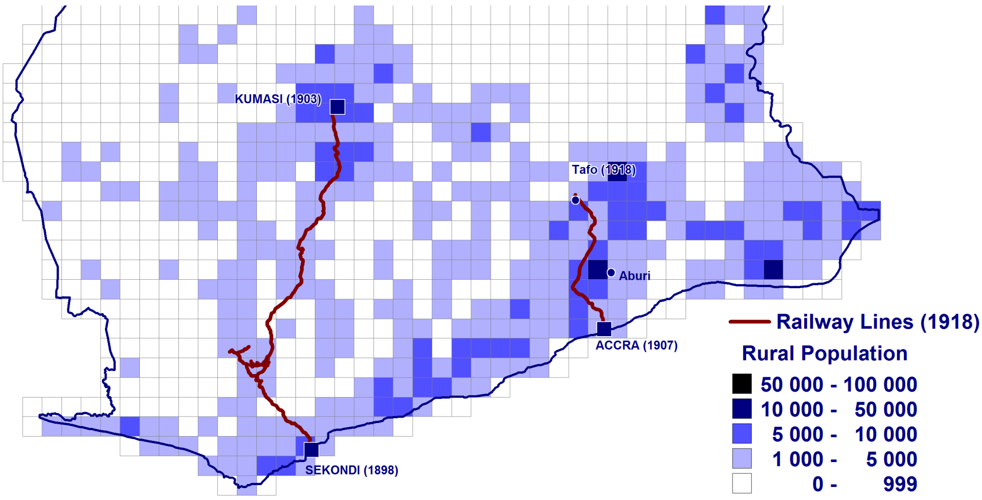
Note: A cell is defined as *suitable* if it contains cocoa soils. It is *highly suitable* if more than 50% of its area consists of forest ochrosols, the best soils for cocoa cultivation, and *poorly suitable* if more than 50% of its area consists of forest oxysols, which are poor cocoa soils. Each red dot represents 100 tons of cocoa production. See Data Appendix A for sources.

Figure 4: Rural Population in 1901.



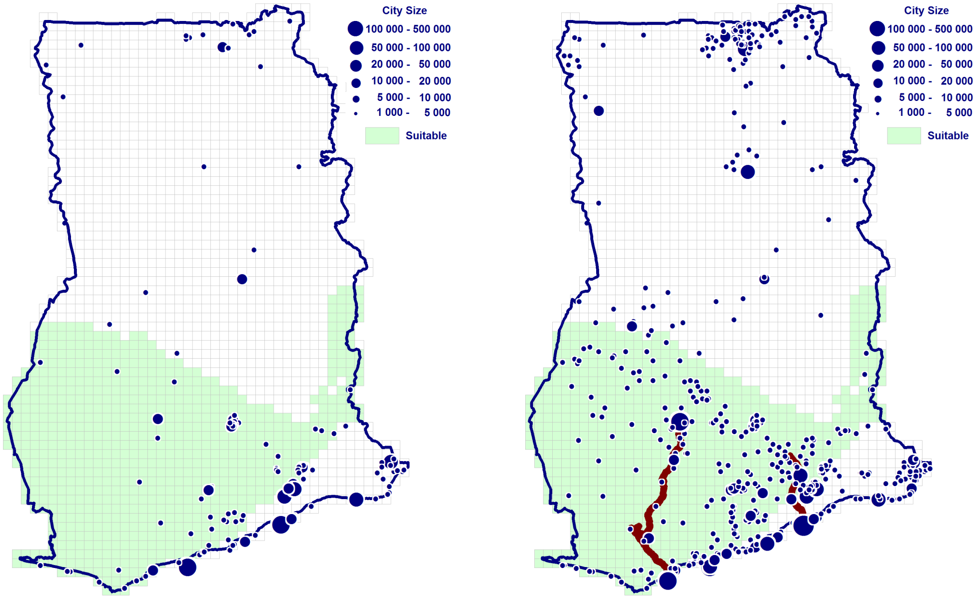
Note: 1901 population data was exhaustively surveyed and georeferenced only for the Gold Coast Colony and part of the Ashanti. For each cell, rural population is the total population of localities with less than 1,000 inhabitants. See Data Appendix A for sources.

Figure 5: Rural Population in 1931.



Note: A map of the distribution of population was published for 1931. For each cell, rural population is the total population of localities with less than 1,000 inhabitants.

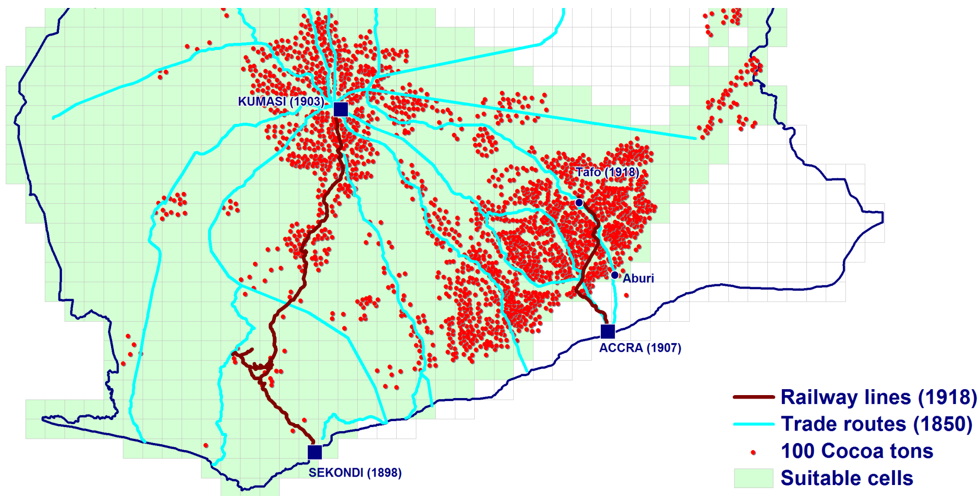
Figure 6: Railway Lines (1918) and Cities 1901-1931.



