THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE ■

Economic History Working Papers   No: 272/2018

The Long-Term Impact Of Italian Colonial Roads In The Horn Of Africa, 1935-2000

Mattia C. Bertazzini
London School of Economics

LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

DEPARTMENT OF ECONOMIC HISTORY

WORKING PAPERS

NO. 272 – JANUARY 2018

The Long-Term Impact Of Italian Colonial Roads In The Horn Of Africa, 1935-2000

Mattia C. Bertazzini
M.Bertazzini@lse.ac.uk
London School of Economics

Abstract

Between 1935 and 1940 the Italians built an extensive road network to facilitate the occupation of Ethiopia and secure control over the Horn of Africa, but were expelled in 1941. This provides a unique case study to examine the long-run effect of cheap transport networks on the concentration of economic activity in developing countries. The results show that cells located next to Italian paved roads are significantly richer today and that the relationship is causal. Persistence is explained by a combination of direct and indirect mechanisms: colonial roads attracted economic activity through lower transport costs until 1960. After that date, the advantage of treated locations persisted only indirectly through increasing returns to scale.

1 Introduction

Roads are important drivers of economic growth in developing countries: they facilitate market integration across regions and create urbanization (Storeygard, 2016), foster agglomeration (Faber, 2014) and improve rural living standards (Stifel, Minten, and Koru, 2016).[1] Compared to railroads, roads are cheaper to build and to maintain. Colonial governments relied on them extensively, in combination with railroads, to facilitate the control and the economic development of their overseas territories (Jedwab and Moradi, 2016; Havinden and Meredith, 1993). After decolonization, roads remained the single most important transportation facility and, still today, absorb significant portions of African governments' budgets (Burgess et al., 2015) and funds from international development programmes (Buys, Deichmann, and Wheeler, 2006). Yet, unlike railroads which have received growing attention (Donaldson, 2010; Chaves, Engerman, and Robinson, 2014; Jedwab and Moradi, 2016; Jedwab, Kerby, and Moradi, 2017), no empirical work exists on the effect of colonial roads for development and sub-regional inequality, nor on their interaction with post-colonial investments and growth.

In this paper, I look at the construction of Italian colonial roads built in the Horn of Africa between 1935 and 1940 to answer two main questions: *What was the impact of colonial paved roads on development and economic concentration in the long-run? What mechanisms explain this effect?* The case study presents some unique features which make it akin to a natural experiment: on the one hand, all major Italian roads were built or planned with a military and strategic purpose in mind, in the context of the Italian-Ethiopian war of 1935-6, and aimed at connecting major cities in straight lines to facilitate the conquest of the country; this makes the endogenous placement of these lines less problematic. On the other hand, the Italian occupation lasted for just five years, which allows us to isolate more effectively the role of infrastructure from other confounding factors that could be associated with a prolonged colonial rule.

To make the results comparable with the related literature on colonial transportation, the paper follows Jedwab and Moradi's (2016) methodology: I construct a gridded dataset containing 15,550 cells, each measuring 11x11km (0.1x0.1 decimal degrees), which covers Eritrea, Ethiopia and Italian Somalia. For each cell, distances from colonial roads and facilities are calculated and matched with population and light density at night, in 2000

---

[1]I have started to work on this topic during my MPhil at the University of Oxford. This paper is an updated version of my MPhil's thesis (Bertazzini, 2015)

and 2010 respectively, which are two proxies for economic development (Ciccone and Hall, 1996; Henderson, Storeygard, and Weil, 2012). I start by performing regression analysis to test whether a robust correlation between proximity to colonial roads and contemporary economic development exists: the estimates show that cells located within 10km of Italian roads are significantly more populous and luminous today, once other factors are accounted for. I corroborate the grid-level analysis with individual level data from the 2011 Demographic and Health Survey of Ethiopia (DHS) and find that people living in treated locations are markedly richer, whether a wealth index or types of employment are introduced as dependent variables.

Since endogeneity is likely to undermine the validity of the results, three main robustness checks are carried out: distance from roads that were planned but never built (due to the outbreak of WWII) is employed as a first placebo test. As a second step, the effect of surfaced colonial roads is compared with gravel tracks that were not paved before WWII, which provides a second falsification exercise. If the effect of Italian roads was causal and due to a specific colonial investment (surfaced roads), we should observe no effect along placebo lines. Third, colonial roads are instrumented by telegraphic and telephonic lines that predated the Italian occupation. These predict the location of Italian roads well as, on the one hand, they approximate the least cost path between cities and, on the other, they influenced the selection of the network's poles. Placebo lines have no effect in the long run and the instrumented coefficients are positive and significant, which underscores a causal link between the colonial road network and contemporary development.

Finally, I study the channels of persistence: the effect of colonial roads on population density had already emerged in 1940, immediately after completion, and reached its peak in 1970 and moderately declined and stabilized in the 1980s and 1990s, around the levels measured in 2000. The jump of the coefficient at the end of the 1930s was due to new settlement along the roads and not to population's relocation from surrounding areas. In order to disentangle the direct effect of roads through lower transport costs, from the indirect one through increased market potential, the conceptual framework of Krugman's New Economic Geography (NEG) model (Krugman, 1991) is employed empirically: the results show that the direct effect of colonial roads only mattered during the 1930s and 1950s while, in the following period, internal market potential captures the effect completely, which suggests that roads caused path-dependence indirectly through increasing returns to scale. In other words, Italian roads produced a resilient spatial equilibrium by 1960, which was hardly modified by subsequent institutional changes.

Additional tests highlight how the effect of roads on population density was moderately reinforced by the coordination of post-colonial investments and, specifically, by the renovation of colonial roads during the 1950s.

While emphasizing the long-lasting effect of colonial roads in shaping spatial equilibria, this article contributes to the debate on the role played by colonialism in shaping contemporary development and inequality. Scholars have emphasized the role of institutions (Acemoglu, Johnson, and Robinson, 2001; Banerjee and Iyer, 2005; Fors and Olsson, 2007; Iyer, 2010; Alesina, Easterly, and Matuszeski, 2011; Bruhn and Gallego, 2012), taxation (Frankema, 2010; Gardner, 2012; Huillery, 2014), human capital (Glaeser et al., 2004; Huillery, 2009; Bolt and Bezemer, 2009; Cogneau and Moradi, 2014; Woodberry, 2012) and physical transportation infrastructure (Andrabi and Kuehlwein, 2010; Donaldson, 2010; Burgess and Donaldson, 2010), but no study has looked at the effect of roads.

With respect to the literature on the spatial impact of colonial railroads (Jedwab and Moradi, 2016; Jedwab, Kerby, and Moradi, 2017), which is the most closely related to this paper, I show how roads had a similar effect despite the substantial differences in sunken and maintenance costs and the demise of railroads after the end of colonialism: this suggests that a transport network's maintenance costs should be proportionate to the needs and capacity of the local economy, if they are to remain operational for long. By explicitly incorporating market potential and post-colonial investments into the analysis, I formally test the relative importance of the various persistence mechanisms that the literature has suggested and give evidence of how direct and indirect forces operated with different intensities over time.

My work also confirms that shocks in transportation costs can persistently alter spatial concentration, thus adding to the debate on the relative importance of geography (Gallup, Sachs, and Mellinger, 1999; Diamond, 1997; Rappaport and Sachs, 2003; Nunn and Puga, 2012; Fenske and Kala, 2015), resilience of location fundamentals (Davis and Weinstein, 2002; Miguel and Roland, 2011) and transport networks and man-made facilities (Bleakley and Lin, 2012; Fafchamps, Koelle, and Shilpi, 2017) in shaping spatial equilibria. The existing contributions that look at transport networks either focus on railroads (Berger and Enflo, 2017; Donaldson and Hornbeck, 2016), or consider top-income countries (Baum-Snow, 2007; Michaels, 2008; Percoco, 2016). Those few studies that consider roads in comparable settings, focus on market integration effects (Limao and Venables, 2001; Storeygard, 2016), or find contradictory results for the core-periphery structure, which calls for further research (Banerjee, Duflo, and Qian, 2012; Faber, 2014). My results underline how infrastructural quality

should be carefully considered in the early phase of construction, as this can have substantially different long-term consequences (only paved roads guaranteed an advantage during the rainy season), in line with what Fourie (2008) suggests. In fact, although there is a certain margin for targeted investments to adjust the spatial equilibrium in the short- and medium-run (Jedwab, Kerby, and Moradi, 2017), increasing returns to scale outweigh other forces and, consequently, even a small initial difference can produce substantially different long-term outcomes.

With a narrower focus, this paper also relates to literature on Italian colonialism, which has often dismissed the importance of Italian colonial investments in general and, specifically, for the post-colonial history of the Horn of Africa (Del Boca, 1992; Bekele, 1995; Labanca, 2002; Podestà, 2004).

The remainder of the paper is structured as follows. Section 2 provides a brief description of the historical framework. Sections 3 illustrates the empirical strategy and the data. The main results and robustness checks are presented in sections 4 and 5, respectively. Section 6 studies the mechanism. Section 7 concludes.

2 Historical framework. Italian road construction in the Horn of Africa

Before Mussolini's decision to annex the Ethiopian Empire to Eritrea and Somalia, the Horn of Africa had seen little investment in transportation. The Italian colonies were too poor to justify extensive road building, while the Ethiopian monarchs had only managed to construct a few gravel roads concentrated around Addis Ababa (Emmenegger, 2012).[2]

The Italo-Ethiopian war completely changed the transportation dynamics in the area: between 1934 and 1935, the Italian military engineers started building new gravel and paved tracks that connected the core of Eritrea and Somalia with the Ethiopian border to allow troops and supplies to reach the front. As soon as new parts of the Ethiopian territory were conquered during the military campaign in 1935 and 1936, these original tracks were extended to allow further expansion (Badoglio, 1937; Graziani, 1938).

When Addis Ababa was finally occupied on 5 May 1936, just a small portion of the Ethiopian territory was actually under Italian control and the Italian garrisons only patrolled major cities such as Addis Ababa, Jimma,

---

[2]For a more detailed analysis of the historical framework, see appendix C

Gondar, Dessié, Harar, together with few minor centres. Different groups of Ethiopian patriots resisted in peripheral parts of the country and were even able to carry out offensives near the main cities (Dominioni, 2008). Moreover, the rainy season was approaching (the first heavy rains were due between November and December) and it would have been impossible for the Italian army to preserve the main strongholds once the supply chain was interrupted by muddy gravel roads (Del Boca, 1979).

[**Figure 1 about here**]

In this desperate situation, the Italian government ordered the implementation of a massive road construction project, connecting the main cities with asphalt roads. Before the start of the rainy season, the main sections of the Asmara-Gondar, Asmara-Dessié-Addis Ababa, Addis Ababa-Jimma, Addis Ababa-Lechemti and Addis Ababa-Harar-Mogadishu roads were completed. In the following years (1937-1939) some other paved roads were built, namely the Addis Ababa-Debra Tabor and Assab-Dessié roads.[3] Other secondary arteries were started but left unfinished or were never started because of the drastic cut in the budget due to the WWII preparations and ultimately the outbreak of the war (Del Boca, 1986), which provides a useful falsification exercise. These major roads, together with those previously built in Eritrea and Somalia and an array of secondary unpaved tracks formed, by 1940, an extensive network which totalled around 6,000km, of which 3,500km was paved and 2,500km unpaved (Leul, Petros, and Kebede, 2008). In the empirical analysis I will focus on the paved roads.

In 1941, the last Italian strongholds surrendered to the combined British and Ethiopian forces. During the war, large parts of the network's surface, together with major bridges and tunnels, were destroyed (Del Boca, 1986) and no comprehensive maintenance was undertaken until 1951, due to lack of capital and economic stagnation. After this date, several infrastructural plans were implemented: initially (during the 1950s), Italian roads were refurbished and no other major arteries were built, whereas after 1960 several plans based on rural road development started to be implemented, which progressively took away the substantial monopoly in transportation that Italian roads had enjoyed until 1960 (Emmenegger, 2012).

This brief description of the history of Italian road building in the area highlights two key features for my identification strategy: on the one hand, Italian roads were designed to serve very clear strategic military needs, whereas their placement was almost completely random with respect to

---

[3]Various (1937-1943). *Gli Annali dell'Africa Italiana*. Ed. by Angelo Piccioli. A. Mondadori

pre-colonial tracks or patterns of settlement. On the other, as opposed to railroads (which are more capital intensive and socially costly to maintain) Italian roads were re-built and kept operational after the end of the Italian occupation, which suggests substantially different dynamics in the post-colonial era, compared to colonial railroads in Kenya or Ghana (Jedwab and Moradi, 2016; Jedwab, Kerby, and Moradi, 2017).

3 Empirical strategy

3.1 Econometric specification

The first question considered in this paper is whether the Italian road network had a long-term impact on the concentration of economic activity of Eritrea, Somalia and Ethiopia. To address this issue, I employ econometric analysis on a spatially explicit gridded dataset, composed by 15,550 cells, each measuring 0.1x0.1 degrees (11x11km). For each observation, I calculate the average value of my dependent variables of interest (population density in 2000 and luminosity in 2010), together with the Euclidean distance of each cell's centroid from colonial roads. I also add various pre-colonial, colonial, environmental and geographical controls.

I start by checking if a robust correlation between proximity to colonial roads and contemporary levels of development is in place by estimating model 1 through OLS:

$$
\begin{aligned}
y_{i,e} = \alpha + &\beta_1 DColonialRoad_{0-10km,i,e} \\
&+\beta_2 DColonialRoad_{10-20km,i,e} \\
&+x'_{i,e}\gamma + \delta_e + \epsilon_{i,e}
\end{aligned}
\tag{1}
$$

Where $y_{i,e}$ represents the two main dependent variables, population density in 2000 (normalized by standard deviations) and luminosity in 2010, both measured in cell i and ethnic group e (Ciccone and Hall, 1996; Henderson, Storeygard, and Weil, 2012; Michalopoulos and Papaioannou, 2013). I consider population density my main dependent variable as this is the only proxy available for the entire period.

$DColonialRoad_{0-10km,i,e}$ and $DColonialRoad_{10-20km,i,e}$ are the explanatory variables of interest, two dummies that equal 1 if cell i in ethnic group e is located within a 10km radius from colonial paved roads and between 10 and 20km, respectively. If colonial infrastructure has affected economic concentration and income distribution persistently, I expect colonial roads to

have a positive and statistically significant effect on my dependent variables. (β_1, $\beta_2 > 0$).

$x'_{i,c}$ is a vector representing the full set of controls. These can be divided into three groups, namely pre-colonial, colonial and geographical & environmental controls. Pre-colonial controls include dummies for being located within 10 and between 10 and 20 km from 1935 tracks and population densities in 1920 and 1930. Controls for other colonial investments include 10km cut-off dummies for colonial facilities whose role has been emphasized by recent literature: railroads (Donaldson, 2010; Jedwab and Moradi, 2016), major and secondary hospitals (Huillery, 2009), schools for Africans (including missions) (Woodberry, 2012; Cagé and Rueda, 2016), colonial provincial capitals, Italian unpaved tracks and a dummy for being in an area of mining interest during the colonial era (Bruhn and Gallego, 2012). Finally, I introduce a standard set of geographical controls that might matter for development: latitude, longitude, altitude, annual average rainfall, FAO land suitability, a dummy for being within 10km of perennial rivers and less than 30km away from the coast. Additionally, I include malaria rates (Weil, 2014), ruggedness (Nunn and Puga, 2012) and temperature (Fenske and Kala, 2015). Table 1 reports main descriptive statistics for all the variables included in specification 1.

