INSTITUTO FIGUEROLA
DE HISTORIA Y CIENCIAS SOCIALES
Universidad Carlos III de Madrid

Universidad Carlos III de Madrid

Working Papers in Economic History

April 2018 WP 18-04

Build It, and They Will Come? Secondary Railways and Population Density in French Algeria

Laura Maravall Buckwalter

Abstract
This paper estimates the effect of gaining access to railways on settler and indigenous population densities in nineteenth-century French Algeria. A growing amount of research shows that railway expansion allowed previously marginalized regions to participate in international trade and thereby to boost growth. However, few studies point out that railways increased marginalization in areas that did not gain access to the infrastructure or that did not have the required geographic characteristics needed to engage in international markets. By taking advantage of unique territorial population data and digitized historical colonization maps in the Constantine region, this paper measures the effect of gaining access in relatively isolated areas where the infrastructure arrived later using a differences-in-differences combined with a propensity score matching methodology. Results show that the indigenous population responded positively to rail infrastructure only in the regions where settler density was already high, while the settler population growth did not respond to the new infrastructure. These results are consistent with an additional IV strategy. A more detailed analysis of freight and passenger transport shows that the potential gains were restricted by tariffs, which mirrored Constantine's geographical restrictions; that is, limited fertile land and the vulnerability of agricultural production to climate.

Keywords: Colonial Railways, Transport and Trade, Agriculture, Population Density, Algeria

JEL Classification: N5, N7, N9, O18, Q17

Laura Maravall Buckwalter: Departamento de Ciencias Sociales. Universidad Carlos III, Calle Madrid, 135, 28903 Getafe, Spain).
E-mail: lmaraval@clio.uc3m.es; laura.maravall@gmail.com

Publisher:
Carlos III University of Madrid. Figuerola Institute of Social Sciences History
UNIVERSIDAD CARLOS III DE MADRID • c/ Madrid 126 • 28903 Getafe (Spain)• Tel: (34) 91 624 97 97
Site: http://portal.uc3m.es/portal/page/portal/instituto_figuerola/home/publications/working_papers

www.uc3m.es/if

Series:
Working Papers in Economic History
ISSN: 2341-2542

Electronic version of these working paper series available on:
http://hdl.handle.net/10016/16

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Build It, and They Will Come? Secondary Railways and Population Density in French Algeria

Laura Maravall Buckwalter[a]*

[a] *Universidad Carlos III de Madrid, Spain*

*lmaravall@clio.uc3m.es and laura.maravall@gmail.com

Draft: Please Do Not Quote or Cite

Abstract

This paper estimates the effect of gaining access to railways on settler and indigenous population densities in nineteenth-century French Algeria. A growing amount of research shows that railway expansion allowed previously marginalized regions to participate in international trade and thereby to boost growth. However, few studies point out that railways increased marginalization in areas that did not gain access to the infrastructure or that did not have the required geographic characteristics needed to engage in international markets. By taking advantage of unique territorial population data and digitized historical colonization maps in the Constantine region, this paper measures the effect of gaining access in relatively isolated areas where the infrastructure arrived later using a differences-in-differences combined with a propensity score matching methodology. Results show that the indigenous population responded positively to rail infrastructure only in the regions where settler density was already high, while the settler population growth did not respond to the new infrastructure. These results are consistent with an additional IV strategy. A more detailed analysis of freight and passenger transport shows that the potential gains were restricted by tariffs, which mirrored Constantine's geographical restrictions; that is, limited fertile land and the vulnerability of agricultural production to climate.

The author thanks Markus Lampe, Joan Rosés, James Simpson, Alfonso Herranz-Loncán, Giovanni Federico, Ewout Frankema, Alexandra López-Cermeño, and the participants at the Economic History Seminar in Vienna University of Economics and Business (May 2016), the Economic History Workshop at the Carlos III University (January 2017), the seminar at the University of Tuebingen (January 2018), and the Economic History Society Annual Conference (May 2018).