(a) Cities in 1901 (b) Cities in 1931

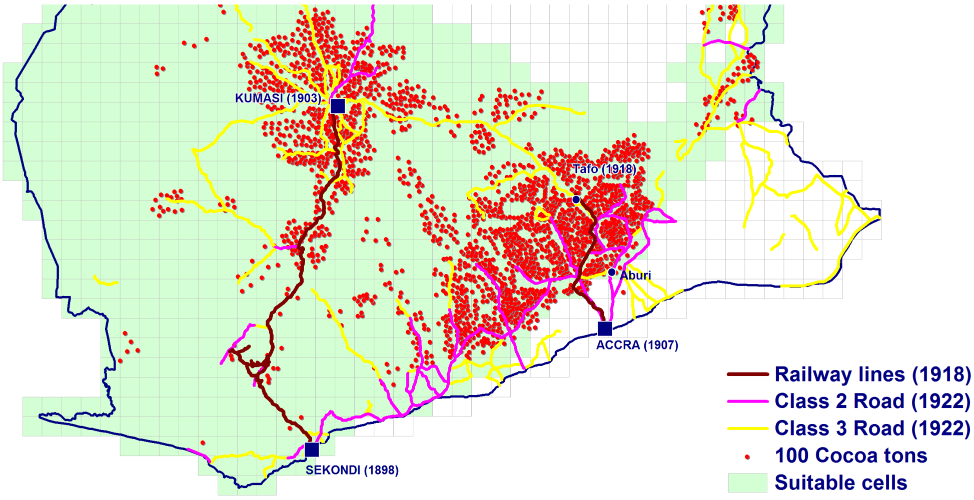
Note: We define as a city any locality superior to 1,000 inhabitants. See Data Appendix A for sources.

Figure 7: Railway Lines in 1918 and Historical Trade Routes in 1850.



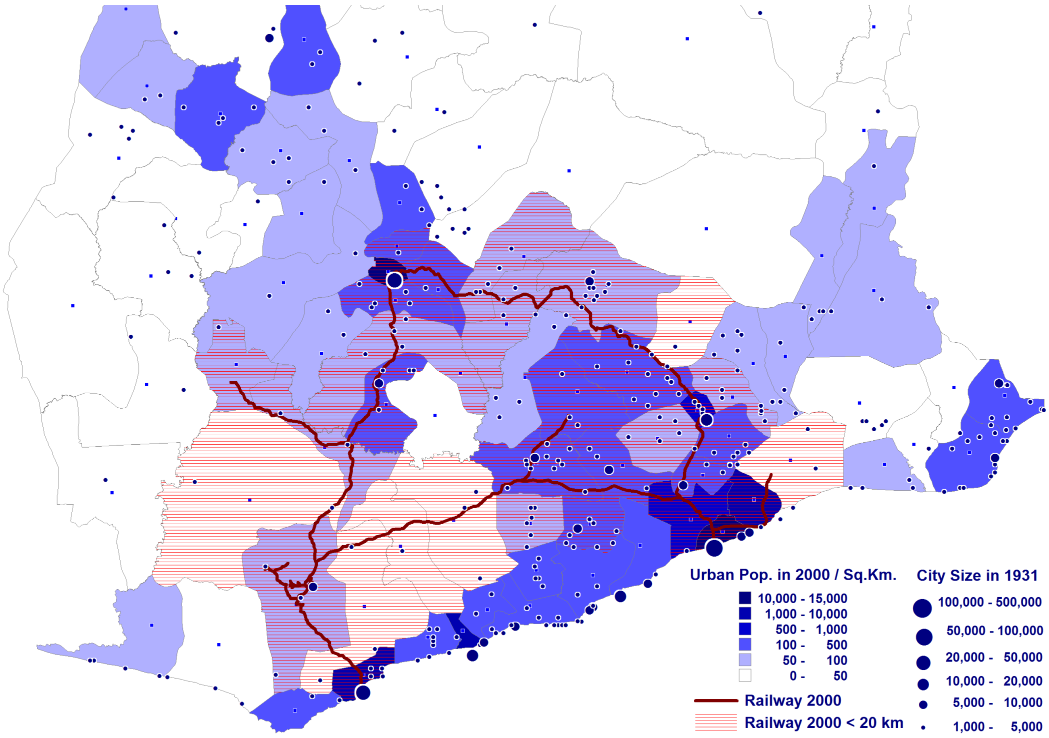
Note: Historical trade routes were forest tracks used by merchants and slave traders. See Data Appendix A for sources.

Figure 8: Railway Lines in 1918 and Roads in 1922.



Note: Class 2 roads are *roads suitable for motor traffic but occasionally closed*, while class 3 roads are *roads suitable for motor traffic in dry season only.* See Data Appendix A for sources.

Figure 9: Urban Density in 2000, Cities in 1931 and Railway Districts.



Note: Railway districts have their centroid less than 20 km from a railway line in 2000. See Data Appendix A for sources.