Table 1 about here

δ_e is my baseline set of fixed effects, the 53 ethnic groups in which Oriental Italian Africa was divided, according to Murdock (1967). These account for potential unobservable characteristics at the tribe level and it is particularly important to capture pre-colonial features (Michalopoulos and Papaioannou, 2013). $\epsilon_{i,e}$ is the error term. In all the specifications, in order to make my estimates robust to spatial correlation, I cluster the standard errors by 202 colonial provinces, called "Residenze", the smallest administrative units in the colonial period. Conley errors, robust to spatial correlation, are also calculated.

3.2 Robustness checks

3.2.1 *Higher living standards, or overpopulation?*

Although population and light densities are good proxies for economic concentration and income largely employed in the literature (Ciccone and Hall, 1996; Henderson, Storeygard, and Weil, 2012), one might still be worried that the estimates could reflect overpopulation and not higher levels of income. Therefore, I test my working hypothesis on the 2011 male individual

Demographic and Health Survey of Ethiopia.

For each individual, I select information on household's total wealth (an index that factors in all assets owned by the household, such as type of house, electric utensils, means of transport etc.) and breadwinner's profession[4] for 13,528 individuals, distributed across 649 villages.[5] For each village, I calculate the same set of explanatory and control variables as for equation 1 and I re-estimate the model with wealth index (OLS) and type of profession (Probit) as left-hand side variables. If treated locations are richer and not just more populated, I expect Ethiopians living next to Italian roads to show significantly higher levels of wealth and to be employed in better remunerated jobs.

3.2.2 *Omitted variable bias*

Given the empirical set-up, endogeneity is likely to bias the estimates and drive spurious correlations. Italian roads, in fact, might have been located following unobservable historical characteristics. I implement several robustness checks to deal with this issue (See figure 5 in the online appendix).

First, I test whether proximity to projected surfaced roads, which were not completed due to the outbreak of WWII, shows levels of contemporary population density similar to treated locations. If pre-colonial characteristics have determined the location of Italian motorways, I expect to observe a positive and persistent effect of these lines.

Second, I perform another placebo test that compares the effect of Italian paved roads with that of unpaved colonial tracks. This exercise tests the effect of a specific technological innovation (asphalt surface): if the same effect could be measured for both types of roads and quality did not matter (Fourie, 2008), it would be difficult to argue for a causal effect of paved roads and endogenous placement would seem more likely.

Third, I implement an instrumental variable estimation. Italians aimed to connect major cities while minimizing cost and time and without having a clear idea of the Ethiopian urban network. I exploit these two facts and instrument colonial roads for pre-1936 telegraphic and telephonic lines, as depicted in Italian military maps: the extent to which Italians considered a city important relied on their knowledge of Ethiopia, which in turn largely depended on the pre-colonial communication network. Furthermore, transmission lines tended to follow straight lines. This strategy therefore, proxies

---

[4]Substantial average salary gaps between sectors exist in Ethiopia, see http://www.wageindicator.org/main/salary/minimum-wage/ethiopia

[5]In theory there are 650 villages but one has incorrect coordinates

both selection of targets and least cost path. I expect my instrument to be a strong predictor of road location and, for the exclusion restrictions to hold, not to influence long-term development through channels other than colonial roads.

3.2.3 Data

The outlined analysis required a novel dataset containing information on a vast array of dependent, historical and geographical & environmental variables. While a detailed description of each variable and data source can be found in part A of the appendix, in this section, I will focus on population density data and Italian colonial investments, my main dependent and explanatory variables.

The main dependent variable, employed for the baseline estimation, is decennial population density for the period 1920-2000. For the years 1960, 1970, 1980, 1990 and 2000 (last available), this is obtainable through the Unep/Grid Sioux Falls database in raster format. This source offers reliable information on population distribution for this period, as it is constructed from granular level demographic statistics and surveys. Conversely, for the years 1920, 1930, 1940 and 1950, micro statistical surveys of population density do not exist; as a first step to solve this issue, I employ reconstructed population densities from the History Database of Global Environment (HYDE), which combines historical population density data, cropland and pasture statistics to build population distribution in raster format (Klein Goldewijk et al., 2011). Given the limited reliability of the HYDE data, I back up the available population density information with the location of villages, towns and cities. Patterns of agglomeration and population concentration largely correlate, although they technically measure two distinct features. Different cartographic sources made it possible to digitize the location of all villages in Eritrea, Somalia and Ethiopia for the years 1925, 1935, 1937, 1939 and 1958. Additionally, I was able extend this to the years 1981, 1988 and 2000 for towns and cities with a population larger than 4,000 inhabitants. The coverage of cartographic material varies significantly over time, but it consistently covers Eritrea and Ethiopia for the entire period.

Figure 2 about here

The principal sources for colonial investments are the 3rd and 4th issues of 1939 and the 1st and 2nd issues of 1940 of "Gli Annali dell'Africa Italiana", a review edited by the Italian Ministry of the Colonies, which describes the public investments implemented by the Italian Government in Oriental Italian Africa. The location of colonial paved and unpaved roads in 1939

comes from a map contained in the 4th issue of 1939. This informs about which road sections had been completed by 1939 and which were under construction, and it also describes the roads surface, whether they were paved or unpaved (employed as both a control and a placebo test). From issue 1 of 1940, I have collected data on the location of colonial schools for natives. Locations of Italian hospitals and clinics have been obtained by geo-referencing a series of maps from the same volume. A variety of other features from the colonial period were collected from diverse sources, including location of colonial railroads, colonial administrative capitals & boundaries and areas of mining potential (see data appendix for a detailed description of the sources). Finally, pre-colonial tracks come from Achille Dardano's map, published in 1935, to the best of my knowledge, the most accurate map ever drawn for pre-1936 Ethiopia. Historical ethnic boundaries are obtained from Murdock's atlas (Murdock, 1967; Nunn, 2008).

4   Results

Are historically better-connected areas richer today? In order to answer the core question of the paper, I begin by estimating equation 1 through OLS. Table 2 reports the estimates from the main specification: standardized population density (*z-scores*) in 2000 is employed as dependent variable in columns 1 to 4, whereas the results for Log light density at night are displayed in columns 5 and 6. For all specifications, ethnic fixed effects are included and standard errors are clustered by the 202 colonial provinces.[6] The estimation in column 1 has no controls, column 2 includes pre-colonial factors, columns 4 and 5 report colonial and geographical controls, respectively. Column 5 displays the restricted model with light density as dependent variable, which is conditioned on population density in 2000, in column 6.

Table 2 about here

Column 4 shows that cells located within 10km of colonial paved roads are, on average, 0.2 standard deviations more densely populated, once the full set of controls has been included. This effect is economically very large, as it implies a statistically significant deviation from the mean of about 20% of the standard deviation, or 29 points. This implies that treated locations have an average population density of about 73 inhabitants per square kilometre, against an average, for the whole area, of roughly 44 (see table 1). The introduction of controls such as pre-colonial population density

---

[6]The baseline estimation without fixed effects is reported in column 6 of table A2

and distance from pre-colonial tracks in column 2 does not significantly reduce the magnitude of the coefficient, partially confirming that pre-colonial characteristics did not play a fundamental role in the location of colonial roads. By contrast, the introduction of colonial controls, in column 3, and of geographical controls, in column 4, reduces the coefficient by about 40% and 20% respectively, pinpointing the importance of other colonial activities and environmental factors in explaining population distribution in 2000.

As underlined above, population density is a good approximation if one wants to estimate the concentration of economic activity, but leaves room for uncertainty when it comes to disentangling the latter from actual living standards. However, the replication of the previous exercise with light density at night will make it possible to say more about levels of development.

Column 5 shows that, conditional on the full set of controls, cells located within 10km of Italian motorways are, on average, about 30% more luminous than the control group. The coefficient remains positive, roughly of the same magnitude and significant at the 1% level, once contemporary population density is added to the equation, as is visible in column 6. In other words, treated cells are not more luminous at night only because more people live in these areas, but they are also significantly richer once population density is accounted for.

5 Robustness checks

5.1 Are treated areas richer, or just more populated?

Although population and light density are proxies that are widely employed in the literature, one might still be sceptical about the extent to which these reflect people's income and not just uneven population distribution, especially in the context of developing countries. In order to tackle this potential issue, I corroborate my results by employing individual data for Ethiopian men from the 2011 Demographic and Health Survey of Ethiopia, which provides direct measures of income.

In table 3, I employ the individual-level dataset to re-estimate equation 1, this time with 13,528 Ethiopian males as units of observation. Panel A reports OLS estimates with wealth index as dependent variable. Panel B shows Probit estimations with binary variables for different types of profession as left-hand side variables. In panel B, marginal effects are reported so that the coefficients can be interpreted as percentages.

Table 3 about here

Consistent with the main results, the data show that people living, at present, in historically better-connected areas are significantly richer: Ethiopian males located less than 10km away from Italian roads have an average wealth score index that is about 0.5 points larger than the comparison group and significant at the 1% level (column 4). Similarly, panel B shows that treated individuals have better chances of being employed in highly remunerated jobs in the tertiary sector: for instance, they are more likely to be professionals (doctors, technicians or teachers), to be employed in services and to be skilled workers (columns 1, 2 and 4, respectively). They are also 13% less likely to be employed in agriculture.

In conclusion, tables 2 and 3 provide evidence that historically better-connected cells are not only more populous today, but also significantly richer.

5.2 Omitted variable bias

Despite the consistency of the estimates across different development indicators and their robustness to the full set of controls, endogeneity could bias my results. In other words, regression analysis cannot rule out the possibility of unobservables driving a spurious correlation between proximity to colonial roads and higher levels of development. To address this issue, I focus on population density, my main dependent variable, and implement a series of falsification tests and IV treatments, reported in table 4 (See figure 5 in the online appendix). All columns include the full set of controls from equation 1 and ethnic fixed effects.

5.2.1 *Placebo tests*

First, I implement two falsification tests that employ cut-off distances from projected paved roads and unpaved colonial tracks. I expect placebo lines to have no effect on population density in the long run, unless the placement of colonial roads was endogenous and related to unobservable pre-colonial factors.

Column 2 shows that cells located next to projected Italian roads that were planned but not built due to WWII, are not significantly more populous today. As the choice of which roads were constructed first was based solely upon strategic military considerations during the Italo-Ethiopian war of 1935-6 and in its immediate aftermath, we can take this as strong evidence that selected locations were not significantly different from the rest, before the construction of colonial infrastructure was completed. This in turn pro-

vides strong evidence that a causal relationship between Italian roads and economic development is in place.

Table 4 about here

Similarly, if one looks at the effect of unpaved colonial tracks, in column 3, no statistically significant effect can be measured within 10km of these lines. Interestingly, despite being very small, the coefficient for the 10-20km cut-off is positive and significant, which suggests that these secondary tracks were probably located in areas that had more economic potential to begin with. This second placebo exercise provides evidence about the fact that only roads of better quality that were usable during the rainy season had a persistent effect over time, whereas unpaved tracks failed to generate agglomeration, despite being allocated to relatively more populous areas.

5.2.2 *IV estimations*

As a second robustness strategy, I exploit the characteristics of the road network's construction process to implement an IV treatment, where I instrument colonial paved roads for pre-1936 telegraphic and telephonic lines. The idea behind this test is, on the one hand, to approximate the least cost path between targets (Italian planners wanted to connect major cities as quickly as possible) and, on the other, to capture the selection of targeted poles (how they chose the cities that were to be connected).

Column 4 reports the first stage and shows that the selected instrument is a good predictor for the location of colonial paved tracks. The second stage in column 5 shows a positive and statistically significant coefficient at the 5% level, which provides further evidence of the causal impact of Italian surfaced roads. By contrast, the small coefficient in the first stage, together with the large one (four times larger than the baseline) in the second stage, suggests that the instrument could be weak. Nevertheless, the Kleibergen-Paap F statistic is relatively large, which indicates that although the IV estimation might not be precise enough to correct for measurement error, it still provides strong evidence of the sign and significance of the effect.

Several additional robustness checks are reported in section B of the online appendix (Tables A2, A3 and A4).

6   Channels of persistence

The results presented so far have shown that the construction of Italian roads rendered neighbouring cells substantially richer. But this raises more

questions about the mechanisms: *When did the effect emerge? What mechanisms explain the persistent effect of colonial roads over time?* In other words: *Why are these areas still richer today?*

Understanding the dynamics triggered by colonial roads and, more generally, by cheap transportation infrastructure in developing regions, will not only make the history of the Horn of Africa more comprehensible but it will also deepen our knowledge of the determinants of sub-regional development for all countries that have experienced technological revolutions in transportation, similar to the former Italian colonies.

In order to explore the historical patterns of agglomeration, I will employ population density across the different estimations, as this is the only proxy for economic development for which data exists for the entire period.

6.1   The effect over time

I start by describing the transformation of the effect over time. I estimate equation 1 with normalised population density between 1920 and 2000 as dependent variables. Coefficients with a 5% confidence interval bands for population density (blue dots) are reported in figure 3. As expected, colonial roads have no significant effect on population density in 1920 and 1930. The effect of Italian roads emerged already in 1940, which suggests an extremely rapid adaptation of the local economy to the construction of new roads. The coefficient jumps from 0.05 to roughly 0.22 and becomes statistically significant at the 1% level. This is consistent with the available qualitative evidence, which portrays rapid agglomeration alongside main roads, already during the phase of construction (e.g. creation of barracks for the workers, grocery stores, gas stations) (Cobolli-Gigli, 1938; Del Boca, 1986) and immediately after completion. In this early phase, new commercial activities and villages mushroomed along Italian roads and some urban centres were relocated to better-connected areas (Benti, 2016). The effect declines moderately between 1940 and 1950, during a decade of financial hardship for the independent Ethiopian government, which was unable to implement any relevant infrastructural development programme (Bekele, 1995).

Figure 3 about here

The coefficient increases significantly between 1950 and 1960, with proximity to paved roads translating into roughly 0.3 standard deviations higher population density in 1960. This coincided with the first period of moderate growth (average of 3% per annum) for the regional economy (Bekele, 1995) and, interestingly, with the "First Highway Plan", launched by the Imperial Ethiopian government between 1951 and 1957, which refurbished the

14

existing Italian network (Emmenegger, 2012). The effect reached its peak in 1970 and then moderately decreased during the 1970s. By 1980, it stabilized to around the same level of 2000. Important construction plans of both major and rural roads were undertaken between 1961 and 1974 (5 distinct Highway Plans) (Emmenegger, 2012; Henze, 2000), so that by 1970, colonial roads had lost their monopoly in terms of transportation networks. In the 1980s and 1990s, during a period of major political changes and economic setback (Henze, 2000), the effect remained stable. During these decades, which witnessed the violent civil war between the supporters of the Communist government (Derg) and the more liberal EPRDF, which seized power in 1991, major infrastructural improvements were achieved in rural roads construction, which became a priority for both governments (Emmenegger, 2012).

The main concern about the coefficients for population density, reported in figure 3, is the fact that it combines two data sources with different characteristics: reconstructed population densities from the HYDE project, between 1920 and 1950, and granular data from Unep after 1950. As more reliable population density estimates for the earlier period are, to the best of my knowledge, not available, I exploit agglomeration data from a variety of cartographic sources (See data appendix for details), to verify whether the jump in the coefficient between 1930 and 1940 is effectively due to Italian road construction and not to population density reconstruction.