1. Introduction

Among the various strategies used by the colonial administration to expand French settlement in Algeria, railway infrastructure was of high relevance; it absorbed, together with the military, almost the entire French colonial budget spent on the colony. It was intended to ease colonial control and expand the settler population (Harter, 2005; Belkacemi, 1984). Nevertheless, the ratio between rural settlers and the indigenous population during the colonial years turned out differently from what the French colonial administration had hoped for. As shown in Figure 1, at the turn of the twentieth century, the rural settlers were outpaced by a persistent indigenous population growth. According to Bennoune (2002, p. 54), the indigenous population crisis experienced between the 1830s and 1870s (i.e., epidemics, droughts, famines, and a significant rebellion in the Kabyle region) had allowed many "French theorists" to predict "the doom of the native 'race.'" However, after the 1870s, the growth figures became alarming for the settler population: despite the increasingly lower margins to improve agricultural production,[1] the indigenous population kept growing.

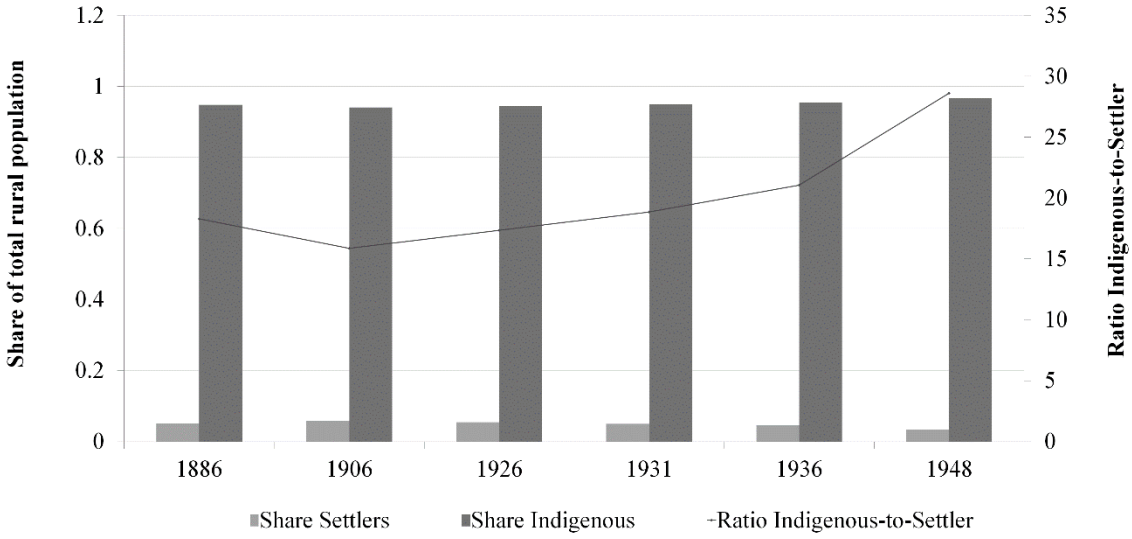
This paper assesses whether colonial railways facilitated settlement and had an impact on the settler and indigenous population densities at the end of the nineteenth century. In line with most authors finding that population growth responded positively to the arrival of the railway (Atack et al., 2010; Hornung, 2013; Jedwab and Moradi, 2016; Gregory and Henneberg, 2010; Berger and Enflo, 2017), it would seem reasonable to predict that gaining rail access had a positive effect in Algeria. As Nouschi (1961) describes, it facilitated competition in French markets (mainly if waterways were absent) thanks to higher regional

[1] Arable land was exhausted, the tribal areas were legally circumscribed, and tax burdens were extremely high.

integration and lower transportation costs, leading to benefits for both settler and indigenous communities. According to Belkacemi (1984, p. 351), settlers should have benefited as "colonization and railways were for the colonists, essential elements in the success of French policies." Also, Bennoune (2002) explains that many historians on French Algeria found that colonial expansion, channeled through railways, improved living standards and decreased mortality rates among local populations too.

[Figure 1]

Figure 1. Total rural indigenous and settler population,6 French Algeria 1872-1936



Sources: *Gouvernement Général de l'Algérie* (1948).