TABLE 1: OBSERVABLES FOR TREATED VERSUS CONTROL CELLS

RHS Variable:	**Dummy Rail 1918, 0-20 km**		
Control Cells:	All	Placebo	C.Coast-Kumasi
	(1)	(2)	(3)
LHS Variable:			
Dummy "the cell contains a mine"	0.03***	0.03*	0.04
Urban population in 1901	98.5	-135.6	144.6
Rural population in 1901	541.7***	359.8**	150.7
Number of large towns in 1901	0.00	-0.04	-0.03
Number of towns in 1901	0.02	-0.19	0.02
Number of headchief towns in 1901	-0.01	-0.03	-0.02
Number of large villages in 1901	1.34***	1.06***	0.34
Number of villages in 1901	6.38***	5.17***	2.43
Share ochrosols class 1 (%)	0.02**	0.02**	0.02
Share ochrosols class 2 (%)	-0.04*	-0.10***	-0.04
Share ochrosols class 3 (%)	-0.24***	-0.29***	-0.42***
Share ochrosols unsuitable (%)	0.04*	0.03	0.02
Share intergrades (%)	0.20***	0.19***	0.27***
Share oxysols (%)	0.05	0.16***	0.16***
Altitude: mean (m)	-16.88*	-17.73	36.11**
Altitude: standard deviation (m)	2.89	3.21	18.39***
Average annual rainfall (mm)	-29.06	50.92**	116.40***
Distance to Accra (km)	-42.68***	41.06***	37.42***
Distance to Kumasi (km)	-10.57*	13.74**	1.43
Distance to Aburi (km)	-39.94***	42.08***	26.57**
Distance to the coast (km)	-29.89***	-2.81	23.10***
N Treated Group:	110	110	110
N Control Group:	501	152	43

Notes: Standard errors are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. We regress each control on a dummy equal to one if the cell is less than 20 km from a 1918 railway line. We thus run 63 different regressions.

TABLE 2: RAILROADS, COCOA PRODUCTION AND POPULATION GROWTH

Dependent Variable:	**Cocoa Prod. 1927** (Tons)	**Population 1931** (Number of Inhabitants)			
	(1)	(2)	(3)	(4)	(5)
Dummy Rail 1918, 0-20 km	329.5***	1,061.3***	556.6**	286.3	363.4
	[63.9]	[298.6]	[272.0]	[299.6]	[297.9]
Dummy Rail 1918, 20-40 km	125.3**	85.7	-106.3	-96.3	-64.1
	[57.1]	[221.7]	[215.5]	[216.1]	[217.6]
Dummy Rail 1918, 40-60 km	-63.9	-208.7	-110.8	-101.3	-88.1
	[49.1]	[203.4]	[190.2]	[190.1]	[187.2]
Cocoa Production (Tons) 1927			1.5***	1.5***	1.4***
			[0.2]	[0.2]	[0.2]
Dummy Rail Station 1918				1,013.7**	435.3
				[478.1]	[409.4]
Cocoa Tonnages at Rail Station 1918					0.1**
					[0.1]
Observations	611	611	611	611	611
R-squared	0.58	0.72	0.76	0.76	0.77
Controls	Y	Y	Y	Y	Y

Notes: Standard errors are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. Additional controls: urban population in 1901, rural population in 1901, number of large towns, towns (> 500), headchief towns, large villages (100-500) and villages (< 100) in 1901, share (%) of ochrosols class 1, class 2, class 3, unsuitable, intergrades, oxysols, mean and standard deviation (m) of altitude, average annual rainfall (mm), Euclidean distances (km) to Accra, Kumasi, Aburi and the coast.

TABLE 3: RAILROADS, COCOA PRODUCTION AND URBAN GROWTH

Dependent Variable:	**Cocoa Prod. 1927** (Tons)	**Urban Population 1931** (Number of Inhabitants in Localities > 1000)			
	(1)	(2)	(3)	(4)	(5)
Dummy Rail 1918, 0-20 km	329.5***	472.7**	311.5	91.1	164.3
	[63.9]	[198.0]	[193.8]	[199.7]	[196.7]
Dummy Rail 1918, 20-40 km	125.3**	-98.4	-159.8	-151.6	-121.1
	[57.1]	[137.3]	[137.2]	[137.3]	[137.5]
Dummy Rail 1918, 40-60 km	-63.9	-98.2	-66.9	-59.2	-46.7
	[49.1]	[135.8]	[131.5]	[131.2]	[130.3]
Cocoa Production (Tons) 1927			0.5***	0.5***	0.4***
			[0.2]	[0.1]	[0.1]
Dummy Rail Station 1918				826.3**	277.4
				[395.0]	[311.2]
Cocoa Tonnages at Rail Station 1918					0.1*
					[0.1]
Observations	611	611	611	611	611
R-squared	0.58	0.69	0.7	0.71	0.73
Controls	Y	Y	Y	Y	Y

Notes: Standard errors are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. Additional controls: urban population in 1901, rural population in 1901, number of large towns, towns (> 500), headchief towns, large villages (100-500) and villages (< 100) in 1901, share (%) of ochrosols class 1, class 2, class 3, unsuitable, intergrades, oxysols, mean and standard deviation (m) of altitude, average annual rainfall (mm), Euclidean distances (km) to Accra, Kumasi, Aburi and the coast.