OLS coefficients with binary variables for a cell having at least one agglomeration (all villages and cities above 100 inhabitants) in 1925, 1935, 1938 and 1939 respectively, are displayed by green dots in figure 3. Cells located within 10 km of Italian roads do not display a significantly higher concentration of villages in 1925 or 1935. Conversely, and consistent with the results for population density, the coefficient jumps from 0, in 1935, to 0.05 and 0.12 in 1937 and 1939: this confirms that the effect emerged immediately after the completion of Italian roads, with the timing suggested by population density data.

6.2   Relocation, or population growth?

The previous section has shown that by 1937 more economic activity was already concentrated next to Italian roads, just one year after the Italian occupation of Ethiopia and two years after the construction of the first military roads. This could have happened through either the relocation of the local population, which moved closer to the better-connected areas from neighbouring cells (Berger and Enflo, 2017), or due to the establishment of

foreign migrants and firms, who either arrived from abroad, relocated from the main cities or from distant areas (Jedwab, Kerby, and Moradi, 2017).

Table 5 about here

As this difference could alter the general interpretation of the results, I test which one of these mechanisms is in place in the early phase after construction: in columns 2 to 5 of table 5, I estimate model 1 with the 1939 dummy for agglomeration as dependent variable, this time shifting the cut-off distance of the explanatory variable of interest progressively away from Italian roads. If relocation was in place, cells located between 20 and 60 km from Italian roads should show a lower average concentration of villages in 1939. The estimates show no significant reduction in agglomeration.

Another way of dealing with this issue is to reduce the comparison group to cells located farther away from colonial infrastructure (Berger and Enflo, 2017). If relocation is in place, the effect should fade away the farther the comparison group is shifted. Columns 6 to 10 show that this is not the case, which implies that the estimated effect was due to the settlement of new firms and individuals.

6.3   Explaining path dependence

6.3.1   Conceptual framework

If the monopoly in transportation provided by Italian infrastructure fully explains why people and firms located next to colonial roads between 1936 and 1940, accounting for the persistent effect of colonial roads until today is less straightforward.

The literature has suggested four possible mechanisms for the persistence of spatial equilibria shaped by revolutions in transportation technology: first, the effect of roads could have persisted directly, if Italian infrastructure has retained partially, or in full, a comparative advantage in transport costs. Second, the present-day effect could be due to a slow depreciation of sunk investments that were allocated to treated areas during the colonial period, through either public or private investments (Bleakley and Lin, 2012). Third, agglomeration and local increasing returns to scale might have consolidated the initial higher concentration triggered by lower transportation costs, through endogenous growth linked to larger market potential (Krugman, 1991; Krugman, 1993; Rosés, 2003). Finally, colonial infrastructure might have helped to solve the spatial coordination of post-colonial investments, this way determining a disproportionate allocation of post-colonial facilities that could have reinforced the initial advantage (Jed-

wab, Kerby, and Moradi, 2017).

In this section, I focus on disentangling the relative importance of these mechanisms between 1940 and 2000, as this makes a substantial difference for the interpretation of the results: if we are observing sunk investment depreciation, this would imply that the region is gradually returning to the pre-colonial spatial equilibrium determined by location fundamentals (Davis and Weinstein, 2002; Miguel and Roland, 2011). This would suggest that colonial roads could change the spatial distribution of economic activity only temporarily. By contrast, if other dynamics are in place (such as co-ordination and increasing returns to scale) this would not only point to the capacity of early colonial investments in cheap transport infrastructure to alter economic concentration permanently, but also to the fact that their spillover effects go beyond a simple reduction in transportation costs.

I employ the conceptual framework of the Krugman-Fujita New Economic Geography (NEG) model, which defines economic concentration as a function of amenities (non-tradeable location fundamentals) and increasing returns to scale ($Ec = f(A; IRS)$) to test the relative importance of direct and indirect mechanisms. As increasing returns to scale are in turn a function of transportation costs and market potential - $IRS = f(InMP; ExMP)$, we can define economic concentration in cell i as $Ec_i = f(A_i; InMP_i; ExMP_i)$ (Ploeckl, 2012; Martinez-Galarraga, 2012). By employing the methodology proposed by Martinelli (2014) (based on the more sophisticated version employed by Donaldson and Hornbeck (2016)), we can define economic concentration in cell i as

$$Ec_i = A_i * (2\frac{GDP_i}{Area_i} + \sum_{j\neq i} \frac{GDP_j}{Distance_j}) \tag{2}$$

Where the Euclidean distance between cells approximates transport costs. Based on equation 2, it is possible to adapt my empirical strategy to test for the relative importance of direct and indirect mechanisms, simply by plugging estimates for internal and external market potential into equation 1:

$$\begin{aligned} Popd_{i,e,t} = \alpha &+ \beta_1 DColonialRoad_{0-10km,e,i} \\ &+\beta_2 DColonialRoad_{10-20km,e,i} \\ &+\beta_3 InMP_{i,e,t-1} + \beta_4 ExMP_{i,e,t-1} \\ &+x'_{i,e} + \delta_{e,i} + \epsilon_{i,e} \end{aligned} \tag{3}$$

Where $InMP_{i,s,t-1}$ and $ExMP_{i,s,t-1}$ are internal and external market potentials for cell i and year $t-1$. In order to construct these estimates,

due to unavailable historical micro-level data on GDP, I employ data on agglomeration (1930-1950) and population density (1960-1990) in year $t-1$ to calculate market potential estimates as in equation 2 (See data appendix for details). Non-tradeable location fundamentals (A_i in equation 2) are included in the control vector $x'_{i,e}$, which represents geographical characteristics as well as the full set of controls as in equation 1. Plugging 10km cut-off distances from post-colonial roads, hospitals and schools will capture the effect of coordination of post-colonial facilities.

This strategy allows increasing returns to scale to be incorporated explicitly, and represents a noticeable improvement compared to related works, which do not formally test the relative importance of the different channels.

6.3.2 Sunk investments or increasing returns to scale?

The first possibility I want to control for is whether path-dependence is due to sunk investments depreciation. The baseline equation (column 4, table 2) already controls for distance from colonial investments, such as hospitals, schools, railroad lines and district capitals. This seems to indicate that sunk investments alone are not driving path dependence and that increasing returns to scale are in place (Bleakley and Lin, 2012; Berger and Enflo, 2017). However, it is also possible that other sunk investments are not observed or that included sunk investments are not fully depreciated by the end of the period and that the effect is doomed to vanish in the near future. If this were the case, two trends should appear in the results: on the one hand, we would expect the effect of roads on population density and urban agglomeration to decrease over time. On the other, we anticipate that the share of the effect explained by included sunk investments should increase consistently over the same period.

I test this possibility in figure 5. The top panel plots the effect of roads over time with both unrestricted (green dots) and restricted model (equation 1) with respect to sunk investments (other colonial investments), whereas the bottom panel reports estimates from the same regression, but with presence of urban agglomerations (villages and cities with more than 4,000 inhabitants) as dependent variable.

Figure 4 about here

As is evident, the effect on population density stabilizes at around 0.2 standard deviations from 1980 onwards, simultaneously to the relative effect of sunk investments, which flattens at about 50% of the unconditional effect. This strongly suggests that path dependency in 2000 is not the result of a slow depreciation of other colonial facilities, but rather a new stable equi-

librium that was reached in 1980 and that provided a steady advantage to the treated locations. Interestingly, if one looks at the coefficients for urban agglomerations (more than 4,000 inhabitants) in the bottom panel, the effect is increasing over time and takes off between 1960 and 1980, in parallel with the convergence towards an equilibrium in population density. The relative share of the effect explained by sunk investments actually reduces over time, which strongly suggests that increasing returns are in place. All together, this shows that persistence is not due to depreciation, but rather to increasing returns and urbanization[7], similar to what the literature has suggested for railroads.

6.3.3 Coordination of post-colonial investments?

Before disentangling the relative importance of the direct effect of Italian roads from endogenous growth led by the initial increase in market potential and population, I test whether colonial roads fostered a relative concentration of post-colonial investments in treated locations, higher than population density levels. If treated areas received more investments because they solved the resource allocation problem, this could indirectly reinforce the effect and should be accounted for.

Conditional on population density, I run Probit regression analysis to test whether colonial roads attracted more roads, hospitals and schools during the post-colonial period. Marginal effects are reported in table 6, so that coefficients can be interpreted as percentage probabilities. Columns 1 to 4 and 7 show that treated locations were significantly more likely to be endowed with paved roads in 1955, 1962, 1973, 1988 and 2000, respectively. The effects are large and significant at the 1% level and span from 10% in 1955 to 11% in 2000 with a peak of 30% in 1988: this underscores how historically better-connected cells continued to receive more investments in infrastructure even after the end of Italian colonization. Similarly, schools were 5% more likely to be allocated to treated cells during the 1980s (columns 5 and 6).

Table 6 gives strong evidence that colonial infrastructure was indeed affecting the spatial allocation of post-colonial investments, a feature that could reinforce the effect of Italian infrastructure over time.

Table 6 about here

---

[7]The fact that individual level effects in table 3 disappear when the DHS survey's dummy for urban environment is introduced is further evidence that urbanization played a role in the persistence of the effect up to the present (not shown).

6.3.4 Direct effect, increasing returns to scale, or coordination?

In table 7, I test the relative importance of the different mechanisms of persistence, namely direct effect through lower transport costs, indirect effect through increasing returns to scale and the coordination of post-colonial investments, as outlined in section 6.3.1. To do so, I estimate equation 3 conditional on 10km cut-off distances from post-colonial investments (dependent variables, table 6). In other words, I add post-colonial investments and market potential estimates over-time as controls to equation 1, in a cross-sectional dynamic framework.

The results show that the direct effect of Italian roads was in place only during two periods, namely in the 1930s and 1950s. The emergence of the effect in 1940 (column 1) is fully accounted for by the direct effect of colonial roads, which directly fostered economic growth within a 20km radius. As expected, market potential in 1930 does not have any significant effect on population density in 1940, which provides further evidence of the substantial impact that Italian infrastructure had on the concentration of economic activity in the region. During the 1940s, which was a period that saw economic stagnation and low investment in infrastructure, the situation was reversed: colonial roads did not have any direct effect on population density in 1950, which is fully captured by internal market potential (column 2).

The 1950s saw major changes in the geographical distribution of the economic activity in the Horn of Africa (column 3): the direct effect of colonial roads on 1960 population density is positive and statistically significant and survives the introduction of the new set controls (column 3). This shows that colonial roads were still attracting population directly, through the comparative advantage they provided in transportation. Interestingly though, as opposed to 1940, external market potential also shows a positive and statistically significant coefficient, which suggests that the indirect effect through increased market potential in the area surrounding each cell plays an important role at this stage. Moreover, new roads built during the early 1950s, which were disproportionately allocated to historically better-connected areas (table 6) also show up as a positive and significant coefficient, which suggests that coordination of post-colonial facilities also plays a role in this phase.

Table 7 about here

By 1970, the direct effect of colonial roads disappears, newly constructed roads become insignificant and the effect measured in figures 3 and 4 is fully

explained by internal market potential in period $t-1$ (columns 4 to 7).[8] This suggests, on the one hand, that the effect of colonial roads only survived indirectly, through increasing returns to scale and, on the other, that the region had reached a resilient spatial equilibrium, which was hardly affected by the major political and institutional changes. The disappearance of the direct effect likely mirrors the loss of the monopoly that the Italian transportation network had enjoyed up to 1960. At the same time, new transportation infrastructure failed to trigger additional agglomeration effects after 1960 (all coefficients between 1970 and 2000 are not statistically different from 0), which further underscores how a stable spatial equilibrium had been reached by 1960 and that only colonial infrastructure managed to alter the regional distribution of economic concentration substantially and permanently.

7   Conclusion

In this paper, I have exploited the quasi-natural experiment provided by Italian road building in the Horn of Africa, to study the effect of colonial roads on economic development. To do so, I have created an 11x11 km grid dataset, covering Eritrea, Ethiopia and Italian Somalia containing spatially explicit information on Italian investments, contemporary proxies for economic development and several controls.

Regression analysis has shown that cells within a 10km radius from Italian paved roads are more developed today, both in terms of economic concentration and income. Falsification tests and IV treatments have confirmed that Italian roads caused these regions to become more developed. As a second step, I have employed the conceptual framework of the New Economic Geography model (NEG) to disentangle the relative importance of the direct effect of roads, through lower transport costs, and the indirect one, through increasing returns to scale and coordination of post-colonial investments.

In synthesis, the analysis of the mechanisms has pinpointed how Italian roads fostered higher concentration of economic activity directly - through lower transport costs - immediately after construction (between 1936-1940) and during the 1950s. In the 1950s, the effect was reinforced by a coordination effect, which led to the refurbishment of colonial infrastructure and to the construction of new post-colonial roads in the same areas. After 1960,

---

[8]I also check the robustness of the estimate by using an alternative measure of market potential, which only captures population of the nearest market and includes population density within a 20km radius. See data appendix and table A5 for details

the effect is fully explained by internal market potential in period $t - 1$: this suggests that a resilient spatial equilibrium was created by then and that, after that date, the effect only persisted indirectly through increasing returns to scale.

This article has focused on the relationship between transportation infrastructure and agglomeration, but it has not addressed the issues of social savings and market integration in the Italian colonies, which remain unexplored topics. Future research should focus on these aspects.

8   List of Abbreviations

- ACS = "Archivio Centrale dello Stato" (Central State Archive)

- AOI = "Africa Orientale Italiana" (Oriental Italian Africa)

- ASAIC = "Archivio Storico dell'Agenzia Italiana per la Cooperazione e lo Sviluppo" (Archive of the former "Oversea Agricultural Development")

- ASMAE = "Archivio Storico del Ministero degli Affari Esteri" (Historical Archive of the Ministry of Foreign Affairs)

- CMI = "Comando Militare Italiano"

- DHS = "Demographic and Health Survey"

- EMA = "Ethiopian Mapping Agency"

References

Acemoglu, Daron, Simon Johnson, and James Robinson (2001). "The colonial origins of comparative development: An empirical investigation". In: *American Economic Review* 91.5, pp. 1369–1401.

ajhethiopia (2000?). *Ethiopian Towns*. GIS online.

Alesina, Alberto, William Easterly, and Janina Matuszeski (2011). "Artificial states". In: *Journal of the European Economic Association* 9.2, pp. 246–277.

Andrabi, Tahir and Michael Kuehlwein (2010). "Railways and price convergence in British India". In: *The Journal of Economic History* 70.02, pp. 351–377.

Badoglio, Pietro (1937). *La guerra d'Etiopia*. Mondadori.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian (2012). "On the road: Access to transportation infrastructure and economic growth in China". In: *NBER working papers series*.

Banerjee, Abhijit and Lakshmi Iyer (2005). "History, institutions and economic performance: the legacy of colonial land tenure systems in India". In: *The American Economic Review* 95.4, pp. 1190–1213.

Baum-Snow, Nathaniel (2007). "Did highways cause suburbanization?" In: *The Quarterly Journal of Economics* 122.02, pp. 775–805.

Bekele, Shiferaw (1995). *An economic history of modern Ethiopia*. Codesria.