However, other authors oppose this view and point that the effect of railways in the Algerian case could have had the opposite effect on population densities. For instance, railways could have contributed to what Samir Amin (1970; p. 32) explained as a crowding out of small rural settlers in favor of big landowners, given that tariffs favored large producers and discouraged small cultivators (Belkacemi, 1984; Nouschi, 1961). Regarding the indigenous population, the effect could be harmful as colonial expansion (facilitated by railways) implied the redistribution of land and massive expropriations in favor of settlers. In

addition, in line with S. H. Coontz,[2] Bennoune (2002, p. 55) explains that the colonial demand for labor drove the positive indigenous demographic growth and, therefore, the Malthusian hypothesis of improved living standards had "to be rejected both on theoretical and empirical grounds." In fact, Ruedy (2005) stated that the persistent indigenous growth, together with higher impoverishment levels, was only feasible if it was accompanied by a growing demand for labor.

This paper contributes to the literature looking into the impact of railways on economic development (Coatsworth, 1979; Fogel, 1979; Fourie and Herranz-Loncan, 2004; Bogart et al., 2015; Jedwab et al., 2017; Jedwab and Moradi, 2016; Atack et al., 2010; Tang, 2014). Due to the availability of data, it focuses in the Constantine region, the largest of the three departments in Algeria.[3] However, assessing the impact of railways is complex. As Banerjee et al. (2012, p. 3) argue, the conclusions will be conditioned by the rail lines analyzed: "the first road to connect the agricultural hinterland to a port is very different from the fifth such road." Thus, this study also restricts the analysis to those regions where the infrastructure arrived relatively later; that is, during a "second wave" of railway construction in the second half of the 1880s (Berger and Enflo, 2017).[4]

To analyze the effects of the railway on settler and indigenous population densities, it uses a differences-in-differences methodology (henceforth, diff-in-diff) that allows estimating the differential effect on population between the regions that gained railway access (treated) and those that did not (control).[5] It takes advantage of geographical

[2] H. Coontz, *Population and the Economic Interpretation* (London, Routledge & Kegan Paul, 1957), p. 192.
[3] In 1848 North Algeria was divided into three *départements*. The other two were Alger and Oran.
[4] The "first wave" began with the construction of the first line built in 1862 and the following boom in the 1870s.
[5] There are different techniques that permit studying the contribution of railways. Fourie and Herranz-Loncán (2004), Bogart et al. (2015), Coatsworth (1979), and Fogel (1979) implement growth accounting and social

information systems that have permitted taking a step further in research and match spatial fundamentals to historical production and demographic data. In order to undertake this technique and claim causality, the paper first argues that obtaining access in these remote areas was exogenous to population growth. It is difficult to make such an assumption if economic development, often proxied by population growth, is found to bring in the railway infrastructure. However, the regions analyzed in this study are settlement centers and a majority of tribal areas where only a minority were starting to settle. The colonial administration did not prioritize settlement in these regions for various reasons: some experienced local resistance during the occupation, others were built to extract natural resources, some were necessary to link to Sub-Saharan Africa, and others were created to achieve a more equal and fair regional network distribution. Thus, the construction was partly motivated to expand settlement, but not as a response to a high population density.

The diff-in-diff methodology is combined with a propensity score matching technique (henceforth, PSM) to ensure the comparability between the treated and control groups as it balances them according to specific baseline characteristics (Stuart et al., 2014; Rosenbaum and Rubin, 1983; Stuart, 2010). In line with current research measuring the effects of railways, it implements an instrumental variable approach as a robustness check to the results (Atack et al., 2010; Banerjee et al., 2012; Berger and Enflo, 2017). The results show that the impact was insignificant on both population groups. Nonetheless, this methodology does not provide an explanation or any insight as to why the effect was insignificant. Therefore, this paper provides a detailed discussion on passenger and freight transportation that helps clarify why the impact was low.

savings methodologies, whereas others like Fourie and Herranz-Loncan (2004), Jedwab et al. (2017), Jedwab and Moradi (2016), Atack et al. (2010), and Tang (2014) use spatial analysis.

The structure of the paper is as follows. The next section examines the literature on the effects of railway networks and provides a brief overview of the development of railways in French Algeria and Constantine. The remaining part of the paper assesses the impact of railways and provides a descriptive analysis of freight and passenger transportation.