TABLE 4: RAILROADS, COCOA INCOME AND POPULATION GROWTH

Dependent Variable:	**Population 1931 (Number of Inhabitants)**					
	Total	Total	Urban	Urban	Rural	Rural
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy Rail 1918, 0-20 km	556.6**	679.4**	311.5	393.9**	245.1	285.5
	[272.0]	[277.7]	[193.8]	[198.3]	[210.2]	[216.1]
Dummy Rail 1918, 20-40 km	-106.3	7.4	-159.8	-83.5	53.5	90.9
	[215.5]	[217.4]	[137.2]	[134.1]	[163.3]	[166.5]
Dummy Rail 1918, 40-60 km	-110.8	-46.8	-66.9	-24.0	-43.9	-22.9
	[190.2]	[192.4]	[131.5]	[131.9]	[142.2]	[144.4]
Cocoa Production (Tons) 1927	1.5***	-3.5*	0.5***	-2.9**	1.0***	-0.6
	[0.2]	[2.0]	[0.2]	[1.4]	[0.2]	[1.7]
Cocoa Income (£, 1931) 1927		0.3**		0.2**		0.1
		[0.1]		[0.1]		[0.1]
Observations	611	611	611	611	611	611
R-squared	0.76	0.76	0.7	0.71	0.67	0.67
Controls	Y	Y	Y	Y	Y	Y

Notes: Standard errors are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. Additional controls: urban population in 1901, rural population in 1901, number of large towns, towns (> 500), headchief towns, large villages (100-500) and villages (< 100) in 1901, share (%) of ochrosols class 1, class 2, class 3, unsuitable, intergrades, oxysols, mean and standard deviation (m) of altitude, average annual rainfall (mm), Euclidean distances (km) to Accra, Kumasi, Aburi and the coast.

TABLE 5: RAILROADS AND COCOA PRODUCTION, PLACEBO REGRESSIONS

Dependent Variable:	Cocoa Production 1927 (Tons)						
Type of Placebo Line:		Planned but not built				Not built yet	
Placebo Line:		C.Coast Kumasi 1873	Saltpond Kumasi 1893	Apam Kumasi 1898	Accra Kumasi 1899	Tafo Kumasi 1923	H.Valley Kade 1927
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dummy Rail 1918, 0-20 km	371.3***	297.1***	352.2***	357.2***	397.1***	347.6***	392.4***
	[54.4]	[59.5]	[61.2]	[60.4]	[56.8]	[54.0]	[54.1]
Dummy Rail 1918, 20-40 km	164.4***	136.5***	148.3***	154.8***	176.5***	146.2***	187.7***
	[47.8]	[48.9]	[49.1]	[48.1]	[47.2]	[46.4]	[45.5]
Dummy Placebo, 0-20 km		-356.5***	-15.9	-40.5	121.1	-56.2	-61.4
		[72.8]	[76.8]	[76.3]	[83.6]	[87.1]	[100.0]
Dummy Placebo, 20-40 km		-28.5	-58.8	-17.1	56.3	-131.6	287.2**
		[113.5]	[93.7]	[93.2]	[118.7]	[86.7]	[129.8]
Observations	611	611	611	611	611	611	611
R-squared	0.58	0.6	0.58	0.58	0.58	0.58	0.59
Controls	Y	Y	Y	Y	Y	Y	Y

Notes: Standard errors are reported in parentheses; * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Additional controls: urban population in 1901, rural population in 1901, number of large towns, towns (> 500), headchief towns, large villages (100-500) and villages (< 100) in 1901, share (%) of ochrosols class 1, class 2, class 3, unsuitable, intergrades, oxysols, mean and standard deviation (m) of altitude, average annual rainfall (mm), Euclidean distances (km) to Accra, Kumasi, Aburi and the coast.

TABLE 6: RAILROADS AND POPULATION GROWTH, PLACEBO REGRESSIONS

Dependent Variable:	**Population 1931 (Number of Inhabitants)**						
Type of Placebo Line:		**Planned but not built**				**Not built yet**	
Placebo Line:		C.Coast Kumasi 1873	Saltpond Kumasi 1893	Apam Kumasi 1898	Accra Kumasi 1899	Tafo Kumasi 1923	H.Valley Kade 1927
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dummy Rail 1918, 0-20 km	1,084.9***	960.5***	1,095.8***	1,112.4***	1,041.7***	1,161.6***	942.8***
	[263.8]	[266.6]	[268.2]	[265.5]	[264.6]	[268.6]	[260.4]
Dummy Placebo, 0-20 km		-769.2**	63.5	164.0	-356.2	518.4	-1,362.5***
		[328.8]	[314.7]	[313.6]	[344.9]	[352.2]	[308.2]
Observations	611	611	611	611	611	611	611
R-squared	0.72	0.73	0.72	0.72	0.72	0.72	0.73
Controls	Y	Y	Y	Y	Y	Y	Y

Notes: Standard errors are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. Additional controls: urban population in 1901, rural population in 1901, number of large towns, towns (> 500), headchief towns, large villages (100-500) and villages (< 100) in 1901, share (%) of ochrosols class 1, class 2, class 3, unsuitable, intergrades, oxysols, mean and standard deviation (m) of altitude, average annual rainfall (mm), Euclidean distances (km) to Accra, Kumasi, Aburi and the coast.