Benti, Getahun (2016). *Urban growth in Ethiopia, 1887-1974: From the foundation of Finfinnee to the demise of the first Imperial era*. Lexington Books.

Berger, Thor and Kerstin Enflo (2017). "Locomotives of local growth: The short-and long-term impact of railroads in Sweden". In: *Journal of Urban Economics* 98, pp. 124–138.

Bertazzini, Mattia C (2015). *The long-term impact of the Italian colonial expenditures in the Horn of Africa*. Unpublished master dissertation, University of Oxford.

Bleakley, Hoyt and Jeffrey Lin (2012). "Portage and path dependence". In: *The quarterly journal of economics* 127.2, p. 587.

Bolt, Jutta and Dirk Bezemer (2009). "Understanding long-run African growth: colonial institutions or colonial education?" In: *Journal of development studies* 45.1, pp. 24–54.

Bruhn, Miriam and Francisco A Gallego (2012). "Good, bad, and ugly colonial activities: do they matter for economic development?" In: *Review of Economics and Statistics* 94.2, pp. 433–461.

Burgess, Robin and Dave Donaldson (2010). "Can openness mitigate the effects of weather shocks? Evidence from India's famine era". In: *The American economic review* 100.2, pp. 449–453.

Burgess, Robin et al. (2015). "The value of democracy: evidence from road building in Kenya". In: *American Economic Review* 105.6, pp. 1817–1851.

Buys, Piet, Uwe Deichmann, and David Wheeler (2006). "Road Network Upgrading and Overland Trade Expansion in Sub-Saharan Africa". In: *World Bank Policy Research Working Paper Series 4097.*

Cagé, Julia and Valeria Rueda (2016). "The Long-Term Effects of the Printing Press in sub-Saharan Africa". In: *American Economic Journal: Applied Economics* 8.3, pp. 69–99.

Chaves, Isaías N, Stanley L Engerman, and James A Robinson (2014). "Reinventing the wheel: The economic benefits of wheeled transportation in early colonial British West Africa". In: *Africa's Development in Historical Perspective.* Ed. by Emmanuel Akyeampong et al. Cambridge University Press. Chap. 10.

Ciccone, Antonio and Robert E Hall (1996). "Productivity and the density of economic activity". In: *American Economic Review* 86.1, pp. 54–70.

CMI (1935). *Impero etiopico - Rete delle trasmissioni (Map 1:2,000,000?)* Comando Militare Italiano.

Cobolli-Gigli, Giuseppe (1938). *Strade Imperiali.* Mondadori.

Cogneau, Denis and Alexander Moradi (2014). "Borders that divide: Education and religion in Ghana and Togo since colonial times". In: *The Journal of Economic History* 74.03, pp. 694–729.

Corni, Guido (1937). *Somalia Italiana (2 volumes).* Mondadori.

CTI, Consociazione Turistica Italiana (1938). *Africa Orientale Italiana (Map 1:1,000,000 - 6 sheets).*

Dardano, Achille (1925). *Africa Orientale (Map 1:2,000,000).* Ministero delle Colonie - Ufficio Cartografico.

— (1935). *Africa Orientale (Map 1:2,000,000).* Ministero delle Colonie - Ufficio Cartografico.

— (1939). *Africa Orientale (Map 1:2,000,000).* Ministero delle Colonie - Ufficio Cartografico.

Davis, Donald R and David E Weinstein (2002). "Bones, bombs and break points: the geography of economic activity". In: *American Economic Review* 92.5, pp. 1269–1289.

Del Boca, Angelo (1979). *Gli italiani in Africa Orientale. La conquista dell'Impero.* Laterza.

— (1986). *Gli Italiani in Africa Orientale. La caduta dell'Impero.* Laterza.

Del Boca, Angelo (1992). *Gli italiani in Africa Orientale. Nostalgia delle colonie.* Laterza.

Diamond, Jared (1997). *Guns, germs, and steel: The fates of human societies.* Norton & Company.

Dominioni, Matteo (2008). "Il sistema di occupazione politico-militare dell'Etiopia". In: *Politiche di occupazione dell'Italia Fascista.* IRSIFAR.

Donaldson, Dave (2010). "Railroads of the Raj: Estimating the impact of transportation infrastructure". In: *NBER working paper series.*

Donaldson, Dave and Richard Hornbeck (2016). "Railroads and American economic growth: a 'market access approach'". In: *Quarterly Journal of Economics*, pp. 799–858.

Dore, Gianni (2013). *Governare l'oltremare: istituzioni, funzionari e societá nel colonialismo italiano.* Carocci.

EMA (1955). *Highway Map of Ethiopia.* Imperial Ethiopian Government.

— (1962). *Highway Map of Ethiopia.* Imperial Ethiopian Government.

— (1973). *Highway Map of Ethiopia.* Imperial Ethiopian Government.

— (1981). *National Atlas of Ethiopia (preliminary version).* Ethiopian Mapping Agency.

— (1988). *National Atlas of Ethiopia.* Ethiopian Mapping Authority.

Emmenegger, Rony (2012). "The Roads of Decentralisation: The History of Rural Road Construction in Ethiopia". In: *NCCR North-South* 39.

Faber, Benjamin (2014). "Trade integration, market mize, and industrialization: Evidence from China's National Trunk Highway System". In: *Review of Economic Studies* 81, pp. 1046–1070.

Fafchamps, Marcel, Michael Koelle, and Forhad Shilpi (2017). "Gold mining and proto-urbanization: recent evidence from Ghana". In: *Journal of Economic Geography* 17.5, pp. 975–1008.

Fenske, James and Namrata Kala (2015). "Climate and the slave trade". In: *Journal of Development Economics* 112, pp. 19–32.

Fors, Heather Congdon and Ola Olsson (2007). "Endogenous institutional change after independence". In: *European Economic Review* 51.8, pp. 1896–1921.

Fourie, Johan (2008). "A note on infrastructure quality in South Africa". In: *Development Southern Africa* 25.4, pp. 481–494.

Frankema, Ewout (2010). "Raising revenue in the British empire, 1870-1940: how 'extractive' were colonial taxes?" In: *Journal of Global History* 5.03, pp. 447–477.

Gallup, John Luke, Jeffrey D Sachs, and Andrew D Mellinger (1999). "Geography and economic development". In: *International regional science review* 22.2, pp. 179–232.

Gardner, Leigh A (2012). *Taxing colonial Africa: the political economy of British imperialism.* Oxford University Press.

Glaeser, Edward L et al. (2004). "Do institutions cause growth?" In: *Journal of economic Growth* 9.3, pp. 271–303.

Graziani, Rodolfo (1938). *Il fronte sud.* Mondadori.

Havinden, Michael and David Meredith (1993). *Colonialism and development: Britain and its tropical colonies, 1850-1960.* Psychology Press.

Henderson, Vernon, Adam Storeygard, and David Weil (2012). "Measuring Economic Growth from Outer Space". In: *The American Economic Review* 102.2, pp. 994–1028.

Henze, Paul B (2000). *Layers of time: A history of Ethiopia.* C. Hurst & Co. Publishers.

Huillery, Elise (2009). "History matters: The long-term impact of colonial public investments in French West Africa". In: *American Economic Journal: Applied Economics* 1.2, pp. 176–215.

— (2014). "The Black Man's Burden: The Cost of Colonization of French West Africa". In: *The Journal of Economic History* 74.1, pp. 1–38.

Iyer, Lakshmi (2010). "Direct versus indirect colonial rule in India: Long-term consequences". In: *The Review of Economics and Statistics* 92.4, pp. 693–713.

Jedwab, Remi, Edward Kerby, and Alexander Moradi (2017). "History, path dependence and development: Evidence from colonial railroads, settlers and cities in Kenya". In: *The Economic Journal* 127.203, pp. 1467–1494.

Jedwab, Remi and Alexander Moradi (2016). "The permanent effects of transportation revolutions in poor countries: Evidence from Africa". In: *Review of Economics and Statistics* 98.2, pp. 268–284.

Kiszewski, Anthony et al. (2004). "A global index representing the stability of malaria transmission". In: *The American journal of tropical medicine and hygiene* 70.5, pp. 486–498.

Klein Goldewijk, Kees et al. (2011). "The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years". In: *Global Ecology and Biogeography* 20.1, pp. 73–86.

Krugman, Paul (1991). "Increasing returns and economic geography". In: *Journal of Political Economy* 99.3, pp. 483–499.

— (1993). "First nature, second nature, and metropolitan location". In: *Journal of regional science* 33.2, pp. 129–144.

Labanca, Nicola (2002). *Oltremare: Storia dell'espansione coloniale italiana.* Il Mulino.

Leul, Kone, Argaw Petros, and Yulianos Kebede (2008). "The status of Road Construction in Ethiopia and its future activities". PhD thesis. aau.

Limao, Nuno and Anthony J Venables (2001). "Infrastructure, geographical disadvantage, transport cost and trade". In: *The World bank Economic Review* 15.3, pp. 451–479.

Marchitto, Nicola (1940). "Prospettive ferroviarie dell'Impero". In: *Rassegna Economica dell'Africa Italiana* 19.2.

Martinelli, Pablo (2014). "Von Thünen south of the Alps: access to markets and interwar Italian agriculture". In: *European Review of Economic History* 18.2, pp. 107–143.

Martinez-Galarraga, Julio (2012). "The determinants of industrial location in Spain, 1856–1929". In: *Explorations in Economic History* 49.2, pp. 255–275.

Michaels, Guy (2008). "The effect of trade on the demand for skill: Evidence from the interstate highway system". In: *The Review of Economics and Statistics* 90.4, pp. 683–701.

Michalopoulos, Stelios and Elias Papaioannou (2013). "Pre-Colonial Ethnic Institutions and Contemporary African Development". In: *Econometrica* 81.1, pp. 113–152.

Miguel, Edward and Gérard Roland (2011). "The long-run impact of bombing Vietnam". In: *Journal of development Economics* 96, pp. 1–15.

Murdock, George Peter (1967). *Ethnographic atlas.* Pittsburgh: University of Pittsburgh Press.

Nunn, Nathan (2008). "The long-term effects of Africa's slave trades". In: *The Quarterly Journal of Economics* 123.1, pp. 139–176.

Nunn, Nathan and Diego Puga (2012). "Ruggedness: The blessing of bad geography in Africa". In: *Review of Economics and Statistics* 94.1, pp. 20–36.

Percoco, Marco (2016). "Highways, local economic structure and urban development". In: *Journal of Economic Geography* 16.5, pp. 1035–1054.

Ploeckl, Florian (2012). "Endowments and market access; the size of towns in historical perspective: Saxony, 1550–1834". In: *Regional Science and Urban Economics* 42.4, pp. 607–618.

Podestà, Gian Luca (2004). *Il mito dell'impero: economia, politica e lavoro nelle colonie italiane dell'Africa orientale, 1898-1941.* Giappichelli.

Rappaport, Jordan and Jeffrey D Sachs (2003). "The United States as a coastal nation". In: *Journal of Economic growth* 8.1, pp. 5–46.

Rosés, Joan R (2003). "Why isn't the whole of Spain industrialized? New economic geography and early industrialization, 1797–1910". In: *The Journal of Economic History* 63.04, pp. 995–1022.

Sillani, Tommaso (1937). *L'Impero (AOI).* La Rassegna Italiana.

Stifel, David, Bart Minten, and Bethlehem Koru (2016). "Economic Benefits of Rural Feeder Roads: Evidence from Ethiopia". In: *The Journal of Development Studies* 52.09, pp. 1335–1356.

Storeygard, Adam (2016). "Farther on down the road: Transport costs, trade and urban growth in Sub-Saharan Africa". In: *Review of Economic Studies* 83, pp. 1263–1295.

URSS (1968). *Ethiopia (Map 1:2,500,000)*. Carthography department - URSS Ministry Committee.

Various (1937-1943). *Gli Annali dell'Africa Italiana*. Ed. by Angelo Piccioli. A. Mondadori.

— (1945). *Some data on Italian activity in the colonies*. Istituto Agricolo Coloniale.

— (1946). *Main features of Italy's action in Ethiopia*. Istituto Agricolo Coloniale.

Weil, David (2014). "The Impact of Malaria on African Development over the Longue Duree". In: *Africa's Development in Historical Perspective*. Ed. by Emmanuel Akyeampong et al. Cambridge University Press. Chap. 3.

Woodberry, Robert D (2012). "The missionary roots of liberal democracy". In: *American Political Science Review* 106.02, pp. 244–274.

9   Figures and Tables

Table 1: Summary Statistics

Statistics	mean	sd	min	max	N
	(1)	(2)	(3)	(4)	(5)
Population density, 2000	44.46	144.1	0	13,762	15,550
ln luminosity, 2010	-4.499	0.771	-4.605	4.103	15,550
Road 1940, 0-10km	0.0523	0.223	0	1	15,550
Road 1940, 10-20km	0.0493	0.217	0	1	15,550
Track 1935, 0-10km	0.224	0.417	0	1	15,550
Track 1935, 10-20km	0.166	0.372	0	1	15,550
Population density, 1920	9.618	24.31	0	2,794	15,550
Population density, 1930	10.50	28.49	0	3,313	15,550
Rail 1940, 0-10km	0.0114	0.106	0	1	15,550
Track 1940, 0-10km	0.0567	0.231	0	1	15,550
District capital 1940, 0-10km	0.0321	0.176	0	1	15,550
Secondary hospital 1940, 0-10km	0.0343	0.182	0	1	15,550
Main hospital 1940, 0-10km	0.00283	0.0531	0	1	15,550
School 1940, 0-10km	0.0120	0.109	0	1	15,550
Dummy mining, 1940	0.143	0.350	0	1	15,550
River, 0-10km	0.320	0.466	0	1	15,550
Coast, 0-10km	0.0121	0.109	0	1	15,550
Rainfall (Cm/year)	11.24	11.53	0	95.17	15,550
Malaria index	6.502	5.047	0	31.17	15,550
Altitude (Meters)	954.1	741.8	-119.5	4,140	15,550
Latitude	8.053	3.827	-0.367	16.67	15,550
Longitude	41.22	4.001	34.53	51.27	15,550
Ruggedness (%,000)	1,505	1,184	0	6,593	15,550
Temperature (C°)	24.42	4.565	0	33.79	15,550
Land Suitability	5.797	1.470	0	9	15,550

Summary statistics of dependent, explanatory and control variables from equation 1. See data appendix for details on individual variables.