2. Literature Review

In recent years several publications have documented how the introduction of the railway permitted countries to specialize into higher value crops and reshape their social and production structures. Railways permitted hinterland regions, previously marginalized from international trade, to participate and benefit from increasing returns to scale, experiencing persistent effects on economic development and growth (Jedwab et al., 2017; Jedwab and Moradi, 2016). For instance, the introduction of railways enhanced real agricultural income, led farmers to increase their investment in the surrounding areas, and improved the trading environments (Donaldson, 2010; Donaldson and Hornbeck, 2012). A growing body of literature finds that colonialism shaped the long-term regional development of the occupied areas through railways and roads (Bertazzini, 2018; Jedwab et al., 2017; Jedwab and Moradi, 2016). For instance, it has been shown that gaining access to colonial railways in Sub-Saharan Africa increased productivity, contributed to local development by attracting settlers and merchants, and determined urban growth and the persistence of cities (Herranz-Loncán and Fourie, 2017; Jedwab et al., 2017; Jedwab and Moradi, 2016).

In line with this idea, the historiography of colonial Algeria has described the positive effects of railways. Yacono (1993) explains that, although the location of numerous routes lacked economic sense or any strategic meaning (such as connecting settlement centers), the overall network had a positive impact by reinforcing the economic development

of the colony. As stated by Nouschi (1961), after 1890 (when tariffs were relatively more unified and lower) the infrastructure allowed settler and indigenous products, which had been negatively affected both by the international grain price drop and the Tunisian wheat competition, [6] to compete in the French market. Furthermore, a detailed study of Algerian railways by Belkacemi (1984) points out that railways contributed to rural and urban growth. They affected the geographical distribution of settlements and facilitated the implementation of the colonial land policy (for instance, to gather information on indigenous land titles, the surveyors would locate along the railway lines).[7] According to the author, railways increased the area cultivated by Europeans and prompted the mining industry (iron ore and phosphates) by means of creating new markets, decreasing transportation costs, increasing maritime trade, mobilizing more significant volumes of goods and people, expanding irrigation, and increasing land values in the neighboring regions.[8]

However, the redistribution of economic activity brought in by the railway can also be unequal. As Coatsworth (1979) noted, the positive effects from railways can lack backward linkages within certain regions and exclusively benefit the export-led sector. In Mexico, in the nineteenth century, the new infrastructure affected the distribution of land and the "balance of social forces," facilitated land grabbing, and generated additional labor surplus (Coatsworth, 1979, p. 958). Also, some regions may lack the geographic characteristics needed to benefit from the infrastructure. For instance, Herranz-Loncán (2011; p. 1) finds that the contribution of railways to GDP in Uruguay between 1870 and

[6] After 1890s the imports of Tunisian wheat were exempt from all taxes in the port of Marseilles.
[7] This was particularly the case for the implementation of the 1873 Warnier Law as the surveyors operated along the roads and railways (Belkacemi,1984, p. 334).
[8] Belkacemi (1984, p.343) provides evidence of cases where irrigation was used as an instrument to improve fertility and ensure the success of the railway infrastructure.

1913 was low as these were unable to take advantage and benefit from export-led growth due to their 'geographic-specific character.'

In the case of Algeria, it is reasonable to expect that colonial railways highlighted social and economic inequalities. The infrastructure was intended to consolidate colonial land policies and ease the expansion of settlement. Thus, the network expansion potentially reinforced Algeria's dual economy, which was characterized by what historians often describe as an indigenous "traditional" subsistence rural sector and a settler "modern" export-led one.[9] In addition, the tariff structure, particularly in Constantine, was detrimental to small farm cultivators who were unable to benefit (Nouschi, 1961). The vast farming estates, such as the *Compagnie genevoise des Colonies suisses*, also faced prohibitive tariffs and were unable to compete with French grain producers (Lützelschwab, 2000; p. 190). The evidence suggests that although passengers and producers paid very high prices, the network experienced frequent delays and lacked specialized labor, rolling stock, and station facilities (Belkacemi, 1984). The remoteness of some areas and the line's inadequacy to cargo flows and population movements restricted the effect of railways (Yacono, 1993). Indeed, based on Auguste Burdeau's arguments in 1891 in the *Chambre d'Agriculture*, Nouschi (1961) explains that, in contrast to the railway experience in the United States which created economic activity, Algeria lacked the economic life necessary to render the network profitable.