TABLE 7: RAILROADS, COCOA PRODUCTION AND POPULATION GROWTH, ROBUSTNESS

Dependent Variable:	**Cocoa Production (1927)** (Tons)				**Population 1931** (Number of Inhabitants)			
Sample:			Placebo	Full			Placebo	Full
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dummy Rail 1918, 0-20 km	371.3***	457.5***	596.3***	374.3***	1,084.9***	1,374.1***	1,411.9***	1,706.0***
	[54.4]	[79.3]	[76.1]	[54.9]	[263.8]	[340.1]	[378.4]	[330.2]
Dummy Rail 1918, 20-40 km	164.4***	300.6***	379.1***	186.4***				
	[47.8]	[60.4]	[68.1]	[42.0]				
Observations	611	611	379	2086	611	611	262	2086
R-squared	0.58	0.13	0.64	0.57	0.72	0.44	0.75	0.38
Controls	Y	N	Y	Y	Y	N	Y	Y

Notes: Standard errors are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. Urban population and rural population in 1931 are included in all regressions. Additional controls that are omitted in columns (2) and (6) are: number of large towns, towns (> 500), headchief towns, large villages (100-500) and villages (< 100) in 1901, share (%) of ochrosols class 1, class 2, class 3, unsuitable, intergrades, oxysols, mean and standard deviation (m) of altitude, average annual rainfall (mm), Euclidean distances (km) to Accra, Kumasi, Aburi and the coast. The following variables are also omitted in columns (4) and (8) as the 1901 population census was only exhaustively conducted and georeferenced in the south of the country: rural population in 1901, number of large towns, towns (> 500), headchief towns, large villages (100-500) and villages (< 100) in 1901

TABLE 8: RAILROADS, COCOA PRODUCTION AND POPULATION GROWTH, ADDITIONAL RESULTS

Dependent Variable:	**Cocoa Production (1927)** (Tons)			**Population 1931** (Number of Inhabitants)		
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy Rail 1918, 0-40 km	233.4***	239.5***	160.9***			
	[42.9]	[44.9]	[49.9]			
Dummy Rail 1918, 0-20 km				1,084.9***	1,109.9***	635.1**
				[263.8]	[266.4]	[271.6]
Dummy Cell Trade Route		13.8			122.7	
		[42.3]			[168.5]	
Dummy Cell Road Class 2			187.5*			1,314.9***
			[108.6]			[425.3]
Dummy Cell Road Class 3			-48.5			293.2
			[53.0]			[198.2]
Dummy Rail 1918 x Road Class 2			53.9			-106.2
			[125.7]			[775.0]
Dummy Rail 1918 x Road Class 3			167.8**			872.1
			[83.1]			[549.6]
Observations	611	611	611	611	611	611
R-squared	0.57	0.57	0.58	0.72	0.72	0.74
Controls	Y	Y	Y	Y	Y	Y

Notes: Standard errors are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. Additional controls: urban population in 1901, rural population in 1901, number of large towns, towns (> 500), headchief towns, large villages (100-500) and villages (< 100) in 1901, share (%) of ochrosols class 1, class 2, class 3, unsuitable, intergrades, oxysols, mean and standard deviation (m) of altitude, average annual rainfall (mm), Euclidean distances (km) to Accra, Kumasi, Aburi and the coast.

TABLE 9: ESTIMATES OF SOCIAL SAVINGS 1927 (COCOA TRANSPORT)

		Social Savings	
		(£)	% of GDP
Total freight (ton miles)	1,4807,648		
Total freight revenue (£)	355,697		
Costs of alternative means of transports (£)			
- Cask rolling (1.9s/ton mile)	1,406,727	1,051,030	2.74
- Head porterage (3s/ton mile)	2,221,147	1,865,450	4.87
- Head porterage (5s/ton mile)	3,701,912	3,346,215	8.74

Notes: Details of calculations and sources are reported in Appendix B.

TABLE 10: RAILWAY CONNECTIVITY AND ECONOMIC DEVELOPMENT IN 2000

RHS Variable:	**Dummy Rail 2000, 0-20 km**	
LHS Variable:	Coeff. (1)	Sample Mean
Urban population in 2000 / Sq.Km.	137.56*	166.8
Share of district pop. $<$ 10 km from primary school (%)	0.87***	99.3
Share of district pop. $<$ 10 km from junior sec. school (%)	1.68**	98.3
Share of district pop. $<$ 10 km from senior sec. school (%)	12.14*	69.0
Share of district pop. $<$ 10 km from health clinic (%)	7.36**	85.6
Share of district pop. $<$ 10 km from hospital (%)	20.42***	57.0
Share of district pop. $<$ 10 km from post office (%)	16.92**	74.3
Share of district pop. $<$ 10 km from telephone (%)	12.09	68.2
Share of district pop. with access to electricity (%)	12.70**	35.6
Share of district pop. with access to tapwater (%)	0.23	27.3
Employment share of primary sector (%)	-13.39***	63.1
Employment share of secondary sector (%)	4.66***	12.9
Employment share of tertiary sector (%)	8.73***	23.9
N Treated Group:	29	
N Control Group:	31	
Controls	Y	

Notes: Standard errors are reported in parentheses; * $p<0.10$, ** $p<0.05$, *** $p<0.01$. We regress each outcome variable on a dummy equal to one if the district centroid is less than 20 km from a 2000 railway line (13 regressions). We use the 2000 *Facility Census* and the 2000 *Demographic and Housing Census* to reconstruct those outcome variables at the district level (110). As we restrict to suitable districts only, we obtain 60 districts. Controls: dummies for containing a national city (Accra, Kumasi) or a provincial capital, or being a coastal district, and Euclidean distances (km) to the coast, Accra and Kumasi.

Appendices

A Data Description

This appendix describes in details the data we use in our analysis.