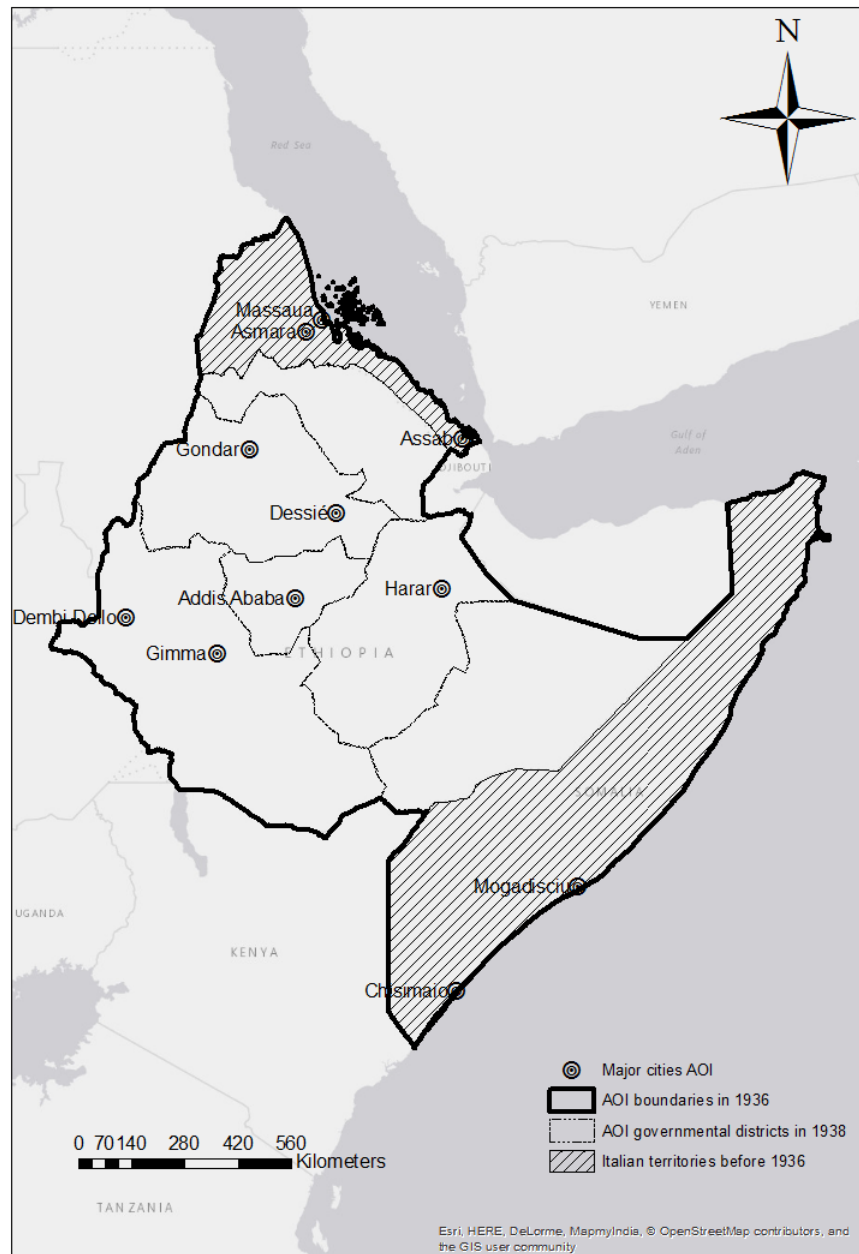Figure 1: The Italian colonies of the Horn of Africa

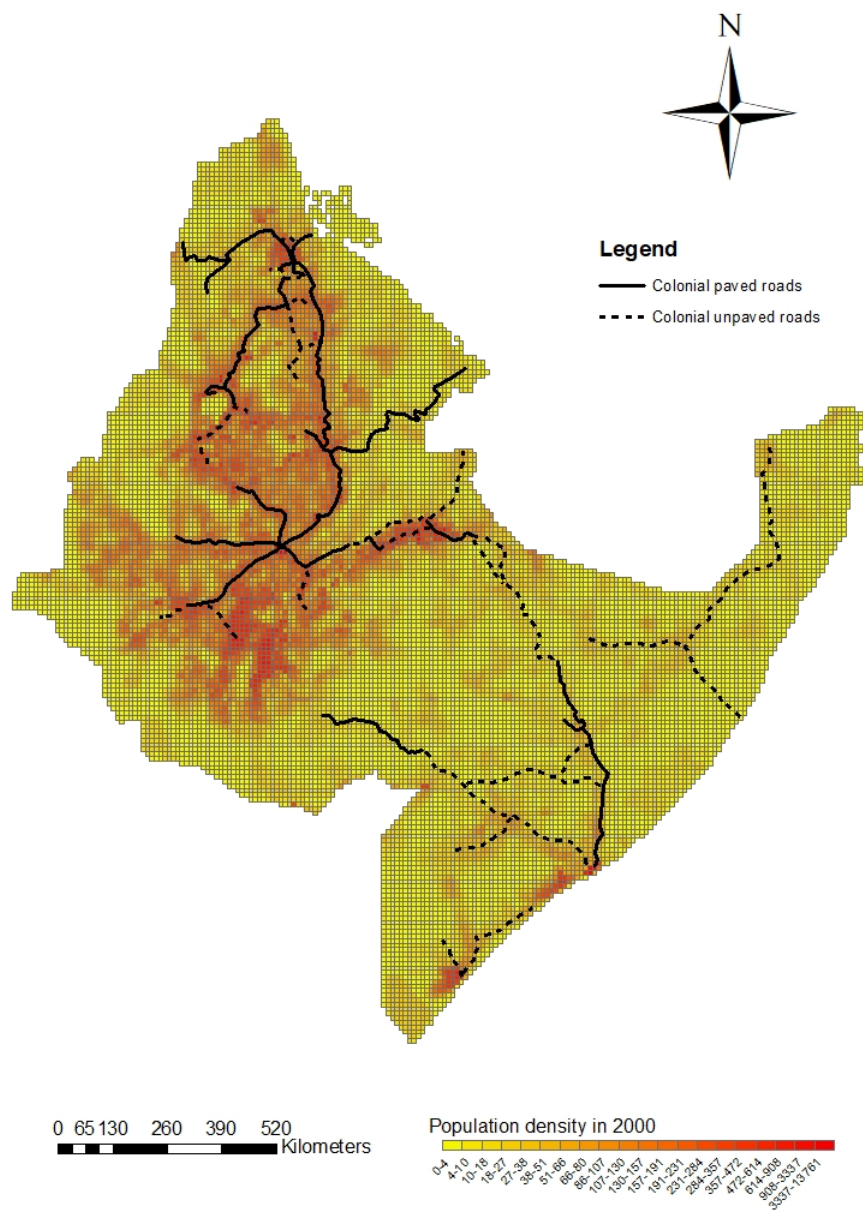Figure 2: Population density in 2000 and colonial roads

Table 2: The long-term effect of colonial roads on economic concentration

Dependent variables:	Pop density 2000, z scores				ln luminosity 2010	
	(1)	(2)	(3)	(4)	(5)	(6)
Road 1940, 0-10km	0.568***	0.468***	0.264***	0.206***	0.299***	0.275***
	(0.150)	(0.138)	(0.061)	(0.054)	(0.080)	(0.080)
	|0.153|	|0.138|	|0.057|	|0.052|	|0.084|	|0.087|
Road 1940, 10-20km	0.278***	0.182***	0.109*	0.068	0.046	0.038
	(0.062)	(0.047)	(0.058)	(0.057)	(0.045)	(0.047)
Pop density 2000 z scores						0.121***
						(0.039)
Track 1935, 0-10km		0.194***	0.139***	0.136***	0.062**	0.045*
		(0.042)	(0.033)	(0.031)	(0.027)	(0.027)
Track 1935, 10-20km		0.074***	0.065***	0.061***	-0.012	-0.019
		(0.019)	(0.018)	(0.016)	(0.015)	(0.015)
Pop density 1920, z scores		0.192***	0.156***	0.145***	-0.020*	-0.037***
		(0.014)	(0.026)	(0.023)	(0.010)	(0.010)
Pop density 1930, z scores		0.017	0.009	-0.001	0.005	0.005
		(0.018)	(0.018)	(0.010)	(0.011)	(0.010)
Rail 1940, 0-10km			0.670	0.792	0.643*	0.547
			(0.618)	(0.616)	(0.380)	(0.344)
Track 1940, 0-10km			0.154**	0.112	0.028	0.014
			(0.078)	(0.076)	(0.047)	(0.042)
District capital 1940, 0-10km			0.095	0.057	0.533***	0.526***
			(0.066)	(0.063)	(0.140)	(0.137)
Secondary hospital 1940, 0-10km			0.135**	0.116*	0.099	0.085
			(0.067)	(0.062)	(0.090)	(0.088)
Main hospital 1940, 0-10km			2.711	2.767	0.689*	0.354
			(1.954)	(1.949)	(0.407)	(0.415)
School 1940, 0-10km			0.554***	0.457**	1.204***	1.149***
			(0.207)	(0.202)	(0.232)	(0.228)
Observations	15,550	15,550	15,550	15,550	15,550	15,550
R-squared	0.16	0.20	0.24	0.27	0.16	0.18
Ethnic FE	YES	YES	YES	YES	YES	YES
Precolonial controls	NO	YES	YES	YES	YES	YES
Colonial controls	NO	NO	YES	YES	YES	YES
Geographical controls	NO	NO	NO	YES	YES	YES

OLS regression estimates in columns 1 to 6 from the grid dataset (11x11 km). Dependent variables: population density (*z-scores*) in 2000 (columns 1 to 4), logarithm of light density at night in 2010 (columns 5 and 6). Robust standard errors, clustered by 202 provincial colonial districts, in parentheses. Vertical brackets report Conley standard errors robust to spatial correlation (row 3), estimated with 1 decimal degree cut-off. All columns include fixed effects for 53 ethnic groups from Murdock's map (Nunn, 2008). Column 1 has no controls, columns 2, 3 and 4 include pre-colonial, colonial and geographical controls, respectively. Columns 5 and 6 display results with the full set of controls. Pre-colonial and colonial controls' coefficients are displayed. Geographical controls include: altitude, longitude, latitude, average annual rainfall, malaria index, average annual temperature, crop suitability for low input rain-fed cereals (FAO), ruggedness and dummies for being within 10km of a perennial river and the coast, respectively. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Colonial roads and contemporary standards of living: DHS individual data

Panel A: Wealth	wealth index			
OLS	(1)	(2)	(3)	(4)
Road 1940, 0-10km	1.586***	1.087***	0.494**	0.563***
	(0.180)	(0.222)	(0.224)	(0.213)
Road 1940, 10-20km	0.187	0.007	0.008	-0.066
	(0.187)	(0.193)	(0.179)	(0.178)
R-squared	0.32	0.34	0.40	0.42
Ethnic FE	YES	YES	YES	YES
Precolonial controls	NO	YES	YES	YES
Colonial controls	NO	NO	YES	YES
Geographical controls	NO	NO	NO	YES
Panel B: professions	professional	agricultural worker	services	skilled manual
PROBIT	(1)	(2)	(3)	(4)
Road 1940, 0-10km	0.0271***	-0.135**	0.0114***	0.0290*
	(0.0100)	(0.0658)	(0.00385)	(0.0159)
Road 1940, 10-20km	-0.00469	0.0442	0.000818	0.0127
	(0.0119)	(0.0577)	(0.00390)	(0.0168)
Observations	13,528	13,528	13,528	13,528
Precolonial controls	YES	YES	YES	YES
Colonial controls	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES

Panel A: OLS regression estimates from the 2011 male DHS dataset, in columns 1 to 4; household's wealth index as dependent variable. All columns include fixed effects for 53 ethnic groups from Murdock's map (Nunn, 2008). Panel B: Probit estimates from the 2011 individual men DHS dataset, in columns 1 to 4; binary variables with type of employment as dependent variables. Marginal effects are reported. *Professional* includes managers, teachers and engineers (column 1). *Agricultural workers* defines workers in the primary sector (column 2). *Services* includes all workers in insurance services, banking and, more generally, the tertiary sector excluding column 1 and sales (Column 3). Columns 4 reports estimates for *skilled manuals*. All columns, in both panels, display results with the full set of controls (see table 2) and report robust standard errors, clustered by 202 provincial colonial districts. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Robustness checks: colonial roads on populations density in 2000

Estimation:	(1) Baseline	(2) Placebo projected	(3) Placebo tracks	(4) First stage	(5) IV Transmission 0-20km
Road 1940, 0-10km	0.206***				0.959**
	(0.054)				(0.438)
Road 1940, 10-20km	0.068				
	(0.057)				
Projected road 1937, 0-10km		0.079			
		(0.054)			
Projected road 1937, 10-20km		0.026			
		(0.047)			
Track 1940, 0-10km			0.129		
			(0.082)		
Track 1940, 10-20km			0.107**		
			(0.050)		
Transmission lines pre-1936, 0-20km				0.077***	
				(0.018)	
Observations	15,550	15,550	15,550	15,550	15,550
R-squared	0.27	0.26	0.27	0.24	0.11
Cragg-Donald Wald F stat	x	x	x	x	244.1
Kleibergen-Paap F stat	x	x	x	x	19.06
Ethnic FE	YES	YES	YES	YES	YES
Precolonial controls	YES	YES	YES	YES	YES
Colonial controls	YES	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES	YES

All columns include fixed effects for 53 ethnic groups from Murdock's map (Nunn, 2008). Robust standard errors, clustered by 202 provincial colonial districts, are shown in parentheses. The full set of controls (see table 2) is included in all columns. Column 2 reports the placebo exercise with projected roads that were not completed due to WWII. Column 3 shows the estimates for colonial unpaved tracks. Column 4 shows the first stage regression, whereas column 5 reports the second stage. A dummy for being between 0 and 20 km from pre-1936 telegraph or telephone lines is employed as instrument (columns 4 and 5). *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 3: The effect of Italian roads over time
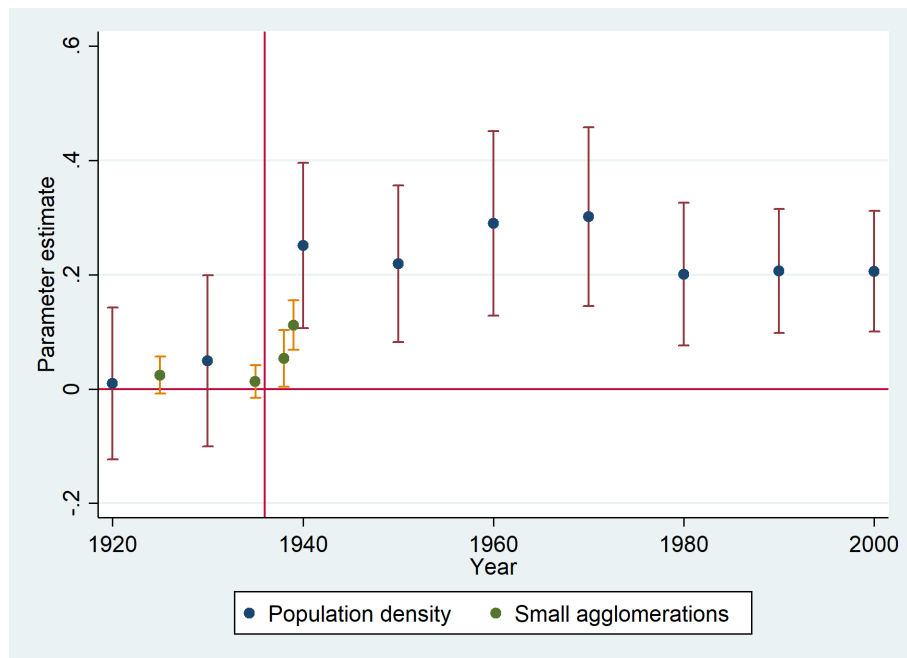
OLS elasticities, grid dataset (11x11 km). Blue dots: elasticities for dummy 0-10km from colonial roads (equation 1) with decadal population densities between 1920 and 2000 as dependent variable. Green dots: elasticities for dummy 0-10km from colonial roads with a binary variable for presence of villages in cell i in 1925, 1935, 1938 and 1939 (restricted to contemporary Ethiopia only). All columns include fixed effects for 53 ethnic groups from Murdock's map (Nunn, 2008) and the full set of controls (see table 2) is included in all specifications. 95% confidence intervals bands are reported. The vertical red line corresponds to 1936.