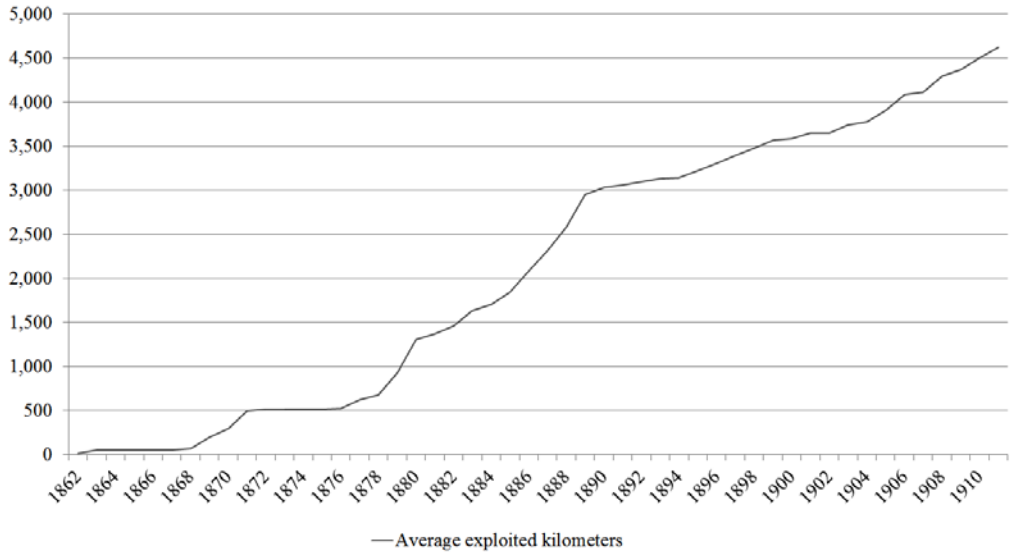
[9] For review on Algeria's dual economy see Prochaska (2004) and Griffin (1976). As shown by Good (1961), by 1954 the non-Muslim population was fully within the "developed" sector, a 10 percent was in the rural sector, and the rest in non-agricultural activities. The Muslims were mainly located in the agricultural sector (around 70 percent of the total) with about 80 percent of them engaged in the traditional sector. The source used by Good (1961) is the Tableaux de l'Economie Algérienne, 1958, p.24 from the *Service Statistique Générale*.

3. The Railroad in French Algeria[10]

The African railroad network expanded from almost 1,750 route kilometers in 1871 to nearly 6,600 in 1884. About 40 percent belonged to the British in South Africa, but almost 30 percent was built in Algeria (Harter, 2005). The latter absorbed over half of the French colonial expenses and distributed them between the military and the payment of the railroad interest stock (*guarantie d'intérêt*) (Bobrie, 1976). According to Belkacemi (1984), the total Algerian colonial investment in the infrastructure amounted to 633 million francs, without including the 367 million interest stock paid to shareholders.

[Figure 2]

Figure 2. Length (in kilometers) of Algerian network



Source: *Statistique des Chemins de Fer Français* (1896) and *Annuaire Statistique de la France* (1914-15).

[10] This section is also based on historical material provided by Bernard Venis and Francis Rambert in the site http://www.alger-roi.fr/.

It was under Napoleon III that the railway construction began in the colony, and the underlying overall structure was completed by the 1890s. Figure 2 shows the evolution of the railway mileage between 1862 and 1910. By 1935, the Algerian network had reached a total of 4,861 kilometers.[11] It was formed by a central trunk route running parallel to the coast linking Oran, Alger, and Constantine to other coastal and interior regions.[12] In Constantine, as shown in Figure 3, the central line united the ports of Bougie, Philippeville, and Bône to the southern hinterland regions Tébessa, Aïn-Beïda, and Biskra. By the end of the 1880s, Algeria's rail system linked Morocco to Tunisia and crossed important inner cities.