Spatial Units:

We assemble data for 2091 grid cells of 0.1x0.1 degrees in Ghana from 1901 to 1931.[1] We choose such a high resolution grid because we have very precise GIS data on railways, cocoa production and urbanization. Each grid cell has the same size, except those cells that are coastal or crossed by a border. We create two dummies equal to one if the grid cell is coastal or bordering another country to control for this issue. Grid cells then belong to 4 provinces.[2]

Railway Data:

We obtain the layout of railway lines in GIS from *Digital Chart of the World*. We then use Gould (1960), Dickson (1968) and Luntinen (1996) to recreate the history of railway construction. For each line, we know when it was surveyed, planned, started and finished, and when each station was reached and opened. From the same sources, we know lines that were built but not planned. Most of those placebo lines follow historical trade routes and have become roads later. Using the GIS road network also available from *Digital Chart of the World*, we recreate in GIS those placebo lines. We then calculate for each grid cell the Euclidean distance (km) from the cell centroid to each real or placebo line. Lastly, we create a set of dummies equal to one if the grid cell is less than X km away from the railway line: 0-20, 20-40 and 40-60 km. We create a dummy equal to one if the grid cell contains a rail station in 1918. We also know how much cocoa production (tons) was brought to the station in 1924.[3]

Cash Crop Production and Price Data:

A very precise map of cash crop production in 1927 was obtained from the *1927 Yearbook of the Gold Coast* and digitized. This map displays dots for each 100 tons of cocoa production.[4] We then use GIS to reconstruct total cocoa production

[1] 0.1x0.1 degrees more or less correspond to 11x11 km.

[2] We cannot use districts as district boundaries are too different between 1901 and 1931.

[3] This information was retrieved for the main railway stations from the 1929 *Yearbook of The Department of Agriculture of The Gold Coast*.

[4] Aggregating all the dots, we obtain 209,100 tons of cocoa production in 1927, which is very

(tons) for each grid cell using 1927 as an approximation for 1931. We then use Bateman (1965) to obtain the international and national producer price. We also have at our disposal producer price data in 1931-36 for 173 cells. This data was obtained from cooperative data collected by Cazzuffi and Moradi (2010). We use this data in GIS to spatially extrapolate the producer price for the 260 producing cells in our sample.[5] We then multiply cocoa production in volume by the producer price to obtain the cell value of cocoa production.

Urban and Population Data, 1901 and 1931:

We collect total and urban population data from the gazetteers of the *Population and Housing Censuses* 1901 and 1931. They list localities and their population size. Defining as a city any locality with more than 1,000 inhabitants, we obtain a geospatialized sample of 439 cities in Ghana. Using GIS, we recalculate total urban population for each grid cell. It was then impossible to find the geographical coordinates of all the villages in 1901 and 1931. Yet, the 1901 census was exhaustively conducted and geospatialized in the South of Ghana (756 cells). We know for each cell the number of *large towns*, *towns* (more than 500 inhabitants), *head chief towns*, *large villages* (100-500 inhabitants) and *villages* (less than 100 inhabitants). Using GIS, we can deduce for each cell the number of villages that are less than 100 inhabitants, the number of villages that have between 100 and 500 inhabitants and the number of villages that have between 500 and 1,000 inhabitants. From the census, we know the average settlement size for each category and we can reconstruct total rural population for each cell in 1901. For 1931, we have a map of the distribution of population for the whole country.[6] This map displays at a very fine spatial level settlements that have less than 500 inhabitants and settlements that have between 500 and 1,000 inhabitants. From the census, we know the average settlement size for each category and we can reconstruct total rural population for each cell in 1931. To conclude, we have the total urban population for each cell in 1901 and 1931. We then have the total rural population for each Southern cell in 1901 and 1931. We can then reconstruct total population for the southern cells in 1901 and 1931.[7] In our main analysis, we restrict our sample

comparable with what we find from national estimates (see Gunnarsson 1978).

[5]Spatial extrapolation was done by universal spatial kriking, which assumes a general linear trend model. We alternatively use a second order trend and a third order trend. We also use inverse distance weight interpolation.

[6]The map was obtained from the 1960 *Ghana Population Atlas*.

[7]We then obtain a total population of 1,085,557 people for Southern Ghana in 1901, compared to 1,043,000 as officially recorded for the Gold Coast Colony. In 1931, we obtain 3,283,660 people for the whole country, against 3,164,000 as officially recorded in the census report. We believe

to those cells for which we have rural population data in 1901. We then show our results are robust to extending our analysis to the rest of the country.

Historical Trade Route and Road Data:

The layout of historical trade routes is obtained from Dickson (1968) (p.215), and digitized in GIS. We use various sources to reconstruct a GIS database of roads in 1922: Gould 1960 and *Map of The Gold Coast with Togoland Under British Mandate*, published in 1930 by the Survey Headquarters. Those road map have a consistent legend showing class 1 roads ("*roads suitable for motor traffic throughout the year*"), class 2 roads ("*roads suitable for motor traffic but occasionally closed*") and class 3 roads ("*roads suitable for motor traffic in dry season only*") and other roads. Other roads are not suitable to motor traffic and are not consider here.

Mining Production and Price Data:

We use annual production data for Ghanaian mines in 1901-1931 for four commodities: gold, manganese and diamond.[8] As we have the geographical coordinates of each mine, we create a dummy equal to one if the cell contains a mine.