Table 5: Population relocation vs population growth in 1939

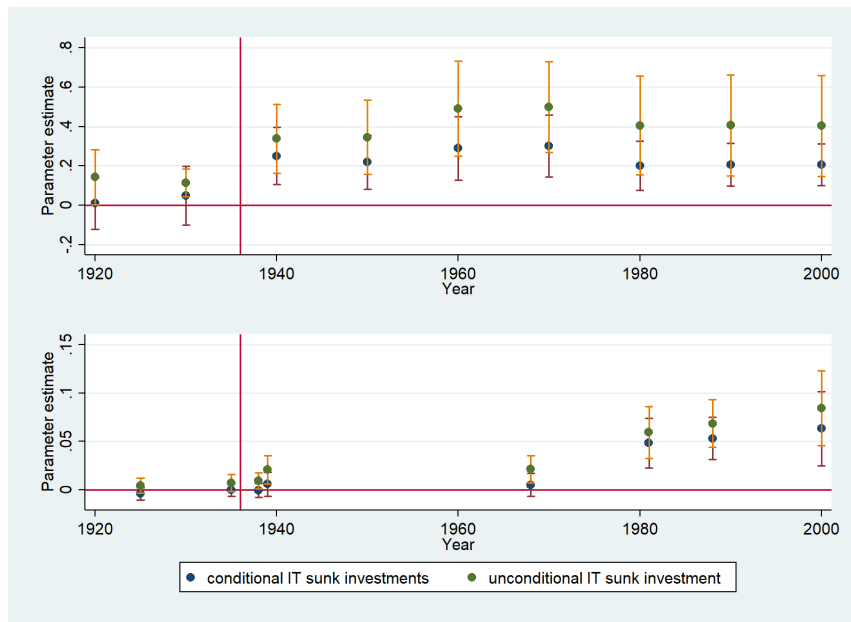| Dependent variable:	Dummy for agglomeration in 1939									
	Treatment cut-off:					Comparison group:				
	0-10km (1)	20-30km (2)	30-40km (3)	40-50km (4)	50-60km (5)	>20km (6)	>40km (7)	>80km (8)	>120km (9)	>160km (10)
Road 1940, 0-10km	0.097*** (0.019)					0.097*** (0.019)	0.099*** (0.019)	0.108*** (0.027)	0.120*** (0.031)	0.079** (0.033)
Road 1940, 10-20km	0.015 (0.017)									
Road 1940, 20-30km		-0.015 (0.016)								
Road 1940, 30-40km			-0.027* (0.014)							
Road 1940, 40-50km				0.000 (0.014)						
Road 1940, 50-60km					-0.021 (0.014)					
Observations	15,550	15,550	15,550	15,550	15,550	14,783	13,321	10,670	8,547	6,667
R-squared	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.08	0.09
Ethnic FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Precolonial controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Colonial controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

OLS estimates, grid dataset (11x11 km) and dummy for presence of village in 1939 as dependent variable for columns 1 to 10. Robust standard errors, clustered by 202 provincial colonial districts, in parentheses. All columns include fixed effects for 53 ethnic groups from Murdock's map (Nunn, 2008). In columns 1 to 5, treatment is shifted farther away from Italian roads, whilst in columns 6 to 10 the comparison group is restricted to cells located farther away from a certain cut-off distance. The full set of controls is employed in all columns (see table 2). *** p<0.01, ** p<0.05, * p<0.1

Table 6: Colonial roads and the coordination of post-colonial investments

VARIABLES	(1) Road 1955 0-10km	(2) Road 1962 0-10km	(3) Road 1973 0-10km	(4) Road 1988 0-10km	(5) School 1981 0-10km	(6) School 1988 0-10km	(7) Road 2000 0-10km	(8) School 2000 0-10km	(9) Hospital 2000 0-10km
Road 1940, 0-10km	0.097*** (0.016)	0.054*** (0.011)	0.170*** (0.019)	0.302*** (0.032)	0.056*** (0.012)	0.047*** (0.014)	0.121*** (0.043)	0.007*** (0.003)	0.027* (0.016)
Road 1940, 10-20km	0.054*** (0.009)	0.029*** (0.007)	0.071*** (0.013)	0.117*** (0.020)	0.054*** (0.011)	0.070*** (0.013)	0.033 (0.034)	0.005* (0.003)	0.024* (0.014)
Pop density 1950 z scores	0.006* (0.003)								
Pop density 1960 z scores		0.010*** (0.003)							
Pop density 1970 z scores			0.040*** (0.007)						
Pop density 1980 z scores				0.108*** (0.015)	0.031*** (0.007)				
Pop density 1990 z scores						0.053*** (0.009)			
Pop density 2000 z scores							0.202*** (0.037)	0.007*** (0.002)	0.111*** (0.010)
Observations	11,382	11,382	11,382	11,382	11,308	11,382	15,550	10,260	10,260
Precolonial Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Colonial Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Geographical Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES

Probit estimates with dummies for being located between 0 and 10 km from different types of post-colonial public facilities, grid dataset (11x11 km). Robust standard errors, clustered by 202 provincial colonial districts, are reported in parentheses. The full set of controls is included in columns 1 to 9 (see table 2). The dependent variables employed in columns 1 to 6 were only available for Ethiopia and Eritrea, which reduces the sample size to 11,382 cells. For columns 8 and 9 data were available for Ethiopia only, which further reduces the sample size to 10,260 units. See data appendix for details on the sources. *** p<0.01, ** p<0.05, * p<0.1

Figure 4: The effect of sunk investments over time

Panel A: OLS estimates (equation 1) with population density over time as dependent variable. Panel B: OLS estimates (equation 1) with large agglomerations (more than 4,000 inhabitants) as dependent variable for contemporary Ethiopia only. For both panels, blue dots report the coefficient of the dummy 0-10km from colonial paved roads conditional on the full set of controls, whereas green dots show the estimates without colonial investments (colonial sunk investments). All columns include fixed effects for 53 ethnic groups from Murdock's map (Nunn, 2008). 95% confidence bands are reported in both panels.

37

Table 7: Colonial Roads, Investments coordination and Market Potential

Dependent variable:	Population density density, z scores						
	1940 (1)	1950 (2)	1960 (3)	1970 (4)	1980 (5)	1990 (6)	2000 (7)
Road 1940, 0-10km	0.226*** (0.077)	-0.061** (0.030)	0.187** (0.087)	0.021 (0.035)	-0.038 (0.027)	0.031 (0.026)	-0.000 (0.003)
Road 1940, 10-20km	0.173*** (0.065)	-0.060*** (0.023)	0.075 (0.067)	0.009 (0.023)	-0.027 (0.018)	0.035* (0.019)	0.002 (0.001)
Road 1955, 0-10km			0.139** (0.064)				
Road 1962, 0-10km				0.039 (0.044)			
Road 1973, 0-10km					-0.050 (0.032)		
School 1981, 0-10km					-0.058** (0.029)		
Road 1988, 0-10km						-0.006 (0.013)	
School 1988, 0-10km						0.020 (0.014)	
Road 2000, 0-10km							-0.001 (0.001)
School 2000, 0-10km							0.013*** (0.004)
Hospital 2000, 0-10km							0.008*** (0.003)
Internal MP, t-1	0.001 (0.001)	0.048*** (0.003)	0.001 (0.001)	0.013*** (0.001)	0.011*** (0.001)	0.007*** (0.000)	0.005*** (0.000)
External MP, t-1	436.909 (355.326)	-3.288 (14.266)	4.036* (2.126)	-0.450 (0.681)	-2.732*** (0.988)	-1.085*** (0.227)	-0.031 (0.027)
Observations	15,550	15,550	15,550	15,550	15,550	15,550	15,550
R-squared	0.67	0.96	0.41	0.95	0.93	0.97	1.00
Ethnic FE	YES	YES	YES	YES	YES	YES	YES
Precolonial controls	YES	YES	YES	YES	YES	YES	YES
Colonial controls	YES	YES	YES	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES	YES	YES	YES

OLS estimates from equation 3, conditional on post-colonial facilities (dependent variables, table 6); grid dataset (11x11 km), with decadal population density (*z-scores*) as dependent variable in columns 1 to 7. Robust standard errors, clustered by 202 provincial colonial districts, in parentheses. All columns include fixed effects for 53 ethnic groups from Murdock's map (Nunn, 2008). Internal market potential estimates (equation 2) are calculated with HYDE and Unep population density in columns 1 to 3 and 4 to 7, respectively. External market potential is calculated from agglomeration data in columns 1 to 3, while Unep population density is employed in columns 4 to 7. See data appendix for details. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

10 Online Appendix

10.1 Appendix A: Data

10.1.1 Dependent variables

1. **Population density 1960-2000**: from the the Unep/Grid Sioux Falls database[9], raster format (0.04 x 0.04 decimal degrees). It is combined with the grid dataset by averaging all raster's cell, whose centroid falls within a grid's cells. The raster dataset is constructed from population statistics and covers the period between 1960 and 2000.

2. **Population density 1900-1950**: reconstructed population density from the History Database of Global Environment (HYDE)[10], it is based upon available historical statistical and back-projected cropland and pasture statistics, available in raster format (0.08x0.08 decimal degrees). The reconstruction is accurate for Eritrea and Somalia, whereas the Ethiopian territory shows significant spatial correlation.

3. **Villages 1925-1939**: data on location and relative size of small human agglomerations have been digitized from different sources: data on villages' locations for Ethiopia, Eritrea and Italian Somalia come from a set of Maps by the Cartographic office of the Ministry of the colonies for the years 1925, 1935 and 1939 (Dardano, 1925; Dardano, 1935; Dardano, 1939). An additional year could be added by using a touristy map of the Consociazione Turistica Italiana from 1938 (CTI, 1938).

4. **Large agglomerations 1925-2000**: data on villages larger than 4,000 inhabitants come from the same cartographic material as for point 3 (1925, 1935, 1938 and 1939). For 1968 data come from a Russian intelligence map and cover Ethiopia and Eritrea only (URSS, 1968). Cities and towns for 1981 and 1988 come the official atlases of Ethiopia (covering Eritrea as well) (EMA, 1981; EMA, 1988). Finally, Ethiopian cities in 2000 can be downloaded from GIS online (ajhethiopia, 2000?)[11].

---

[9]Available at http://na.unep.net/siouxfalls/datasets/datalist.php, last accessed on 20th of April 2016

[10]Available at http://themasites.pbl.nl/tridion/en/themasites/hyde/download/index-2.html, last accessed on 20th of April 2016

[11]Last accessed on 21st November 2017

5. **Satellite Light Density 2010**: the employed data are accessible from the National Geophysical Data Center (NOAA) in raster format (0.008x0.008 decimal degrees)[12]. I converted pixels to points and, for each cell, I calculated the average value of the contained points. This raster depicts the measured average nightly luminosity produced by humans and it is corrected for potential natural lights and atmospheric phenomena.

6. **Individual living standards 2011**: the individual dataset has been downloaded from the USAID website[13] and was collected in the context of the Ethiopian "Demographic and Health Survey" from 2011. For each individual, this source provides categorical information about profession, frequency in accessing media (television, radio and newspapers) and preferred types of payment, together with a wealth index which combines all goods and assets owned by the individual's household. The database contains 14,110 men, but due to a village having the wrong coordinates, I was unable to use the full sample, which left me a total of 13,528 observations. Each individual is linked to one of the 649 geo-coded Ethiopian villages available; their location has been modified for anonymity purposes within a 5km radius.

10.1.2 Historical data

1. **Colonial road network**: paved and unpaved colonial roads completed and under construction in 1939 were digitized from the map *Carta progressiva delle costruzioni stradali* contained in the 4th issue of 1939 (outside text). Only distance from finished roads, categorized as "completed" in this map, was considered. The map was digitized and the Euclidean distance of each cell's centroid from paved and unpaved roads was calculated.

2. **Colonial railroads**: with respect to the location of colonial railroads, for Eritrea and Ethiopia the data come from DIVA GIS[14] (functioning railroads have not changed since 1941) whereas, for Somalia, where colonial trunks were dismissed, I have geo-referenced the map *Carta dei commissariati, delle residenze, delle vie di comunicazione e delle*

---

[12]Available at http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html, last accessed on 20th April 2016

[13]Available at http://www.dhsprogram.com/Data/, last accessed on 20th of April 2016

[14]Available at http://www.diva-gis.org/, last accessed on 20th of April 2016

popolazioni (outside text) in Guido Corni's 1937 book *Somalia Italiana*.

3. **Colonial facilities in 1940**: colonial primary and secondary hospitals and colonial schools for indigenous people can be found in a series of maps (one per government per type of investment) in issue 1 of 1940 of "Annali dell'Africa Italiana".

4. **Administrative units in 1940**: colonial district capitals, along with provincial (Residenze), regional (Commissariati) and governmental (Governi) boundaries come from the map "Circoscrizioni amministrative dell'Arica Orientale Italiana", contained in the 1939's 3rd issue of *Annali dell'Africa Italiana* (outside text).

5. **Areas of mining interest in 1940**: information on whether a cell's centroid falls within a zone of mining interest come from the map *Carta mineraria dell'AOI* from issue 2 of 1940 (outside text) of *Annali dell'Africa Italiana.* All cell's centroids falling within these areas are considered suitable for mining exploitation.

6. **Post-colonial transportation network**: post-colonial data on the road network come from various maps and atlases published by the Ethiopian government, which I have geo-coded and digitised. These only cover the territories of Ethiopia and Eritrea. The road network in 1955, 1962 and 1973 has been reconstructed from a series of maps produced by the Ethiopian Imperial Highway Authority (EMA, 1955; EMA, 1962; EMA, 1973), while maps from the "National Atlases of Ethiopia" were employed to reconstruct the transportation network in 1981 and 1988 (EMA, 1981; EMA, 1988). Contemporary roads can be downloaded from DIVA-GIS and they are open source.

7. **Post-colonial hospitals and schools**: the location of schools in 1980 and 1988 is reported in the national atlases of 1980 and 1988 (EMA, 1981; EMA, 1988). Data on the location of contemporary (2000?) schools and hospitals were accessed from online sources: locations of contemporary hospitals in Ethiopia come from the map "Ethiopia health facilities", available online on the GIS online platform. For contemporary schools, instead, due to the unavailability of similar sources, I had to rely on a list of Ethiopian schools taken from an Ethiopian social network, used by former students to keep in

contact after graduation[15].

8. **Placebo lines**: roads that were projected in 1936-7, but never completed under Italian occupation due to WWII, are reported in the map "Strade in AOI" (outside text) (Sillani, 1937). For unpaved colonial roads, see point 1.

9. **Pre-1936 transmission lines**: telegraph and telephone lines were digitized from the map "Impero etiopico - Rete delle trasmissioni" (CMI, 1935).

10.1.3 Geographical and environmental controls

1. **Altitude**: altitude in meters was downloaded, in raster format (0.008x0.008 decimal degrees), from DIVA-GIS.[16]

2. **Rainfall**: annual rainfall average (centimetres/year) from the "World-climate" project, available in raster format (0.04x0.04 decimal degrees) on DIVA GIS.[17]

3. **Malaria**: malaria transmission index is available, in raster format (0.5x0.5 decimal degrees), from Kiszewski's dataset (Kiszewski et al., 2004). This informs on the rapidity and likelihood of malaria transmission in each cell.