To finance the construction of railways and attract private capital, after the mid-1870s, the State relied on a fixed rate system (*guarantie d'intêret*) and granted concessions to companies in exchange for an annual fixed payment.[13] The *Compagnie des Chemins de Fer Algériens* (CCFA) was the first company to obtain line concessions in 1860.[14] It was a joint-stock company created to consolidate French presence and allow the transport of natural resources from the hinterland regions to the ports. In 1862, the company inaugurated eight steam locomotives that carried both passengers and freight at a speed of 20 to 25 kilometers per hour between Alger and Blida. Nonetheless, only the Algiers–Blida line was completed as financial problems soon affected the company (Belkacemi, 1984). Consequently, the government relied on additional companies and, by the late 1880s, the six

[11] *Direction des Chemins de Fer*, 1935.
[12] The central line, together with its coastal vertebrates, was established by the first of the three development plans designed to create the infrastructure: the first in 1857, the second in 1879, and the third initiated in 1907 and lasted up to1909.
[13] The payment lasted the whole concession term of 99 years and was proportional to the initial capital investement.
[14] For the Philippeville–Constantine, Algiers–Blida, and Oran–Sig routes.

railway lines built in Algeria were owned by the East Algerian Company (CEA), the Franco-Algerian Company, the Parys-Lyon, the Mediterranean Company, the Bône-Guelma Company (CBG), the West Algerian Company, and Mokta-el Hadid (Harter, 2005).

In Constantine, the ownership of lines was distributed between the CEA and the CBG. The CEA owned a significant portion of Constantine's lines. It opened to traffic the Philippeville–Constantine, and Constantine–Sétif lines in 1870 and 1879 respectively.[15] During the 1880s, the company finished the Eastern network: the Ménerville–Sétif line, completed with the inauguration of the El-Achir–Sétif section in 1882, and the El-Achir–Ménerville, finished in 1886. The Bougie–Sétif line was completed with Bougie–Beni-Mançour in 1888 and Tazmalt–Beni-Mançour in 1889.[16] The company also expanded into the south by building the track El-Guerrah–Biskra in different sections: Batna–Aïn-Touta in 1886, Aïn-Touta–El-Kantara in 1887, and El-Kantara–Biskra in 1888.[17] It then inaugurated Aïn-Beïda–Ouled-Rahmoun in 1889 and later, in the 1900s, expanded to Khenchela.[18] The rest of the Constantine network was assigned to the CBG,[19] launching before the 1880s the Bône–Guelma and Guelma–Khroubs routes.[20] The company then opened Souk-Ahrás–Duvivier in 1881,[21] spread to the Tunisian border in 1884, and completed the Souk-Ahrás–Tébessa track in 1888.[22]

---

[15] The CEA was also assigned other tracks between Alger and Constantine, such as Maison–Carrée-Alma in 1877 (in circulation 1879) and Alma–Ménerville in 1878 (in circulation 1881).
[16] Both conceded in 1884.
[17] The concession for all these lines was granted in 1880.
[18] The first line is conceded in 1885 and the second in 1900. The prolongation after the 1900s to Khenchela is not included in this study.
[19] Also named the *Société de Construction des Batignolles* to which the initial concession was granted in 1874 and who also received the concession for the line from Tunisia to the Algerian border in the late 1870s.
[20] Within the line Bone–Guelma, the track Bône–Duvivier opened to transit in 1876 while Duvivier–Guelma opened in 1877. The concession was given in 1877. Khroubs was linked to Constantine by the completion of the Constantine–Sétif track.
[21] Conceded to the company in 1877.
[22] Declared of public utility in order built it in 1885.

4. Data and Empirical Model

This section analyzes the impact of the railway on population densities using a diff-in-diff strategy between two-time intervals: 1884-1892 and 1884-1897. The purpose of this methodology is to look at the differences in the outcome variables between a treatment group, formed by areas that gained railway access, and a control group, represented by the ones that never gained access. It applies a PSM technique to reduce the potential bias caused by covariates that could also determine the probability of gaining access to the network.

Build it, and they will come?[23]

The methodology measuring the effect of the railway requires clarifying that the reasons behind the expansion of the infrastructure were exogenous to population growth. This section, mostly based on Belkacemi's (1984) thesis *French Railways in Algeria, 1805-1990*, argues that the lines built in the 1880s (see the darker lines in Figure 3) were established to increase settlement, for strategic purposes (i.e., mainly to consolidate settlement), and as a matter of regional equality. Therefore, the construction did not respond to the already existent population levels.