Geographical Data:

Forest data comes from land cover GIS data compiled by Globcover (2009). The data displays those areas with virgin forest or mixed virgin forest/croplands, which were areas with virgin forest before it was cleared for cash crop production. Soil data comes from the *1958 Survey of Ghana Classification Map of Cocoa Soils for Southern Ghana.* This map was digitized in GIS and we calculated for each cell the share of land which is suitable to cocoa cultivation. We also know the respective shares of land which consists of ochrosols (first class, second class, third class, unsuitable), oxysols and intergrades. A cell is defined as suitable if it contains cocoa soils. It is then highly suitable if more than 50% of its area consists of forest ochrosols, the best soils for cocoa cultivation, and poorly suitable if more than 50% of its area consists of forest oxysols, which are very poor cocoa soils. Climate data comes from *Terrestrial Air Temperature and Precipitation: 1900-2007 Gridded Monthly Time Series, Version 1.01*, 2007, University of Delaware. We estimate for each grid cell average annual precipitations (mms) in 1900-1960. Topography comes from SRTM3 data. We estimate for each grid cell the mean and standard deviation of altitude (meters). The standard deviation captures the

our estimates are entirely reliable.

[8]Mining production and price data is collected from the following documents: *The Mineral Industry of the British Empire and Foreign Countries 1913-1919*; *Reports of the Mines Department of the Gold Coast* 1931-1958.

slope and ruggedness of the terrain. Lastly, we use GIS to obtain the Euclidean distance (km) from each cell centroid to the coast.

Economic Geography Data:
For each grid cell, we use GIS to get the Euclidean distances (km) to Accra, the capital city, Kumasi, the largest hinterland city, and Aburi, the city where Ghanaian cocoa production originated.

Urbanization, Infrastructure and Employment Data:
We use the 2000 *Facility Census* and the 2000 *Population and Housing Census* to recreate welfare data at the district level (110) in 2000: urban density (number of people in localities with more than 1,000 inhabitants per sq km), share of people (%) living less than 10 km from various types of infrastructure, and employment shares in the primary, secondary and tertiary sectors. We divide districts between suitable (having some cocoa soil) and non-suitable districts. We create a dummy equal to one if the district centroid is less than 20 km from a railway line in 2000.

References

Bateman, M. (1965): *Cocoa in the Ghanaian Economy.* Unpublished Ph.D. dissertation, MIT.

Cazzuffi, C., and A. Moradi (2010): "Why Do Cooperatives Fail? Big versus Small in Ghanaian Cocoa Producers' Societies, 1930-36," CSAE Working Paper Series 2010-18, Centre for the Study of African Economies, University of Oxford.

Dickson, K. (1968): *A Historical Geography of Ghana.* Cambridge: Cambridge University Press.

Globcover (2009): *Globcover Regional: Africa.* Louvain: European Space Agency Ionia GlobCover.

Gould, P. R. (1960): *The Development of the Transportation Pattern in Ghana.* Chicago: Northwestern University Press.

Gunnarsson, C. (1978): *The Gold Coast Cocoa Industry 1900-1939: Production, Prices and Structural Change.* Lund: Published Thesis.

Luntinen, P. (1996): *Railway on the Gold Coast: A meeting of two cultures : a colonial history.* Helsinki: Suomalainen Tiedeakatemia.

B Placebo Lines

This appendix presents background information on the placebo lines. It draws heavily from Gould (1960), Tsey (1986) and Luntinen (1996). Figure 1 shows the location of the lines.

B.1 Lines Proposed But Never Built

Several private initiatives submitted proposals trying to convince the Colonial Office of the profitability of their railway schemes. All asked for a government guarantee of interest on capital outlay. The government opposed this, with reference to the obvious incentive problems of such a guarantee, subsidising British investors and leading, as in India, to over-capitalisation. These initiatives did not enter a phase of concrete planning, nor were they able to raise the capital necessary for their schemes. Eventually, the state favoured public ownership. The government proposals described below can be considered as alternatives to the two routes built.

Cape Coast-Kumasi 1873

The first proposal to build a railway was made in 1873 to connect Cape Coast to Kumasi via Prasu. The British planned to use the line to send troops to fight the Ashanti. Some railway materials were landed at Cape Coast, but the project was dropped, since it was not possible to build the line in time for the military operation. The line was also proposed in 1891 and 1892 by private consortia (Luntinen 1996, p. 18). Cape Coast was the capital of the Gold Coast colony before it was moved to Accra in 1877. The latter had a drier climate and was believed to be a healthier place for Europeans. In 1901 Cape Coast was similar in size to Accra (respectively 28,948 and 30,144). Moreover, Cape Coast was the starting point of an important historical trade route to Kumasi. Villages clustered along this road like pearls on a string (see Figures 4 and 7). This permitted Cape Coast to achieve the largest trade volume of the coastal towns in 1900 (Gould 1960, p. 17). Hence, in terms of existing traffic, the line also had some potential.

Saltpond-Oda-Kumasi 1893

In 1893 the colonial goverment commissioned a survey for a railroad network that would benefit the whole country, not only the mining industry. Government and engineers favoured a trunk route. Saltpond was chosen as starting point for its central location and because construction materials could be landed easily. The line was to reach Kumasi, crossing densely populated and rich palm kernel and palm oil areas. The line had the support of Governor Griffith. The new governor in 1895, however, changed course and again favoured the mining industry. The project was eventually dropped because of the relatively higher capital outlay compared to Apam as starting point.

Apam-Oda-Kumasi 1898

Apam was considered as an alternative starting point. The line was slightly shorter, thus cheaper, and had all the advantages of a central railway route. Calculations of the consulting engineer indicated the profitability of the line.

Accra-Oda-Kumasi 1899

There was an obvious interest to connect the capital. A line from Accra to Apam was discussed, but this would have implied two rivaling ports. Harbour engineers and new governor Hodgson therefore favoured Accra as terminal point. Eventually, the possible extension of the Sekondi-Tarkwa line - though originally intended as a short, local mining railway - to Kumasi undermined the central route strategy.