4. **Land suitability**: agricultural suitability for rain-fed low input cereals for each cell was calculated from the GAEZ-FAO land fertility database[18], available in raster format (0.5x0.5 decimal degrees). This index ranges from 0 to 9 (0 highest, 9 lowest) and factors in temperature, soil moisture and rainfall patterns.

5. **Ruggedness**: terrain ruggedness in thousands of percentage points comes, in raster format (0.08x0.08 decimal degrees) from Nunn and Puga (2012)'s database .[19]

---

[15]Available on the website graduates.com, http://www.graduates.com/Schools/Ethiopia, last accessed on 21st of April 2016

[16]Available at http://www.diva-gis.org/, last accessed on 20th of April 2016

[17]Available at http://www.diva-gis.org/, last accessed on 20th of April 2016

[18]Available at http://www.fao.org/nr/gaez/about-data-portal/en/, last accessed on 21st of April 2016

[19]Available at http://diegopuga.org/data/rugged/, last accessed on 3rd of October 2017

6. **Temperature**: monthly (September) average temperature data between 1960 and 1990 come - in raster format (0.008x0.008 decimal degrees)- from the "Worldclimate" project.[20]

7. **Perennial rivers**: perennial waterways can be downloaded by country as a shapefile from DIVA GIS.[21]

10.1.4 Market potential

8. **Internal market potential** (Table 7): for all years between 1920 and 2000 internal market potential is calculated as $2 * Population density_{i,t-1}$. The lag is employed to deal with endogeneity.

9. **External market potential** (Table 7): for years 1940 and 1950, external market potential for each cell is calculated as the sum of relative village's size (As reported in Dardano's series of maps: 1935 for 1930 and 1939 for 1940) of all other cells in $t-1$, divided by the euclidean distance between centroids $\sum_{j \neq i} \frac{SumSizeAgglomeration_{j,t}}{Distance_j}$. This solution is preferred over simply using reconstructed population densities for years 1930 to 1940 as population statistics from HYDE suffer from spatial errors, which makes them less suitable to construct spatial lags. The same was not possible for 1950 when I had to use HYDE population density due to lack of agglomeration maps at a sufficient level of disaggregation. For the years 1950 to 2000, the same formula is applied, but with population density as nominator. The results are robust also when alternative measures are employed (see table 12 - appendix).

10.2 Appendix B: Additional estimations

10.2.1 *Additional robustness checks*

Table A2 in the appendix reports the first and second stages of an alternative IV estimation, which employs straight lines connecting regional capital, a rather mainstream strategy in the literature (Michaels, 2008; Jedwab and Moradi, 2016): this confirms the causal nature of the effect, despite being less efficient than the main IV (Table 4, columns 2 and 3). Finally, columns 4 and 5 show that both IV estimations survive the drop of targeted areas, a possible violation of the exclusion restrictions.

---

[20]Available at http://www.worldclim.org/version1, last accessed 3rd October 2017

[21]Available at http://www.diva-gis.org/, last accessed on 20th of April 2016

Tables A3 and A4 in the appendix report additional robustness checks for population density in 2000 and luminosity at night. First, endogeneity might arise due to the targeting of specific cities (and surrounding areas), which was, by definition, endogenous. Column 2 in table A3 shows that the results survive the exclusion of targeted provinces from the sample. Second, one might be concerned that the measured affect is due to selection bias in the comparison group (e.g. if the majority of non-treated locations were desert or sparsely populated due to unobservables, the coefficients would be upward biased); in columns 3 to 5, I drop those cells that were not located in the Ethiopian highlands (the most fertile and healthy areas) and those that lie further than 200 and 100km from paved roads, respectively. Third, in columns 6 to 8, I replicate the estimation including different levels of colonial administrative boundaries as fixed effects, to verify that district-level unobservables are not driving the results. The coefficients survive these various tests.

In table A4, I show that the main placebo and instrumental variable treatments employed in table 4 also work when luminosity at night is employed as dependent variable (columns 2 to 5). Finally, to confirm that the artificially selected units of observation are not driving the results, in columns 6 and 7, I estimate equation 1 with a grid dataset composed by larger (30x30 km) grid cells. If anything, this final test shows that my estimates underestimate the magnitude of the effect.

Table 8: Table A1: descriptive statistics, individual men dataset

	mean (1)	sd (2)	min (3)	max (4)	N (5)
wealth index	3.233	1.529	1	5	13,528
professional/technical/managerial	0.0553	0.229	0	1	13,528
agriculture	0.620	0.485	0	1	13,528
services	0.0185	0.135	0	1	13,528
skilled manual	0.0850	0.279	0	1	13,528
Road 1940, 0-10km	0.282	0.450	0	1	13,528
Road 1940, 10-20km	0.105	0.306	0	1	13,528
Track 1935, 0-10km	0.408	0.491	0	1	13,528
Track 1935, 10-20km	0.120	0.325	0	1	13,528
Population density, 1920	20.56	15.99	0	66	13,528
Population density, 1930	22.20	17.27	0	69	13,528
Rail 1940, 0-10km	0.151	0.358	0	1	13,528
Track 1940, 0-10km	0.108	0.311	0	1	13,528
District capital 1940, 0-10km	0.249	0.433	0	1	13,528
Secondary hospital 1940, 0-10km	0.245	0.430	0	1	13,528
Main hospital 1940, 0-10km	0.187	0.390	0	1	13,528
School 1940, 0-10km	0.227	0.419	0	1	13,528
Dummy mining 1940	0.344	0.475	0	1	13,528
River, 0-10km	0.617	0.486	0	1	13,528
Rainfall	11.30	10.31	0	72	13,528
Malaria index	3.640	5.392	0	21	13,528
Ruggedness	2,167	1,031	1	5,408	13,528
Temperature	19.89	4.147	9.900	32	13,528
Land Suitability	4.754	1.304	1	9	13,528
Latitude	9.670	2.212	3.536	14.55	13,528
Longitude	38.73	2.271	33.28	43.79	13,528
Altitude	1,735	630.7	0	3,322	13,528

Descriptive statistics for the variables of equation 1, individual DHS dataset (table 3).

Table 9: Table A2: additional IV tests for population density in 2000

Estimations:	Baseline	IV straight lines	IV straight lines	IV straight lines no targets	IV transmission lines 0-20km no targets	Baseline no FE
	(1)	(2)	(3)	(4)	(5)	(6)
Road 1940, 0-10km	0.206***		1.358**	1.390*	0.928*	0.221***
	(0.054)		(0.631)	(0.727)	(0.481)	(0.061)
Road 1940, 10-20km	0.068	-0.219***	0.306**	0.295*	0.198*	0.085
	(0.057)	(0.022)	(0.136)	(0.151)	(0.117)	(0.053)
Straight lines all, 0-10km		0.113***				
		(0.035)				
Observations	15,550	15,550	15,550	15,205	15,205	15,550
R-squared	0.27	0.24	0.08	0.04	0.08	0.22
Cragg-Donald Wald F stat	x	x	252.7	231.7	214.4	x
Kleibergen-Paap F stat	x	x	10.62	9.720	16.65	x
Ethnic FE	YES	YES	YES	YES	YES	NO
Precolonial controls	YES	YES	YES	YES	YES	YES
Colonial controls	YES	YES	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES	YES	YES

The full set of controls and fixed effects are included in columns 1 to 5 (see table 2, main text, for a description). Column 1 reports the baseline estimation (column 4, table 2). Column 6 presents the estimates from equation 1 without fixed effects. Robust standard errors, clustered by 202 colonial provincial districts, in parentheses. Column 2 reports the first stage regression of an IV treatment that employs a dummy for being within 10km from straight lines connecting main cities as an instrument (Jedwab and Moradi, 2016). Column 3 shows the second stage. In columns 4 and 5, targeted provinces (those containing a major town that became a pole of the road network) are dropped in the second stage estimations for straight lines and telegraphic & telephonic lines (table 4), respectively. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 10: Table A3: changing the comparison group

Population density, 2000	(1) Baseline	(2) No targets	(3) No low-lands	(4) < 200km	(5) <100km	(6) Province FE	(7) Regional FE	(8) State FE
Road 1940, 0-10km	0.206***	0.147***	0.256***	0.190***	0.174***	0.245***	0.186***	0.185***
	(0.054)	(0.046)	(0.058)	(0.053)	(0.052)	(0.053)	(0.050)	(0.057)
Road 1940, 10-20km	0.068	0.035	0.153***	0.062	0.056	0.089	0.068	0.053
	(0.057)	(0.063)	(0.052)	(0.058)	(0.061)	(0.073)	(0.077)	(0.070)
Track 1935, 0-10km	0.136***	0.136***	0.100***	0.150***	0.161***	0.133***	0.128***	0.132***
	(0.031)	(0.031)	(0.035)	(0.039)	(0.051)	(0.030)	(0.028)	(0.035)
Track 1935, 10-20km	0.061***	0.065***	0.050*	0.073***	0.042	0.064***	0.062***	0.063***
	(0.016)	(0.016)	(0.025)	(0.021)	(0.026)	(0.018)	(0.019)	(0.023)
Pop density 1920 z scores	0.145***	-0.043	0.055	0.141***	0.131***	0.118***	0.098***	0.151***
	(0.023)	(0.203)	(0.081)	(0.024)	(0.026)	(0.022)	(0.032)	(0.025)
Pop density 1930 z scores	-0.001	0.381	0.223***	-0.002	-0.007	-0.025***	-0.046	0.006
	(0.010)	(0.239)	(0.076)	(0.010)	(0.007)	(0.008)	(0.028)	(0.016)
Rail 1940, 0-10km	0.792	0.645	0.061	0.774	0.810	0.537	0.519	0.798
	(0.616)	(0.661)	(0.086)	(0.606)	(0.650)	(0.521)	(0.444)	(0.587)
Track 1940, 0-10km	0.112	0.090	0.230*	0.146	0.206	0.107*	0.092	0.126
	(0.076)	(0.077)	(0.124)	(0.108)	(0.137)	(0.064)	(0.060)	(0.082)
District capital 1940, 0-10km	0.057	0.058	0.061	0.085	0.040	0.115**	0.098**	0.069
	(0.063)	(0.069)	(0.068)	(0.066)	(0.074)	(0.048)	(0.049)	(0.065)
Secondary hospital 1940, 0-10km	0.116*	0.128**	0.082	0.138*	0.080	0.088	0.155***	0.168**
	(0.062)	(0.063)	(0.106)	(0.071)	(0.086)	(0.068)	(0.055)	(0.072)
Main hospital 1940, 0-10km	2.767	3.174	1.125***	3.034	3.392	3.022	2.945	2.849
	(1.949)	(2.287)	(0.328)	(2.147)	(2.334)	(2.097)	(2.051)	(1.954)
School 1940, 0-10km	0.457**	0.480*	0.221*	0.428**	0.392**	0.354*	0.332	0.366*
	(0.202)	(0.251)	(0.115)	(0.216)	(0.198)	(0.187)	(0.214)	(0.197)
Observations	15,550	15,205	5,830	11,405	6,815	15,550	15,208	15,550
R-squared	0.27	0.25	0.55	0.26	0.23	0.31	0.30	0.22
Ethnic FE	YES	YES	YES	YES	YES	NO	NO	NO
Precolonial controls	YES	YES	YES	YES	YES	YES	YES	YES
Colonial controls	YES	YES	YES	YES	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES	YES	YES	YES	YES
State FE	NO	NO	NO	NO	NO	NO	NO	YES
Regional FE	NO	NO	NO	NO	NO	NO	YES	NO
Provincial FE	NO	NO	NO	NO	NO	YES	NO	NO

The full set of controls and fixed effects are included in columns 1 to 8 (see table 2, main text, for a description). Robust standard errors, clustered by 202 colonial provincial districts, in parentheses. Column 1 reports the baseline estimation (column 4, table 2). In columns 2 to 5 the comparison group is restricted in various ways: in column 2, targeted provinces (those containing a major town that became a pole of the road network) are excluded. In column 3, the sample is restricted to those cells that had an average altitude between 800 and 2,000 meters above see level, the most suitable cells for living and farming. Columns 4 and 5 restrict the sample to cells located within 200 and 100 km from Italian paved roads, respectively. In columns 6 to 8, I employ different sets of fixed effects, using different administrative colonial units instead of ethnic fixed effects: in column 6, 202 provincial boundaries (residenze); in column 7, 73 regional boundaries (commissariati); in column 8, 6 state boundaries (Governi). *** p<0.01, ** p<0.05, * p<0.1

Table 11: Table A4: Light density robustness and different cells size

| Dataset:	10x10km grid cell					30x30km grid cell	
Specification:	Baseline OLS	Placebo 1 projected	Placebo 2 tracks	IV 1 all lines	IV 2 transmissions 20km	Baseline OLS	Baseline OLS
Dependent variable:	Light density at night						Population 2000
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Road 1940, 0-10km	0.299***			0.896*	0.725*	0.305***	0.459***
	(0.080)			(0.501)	(0.386)	(0.111)	(0.174)
Road 1940, 10-20km	0.046			0.169	0.134	0.196	0.401***
	(0.045)			(0.110)	(0.092)	(0.131)	(0.124)
Projected road 1937, 0-10km		0.069*					
		(0.040)					
Projected road 1937, 10-20km		-0.008					
		(0.034)					
Track 1940, 0-10km			0.048				
			(0.050)				
Track 1940, 10-20km			0.128***				
			(0.046)				
Observations	15,550	15,550	15,550	15,550	15,550	1,719	1,719
R-squared	0.16	0.16	0.16	0.14	0.15	0.19	0.44
Cragg-Donald Wald F stat	x	x	x	252.68	244.12	x	x
Kleibergen-Paap F stat	x	x	x	10.62	19.06	x	x
Ethnic FE	YES	YES	YES	YES	YES	YES	YES
Precolonial controls	YES	YES	YES	YES	YES	YES	YES
Colonial controls	YES	YES	YES	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES	YES	YES	YES

The full set of controls and fixed effects are included in columns 1 to 7 (see table 2, main text, for a description). Robust standard errors, clustered by 202 colonial provincial districts, in parentheses. Column 1 reports the baseline estimation (column 5, table 2). Columns 1 to 5 employ the 11x11km grid dataset, columns 6 and 7 replicate baseline estimations for Light density and population density (columns 4 and 5, table 2, main text) with a 33x33km (0.3x0.3 decimal degrees) grid dataset. In columns 2, 3, 4 and 5, I replicate the main robustness checks (table 4, main text) with the logarithm of light density at night in 2010 as dependent variable. In columns 2, roads projected but not completed are employed. In column 3, colonial unpaved roads are used. Columns 4 and 5 report IV estimates for straight lines connecting major cities and telegraphic & telephonic lines before 1936, respectively.