B.2 Lines Not Yet Built

The rail network was subsequently expanded. Hence, we also consider lines that were actually built, but not in time to affect cocoa production in 1927. Note that cocoa is a perennial crop. Pod production of the type of cocoa predominantly grown in Ghana starts after 5 years, peaks after 25 years, and declines thereafter. Hence, for observing an impact on cocoa production in 1927, farmers must have grown cocoa trees before 1923. There is no qualitative evidence that this happened to a significant degree. If the prospect of railroad connection did indeed induce much production in advance, it would add to the positive correlation expected from reverse causality (cocoa production attracted the railroad). However, we do not find a positive effect of the two placebo lines below (last two columns, Table 5).

Tafo-Kumasi 1923

The eastern railroad line, with Accra as terminus, reached Tafo in 1916, when war time restrictions on construction materials suspended all further railway projects. Bauxite discoveries, midway between Tafo and Kumasi in 1917, led to the decision to extend the Accra-Tafo railway to Kumasi (Tsey 1986, p. 64). Actual construction, however, only started in 1920. The line was completed in 1923.

Huni Valley-Kade 1927
The line ran parallel to the coast, about 80 kilometres inland. It connected the diamond mines at Kade and was supposed to tap and encourage cocoa, kola, palm oil and timber exports. By conveying more traffic to the newly developed harbour at Takoradi it was hoped to make the port viable. Construction begun in 1923. Several roads already connected the area to the coast, but they were of poor quality. Railroad surveyors found them suitable for light traffic only and believed that lorry traffic could not operate profitably beyond 50 kilometres from the coast. This turned out to be wrong. The short distance to the coast made lorry transport very competitive reviving the old ports of Cape Coast, Saltpond and Winneba directly in the South.

References

Gould, P. R. (1960): *The Development of the Transportation Pattern in Ghana.* Chicago: Northwestern University Press.

Luntinen, P. (1996): *Railway on the Gold Coast: A meeting of two cultures : a colonial history.* Helsinki: Suomalainen Tiedeakatemia.

Tsey, K. (1986): “Gold Coast Railways: the Making of a Colonial Economy, 1879-1929,” PhD Dissertation, University of Glasgow.

C Social Savings

The total freight revenues from cocoa were taken from the 1928 *Gold Coast Railway and Habour Administration Report*. The same source stated the quantity of cocoa loaded by station; multiplying this with the distance of those stations to the port we obtained total freight ton miles. The GDP figure of £38,290,000 (current £) refers to 1930 and was derived from Omaboe (1960), who revised figures from Cardinal (1931). Chaves, Engerman, and Robinson (2010), in contrast, arrived at a GDP of £60,737,040 in 1925-26 by extrapolating an estimate of Szereszewski (1965) using an obscure 5% annual growth rate between 1911 and 1925. Costs of cask rolling were reported by Gould (1960) (p.25). Cask rolling required suitable roads (Dickson 1968, p.38; Hogendorn 1969, p.323). Hence, the rates understate the true full costs of this means of transport. Chaves, Engerman, and Robinson (2010) assumed head porterage costs of 3s per ton mile. Their estimate is based on Lugard (1922) and his description of the situation in Africa in general, whereby a carrier could carry 65lbs (30kg) and walk 12 miles a day; costs were then derived from a daily wage of 9d plus some allowance for subsistence and supervision. Chaves, Engerman, and Robinson (2010) justified their choice of 9d as this was generally paid in Sierra Leone and Ghana's Northern Territories - relatively poor regions - and for compulsory labour in Nigeria. Clearly, forced labour is associated with a loss in social welfare as the carriers' opportunity costs were almost certainly higher than the wage paid. We consider their estimate of 3s per ton mile to be far too low. *The Gold Coast Handbook* 1924 (p.222) reports minimum costs of head porterage of 5s per ton mile instead, an estimate consistent with a wage of 1s3d per day, which was the going rate for carriers in Ghana at that time (*Gold Coast Blue Book* 1928, Gould (1960), p.16). According to Kay and Hymer (1972) (p.145), carriers earned a wage of 1s6d per day plus 6d for the return journey without a load, which was considered "very moderate for the cocoa season".

References

Cardinal, A. W. (1931): *The Gold Coast*. Government Printer.

Chaves, I., S. L. Engerman, and J. A. Robinson (2010): "Reinventing the Wheel: The Economic Benefits of Wheeled Transportation in Early Colonial British West Africa," Unpublished manuscript, Department of Political Science, Stanford University.

Dickson, K. (1968): *A Historical Geography of Ghana*. Cambridge: Cambridge University Press.

Gould, P. R. (1960): *The Development of the Transportation Pattern in Ghana*. Chicago: Northwestern University Press.

Hogendorn, J. N. (1969): "Economic Initiative and African Cash Farming: Pre-Colonial Origins and Early Colonial Developments," in *Colonialism in Africa, 1870-1960: The Economics of Colonialism*, ed. by P. Duignan, L. Gann, and V. Turner, pp. 283–328. Cambridge: Cambridge University Press.

Kay, G., and S. Hymer (1972): *The Political Economy of Colonialism in Ghana*. Cambridge: Cambridge University Press.

Lugard, F. (1922): *The Dual Mandate in British Tropical Africa*. London: William Blackwood and Sons.

Omaboe, E. (1960): "Ghana's National Income in 1930," *Economic Bulletin of Ghana*, 4, 6–11.

Szereszewski, R. (1965): *Structural Changes in the Economy of Ghana, 1891-1911*. London: Weidenfeld & Nicolson.