Table 12: Table A5: Alternative measures of market potential

Population density, a-scores	1940 (1)	1950 (2)	1960 (3)	1970 (4)	1980 (5)	1990 (6)	2000 (7)
Road 1940, 0-10km	0.267***	-0.045	0.218**	0.017	-0.064**	0.018	-0.001
	(0.077)	(0.028)	(0.092)	(0.035)	(0.032)	(0.027)	(0.003)
Road 1940, 10-20km	0.202***	-0.048**	0.105	0.007	-0.061**	0.022	0.002
	(0.062)	(0.020)	(0.069)	(0.020)	(0.025)	(0.018)	(0.001)
Road 1955, 0-10km			0.166***				
			(0.063)				
Road 1962, 0-10km				0.040			
				(0.046)			
School 1981, 0-10km					-0.118**		
					(0.050)		
Road 1973, 0-10km					-0.086*		
					(0.047)		
Road 1988, 0-10km						-0.020	
						(0.015)	
School 1988, 0-10km						-0.005	
						(0.015)	
Road 2000, 0-10km							-0.001
							(0.001)
School 2000, 0-10km							0.012**
							(0.005)
Hospital 2000, 0-10km							0.007***
							(0.003)
Internal MP, t-1	0.000	0.046***	-0.001	0.013***	0.011***	0.007***	0.005***
	(0.002)	(0.002)	(0.002)	(0.000)	(0.002)	(0.000)	(0.000)
External MP, t-1	0.027**	0.017***	0.020**	-0.003	-0.001	-0.002*	-0.000
	(0.012)	(0.004)	(0.008)	(0.003)	(0.002)	(0.001)	(0.000)
Observations	15,550	15,550	15,550	15,550	15,550	15,550	15,550
R-squared	0.68	0.97	0.42	0.95	0.92	0.97	1.00
Ethnic FE	YES	YES	YES	YES	YES	YES	YES
Precolonial controls	YES	YES	YES	YES	YES	YES	YES
Colonial controls	YES	YES	YES	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES	YES	YES	YES

Same specification of table 7. As a sensitivity check for main market potential estimates (table 7), I calculate alternative measures of internal and external market potential. Internal market potential is now estimated as the sum of the population density of each cell plus the average population density of those cells whose centroids lie within 20km from its centre: $InMP_{i,t} = Popd_{i,t} + AvPopd_{i<20km,t}$. External market potential is calculated as the internal market potential of the closest colonial district capital (with whom the cell is most likely to trade) divided by their Euclidean distance: $ExMP_{i,t} = \frac{InMP_{closestDC,t}}{Dist_{i,closestDC}}$. *** p<0.01, ** p<0.05, * p<0.1

Figure 5: Robustness checks
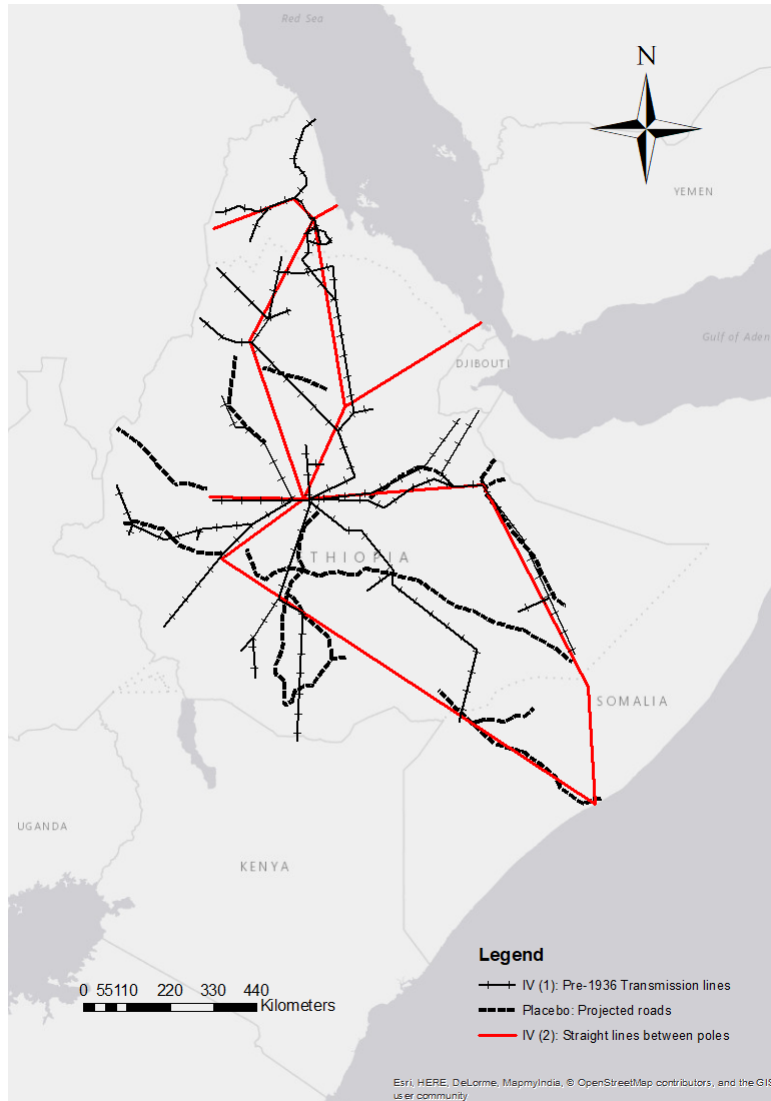
10.3 Appendix C: History of road construction in the Horn of Africa

10.3.1 Italian colonialism in the Horn of Africa

The Italian colonial occupation of the Horn of Africa started with the establishment of private commercial bases in Eritrea and Somalia. Initially, in fact, the Italian government was not willing to intervene directly and only supported private enterprises. Eritrea (first Italian private settlement in 1869) received colonial legal status in 1882, whereas Somalia (first private settlements in the early 1890s) became a colony in 1905. Between 1880 and 1920, few Italian settlers and businesses moved to the horn, due both to the limited resources that Eritrea and Somalia had to offer and the modest commitment of the liberal governments towards the economic development of these regions.

After Mussolini seized power in 1922, the colonies received more attention and resources. The renewed emphasis that the dictator placed on colonial expansion was due, on the one hand, to the fact that the colonies could provide the regime with military successes and the mirage of unlimited land for the Italian people, both extremely important elements of the internal propaganda (Labanca, 2002, pp.237-53). On the other hand, especially after the autarchic shift of the early 1930s, Mussolini was convinced that a large colonial empire could furnish raw material for manufacturing and protected export markets for Italian goods (Dore, 2013, p.62). Consequently, the government implemented a more aggressive colonial policy, that culminated in the invasion of Ethiopia in 1936. After this episode, which arguably has no precedent in colonial history both for the magnitude of the deployed army and violence against any form of organised resistance, Mussolini declared the foundation of the "empire" on May 9th 1936 (Del Boca, 1986, pp. 710-1).

The annexation of Ethiopia to Eritrea and Somalia substantially enlarged the Italian territories in the area, which were re-organised in the "Africa Orientale Italiana" (AOI). It was within the context of the Italo-Ethiopian war and immediately after the annexation that road building started to be implemented on a large scale, predominantly for military purposes.

10.3.2 Road construction before 1936

The first organic intervention on the Ethiopian road network dates back to the second half of the nineteenth century and followed the formation

and consolidation of the Amhara Empire. Emperor Tewodros (1855-68) was responsible for the creation of an embryonic centralised network which gravitated around the Amhara region in the northern part of the country. This network featured simple gravel tracks, appropriate for a more rapid deployment of infantry troops (Emmenegger, 2012, p.15). A net improvement, both in terms of quality and extension of the existing network, was achieved under the emperors Menelik II (1889-1913) and Hailé Selaissé (1920-1936), who considered a modern transportation network fundamental to achieve military and political control over the heterogeneous Ethiopian empire. The road network was then centred on Addis Ababa, which was elevated to the rank of capital of the empire by Menelik in 1889. Hailé Selaissé instituted the Ministry of Public Transport and promoted more structured road building plans, starting in the late 1920s. This project, which aimed to connect the major centres of the empire with Addis Ababa, was not completed due to the outbreak of the war as well as the lack of financial resources, materials and know-how (Emmenegger, 2012, p.16). As a result, at the outset of the Italian-Ethiopian war, few roads were completed and none had modern asphalt surface. Only two unpaved tracks for lorries existed in 1935, the Jimma-Addis Ababa and the Addis Ababa-Dessié.[22]

The situation in Eritrea and Somalia before 1935 did not greatly differ: these colonies were neither particularly densely populated nor endowed with valuable natural resources that could justify significant investments in infrastructure. In Somalia, only 4 major lorry tracks existed in 1925, and none of them was paved: these connected Mogadishu with Brava, Lugh Ferrandi, Oddur and Fer Fer.[23] In the same year in Eritrea, only the region of Asmara had an extensive network: paved roads with rudimentary asphalt surface connected Asmara with the port of Massawa and with the towns of Cheren, Adi Qualá and Adi Caieh.[24]

Between 1934 and 1935, several interventions were carried out to provide logistic support during the Ethiopian military campaign. In Eritrea, the artery connecting Massawa to Asmara was re-furbished. Military engineers rapidly built the road Nefasit-Decameré *ex nihilo*, which aimed to direct the traffic coming from Massawa directly to Adigrat without passing through Asmara (Del Boca, 1979, p.297). After the occupation of Macallé, in 1935,

---

[22]Various (1946). *Main features of Italy's action in Ethiopia.* Istituto Agricolo Coloniale, plate 1

[23]Various (1945). *Some data on Italian activity in the colonies.* Istituto Agricolo Coloniale, plate 69

[24]Various (1945). *Some data on Italian activity in the colonies.* Istituto Agricolo Coloniale, plate 66

the road Adigrat-Macallé was quickly built in order to support military penetration (Badoglio, 1937, pp.34-7). At the battle of the Ascianghi Lake (Mai Ceu, end of March 1936), the last part of the Ethiopian army was destroyed and General Badoglio, the Italian commander in chief, started planning the march on Addis Ababa: all available men were employed to refurbish the track Macallé-Quoram-Dessié-Addis Ababa (Badoglio, 1937, pp.197-8).

Similarly, in Somalia, General Graziani employed remarkable means for the construction of strategic roads aimed to support the Italian troops. In 1935, the paved road Mogadishu-Fer Fer-Mustahil was built; this satisfied the double need of connecting the outposts with the capital of the colony and of facilitating the march towards Harar and Dire Daua in the planned offensive of 1936. During the two main operations on the Southern front, the one leading to the occupation of Neghelli in January 1936 and the one aiming at the occupation of Gorrahei, Dagabur and Harar in the following April, Graziani ordered the construction of new tracks in these direction. Mogadishu was therefore connected with Neghelli and Fer Fer with Harar (Graziani, 1938, pp.284-5).

10.3.3   Italian road construction 1936-1941

With the conquest of Addis Ababa on the 5th of May 1936, the war was officially concluded. Mussolini immediately stated that the priority in the newly founded empire was to build an extensive highway network as quickly as possible. Several reasons motivated him: firstly, the peripheral regions such as those of Jimma and Gore were completely out of control, due to the impossibility of moving troops efficiently. Secondly, some of the major cities of the empire, such as Gondar, Harar, Dire Daua and even Addis Ababa, were at risk of remaining isolated during the rainy season (between June and September). Thirdly, the army and the civilian personnel in the main centres needed a constant flow of supplies from Italy. Roads were cheaper to build than railroads, faster to realise and, moreover, it was clear that for at least the next 20 years the trade volumes would have not justified the construction of railroads (Marchitto, 1940, pp.78-81). Mussolini personally laid out the infrastructural plan, that was officially communicated on the 19th of May 1936. Five main arteries had to be realised immediately. These were the Om Ager-Gondar-Debra Tabor-Dessié (650 km), the Debra Tabor-Debra Marcos-Addis Ababa (500 km), the Adigrat-Dessié-Addis Ababa (850 km), the Assab-Dessié (500 km) and the Addis Ababa-Jimma (350 km). The Addis Ababa-Allata-Neghelli-Dolo (1,100 km) had to be constructed in

a second phase.[25] Some of these projects were started but not completed due to the outbreak of WWII.[26]

The simple list of the projected roads gives an idea of the magnitude of the investments undertaken. The vast majority of the works were managed by the AASS (Azienda Autonoma Statale della Strada), a public company created by the Minister of the public works (Sillani, 1937, p.242). The AASS obtained an incredibly large budget from Rome: not only the six-year development plan allocated more than 7.7bn Italian Lire (out of the total 12bn) for road construction (Sillani, 1937, p.233), but the AASS even received additional 3.1bn Lire, for the financial year 1936-7 (Sillani, 1937, p. 243).[27] In 1939, the newly built colonial transportation network totalled roughly 4,625 km paved roads and 4,877 km unpaved tracks.[28]

The construction of this vast road network did not prevent Italians from being expelled from the Horn of Africa in November 1941. During the military events of 1940-1, the Italian troops were progressively forced to abandon all the territories of the AOI. In order to slow down the British offensive, the army destroyed several bridges, tunnels and large parts of the asphalt surface. Some of the major threats to network's integrity were the disruption of Mussolini Bridge between Agordat and Cheren, of the ramp that gave access to Cheren from the North-West (Del Boca, 1986, p.411), the dismantlement of the surface of the Mogadishu-Harar (Del Boca, 1986, p.451), of the Addis Ababa-Dessié (Del Boca, 1986, p.486) and of all the arteries that gave access to Gondar (Del Boca, 1986, p.507).

10.3.4 Post-colonial road construction (1941-2000)

Emperor Hailé Selaissé regained control over Ethiopia and Eritrea by the end of 1941. The economic stagnation that hit the area in the 1940s along with the lack of available international capitals, prevented the independent government from undertaking any maintenance of the existing network. This effect cumulated with the disruption caused by the war and yielded a substantial deterioration of the road network that shrank from about 6,000 km

---

[25]Various (1937-1943). *Gli Annali dell'Africa Italiana*. Ed. by Angelo Piccioli. A. Mondadori, year 2, issue 4, pp. 321-2

[26]This original plan was marginally modified after the approval of the six-year plan on the 21st June 1937, but the fundamental arteries remained the same

[27]The plan, which was passed in 1936, was modified into a 12-year plan in 1937. Therefore, the total public capital invested in road construction amounted to roughly 7bn Italian Lire: 1bn for the 4 financial years 1936/7-1939/40, plus the 3bn allocated in 1936.

[28]Various (1937-1943). *Gli Annali dell'Africa Italiana*. Ed. by Angelo Piccioli. A. Mondadori, year 2, issue 4, p.365

in 1940 to about 1,000 km in 1951 (Leul, Petros, and Kebede, 2008, p.9). Between 1951 and 1974, when the Communist coup, led by Menghistu, overthrew the imperial regime, four highway plans were implemented. The first one (1951-7) focused on reconstructing the Italian road network and restored about 4,200 km of paved roads, whereas the next three started building new sections, often paving colonial and pre-colonial tracks. The road network was expanded from 6,400 km, in 1951, to 9,260 km, in 1974 (Emmenegger, 2012, p.17).

As opposed to the imperial government, which had given precedence to the creation of a large highway network, the communist regime (1974-1991) placed more emphasis on rural roads. These were considered fundamental to improve market accessibility for farmers and constituted a valuable propaganda tool in the countryside. Consequently, the Highway programmes were progressively replaced by Sector Programmes (First Sector Programme 1977-1982), that extended the road network from 9,260 km to 19,020 km in 1991 (Emmenegger, 2012, p.17). During the 1990s, the last decade considered in this study, the liberal EPRDF government (Ethiopian People's Revolutionary Democratic Front) launched the First Road Sector Development Programme (1997-2007), which to some extent continued the investments in rural roads, recognizing the absence of a suitable transportation network in the countryside as a major limit to Ethiopia's development.