Railway expansion was primarily a tool used by the colonial administration to advance settlement. Belkacemi (1984, p. 322-323) quotes the following report from Burdeau's House of Representatives:

[23] The quote is commonly attributed to the movie Field of Dreams (1989). In the movie the original quote is `If you build it, he will come. ´

> Draw a map of Algerian railways and another of the density of European population […] and you will notice that the railway is a perfect colonizing river which carries new settlers and sets them down along its banks.[24]

According to the author, this was particularly the case during the 1880s in the new open- to-settlement territories, such as those established between 1881 and 1891 in the High Plains of Sétif, Batna, and the lands surrounding the Beni-Mançour–Bougie tracks (Belkacemi, 1984, p. 323). In these regions, which are those included in the analysis, the settlers came after the line and settled in the proximities of a projected line.

There was also an active military component that explained the construction of railway lines in Constantine. For instance, Belkacemi (p. 321) notes the following statement from the *Courrier d'Oran* in 1881:

> [France] should construct railways in the territories of hostile (or suspect) tribes and make the stations military strong point and fortified blockhouses from which to carry out surveillance and control of the surrounding areas.[25]

Most of the tracks analyzed in this chapter (i.e., El-Guerra–Batna, Batna–Biskra, Ouled-Rahmoun–Aïn Beïda, and Duvivier–Souk-Ahrás–Tébéssa) "were all conceived as *lignes de pénétration* of only mediocre economic value" and were designed to restrict uprisings and secure the territory (Belkacemi, 1984, p. 318). The dotted areas in Figure 3 account for the Kabyle and the southern Aurés and Oases (in the proximities of Batna and

[24] Burdeau, A. *L'Algérie en 1891. Rapport de discours à la chambre des députés*, Paris, 1892.
[25] From *Courrier d'Oran*, 19 October 1881.

Biskra), which were to be secured as they experienced most indigenous uprisings.[26] The map also illustrates that the lines connecting Bougie, Constantine, and Philippeville formed a "defensive outer circle" surrounding the rebellious regions (Belkacemi, 1984, p. 315). The lines linking Tunisia and going south towards Tébessa were built as a response to the needs of the ministry of war to secure the Tunisian border, facilitate occupation, and transport troops directly from Bône (Belkacemi, 1984).

Finally, the railway expansion in Constantine was often based on matters of fairness rather than economic ones. The unequal regional distribution of lines – as opposed to the road infrastructure, which was homogeneous throughout the departments – often became the primary cause of conflict, leading to the formation of pressure groups in the *Conseils Généraux* and the Chambers of Commerce. For instance, Belkacemi (1984) notes that traffic to the port of Bougie diminished given that most of the products from the plateau of Sétif were transported by railway to the ports of Alger and Philippeville.[27] According to the author, the lines El Guerrah–Batna, Batna–Biskra, Souk-Ahrás–Tébessa, and Beni-Mançour–Bougie lines were also built based on equity issues.

Data Description

Although a significant number of lines were built in the 1870s, the darker lines in Figure 4 show those that opened to service between 1884 and 1892. These lines allow measuring the impact of railways on the population density for the years 1892 and 1897. The units of

[26] The dotted area in Figure 3 displays the regions endowed with the highest densities of Kabyle or Berber-speaking local populations. The data was obtained from the SA (1904/05)

[27] The three of them were initially built during French occupation for military purposes. The ports Djidjelli, Collo, Herbillon and La Calle were of tertiary importance.

observation are settlement centers and tribal areas known as *douars* (i.e., municipally self-governing territorial units to which the *sénatus- consulte* had been applied)[28] that did not experience any territorial changes between 1884 and 1897. The dependent variable is the population density per hectare, accounting separately for the settler and the indigenous populations. This is in line with numerous authors who, to measure the effect of the railway infrastructure, use population density and urbanization levels as proxies for economic development and growth (Atack et al., 2010; Hornung, 2013; Jedwab and Moradi, 2016; Gregory and Henneberg, 2010; Berger and Enflo, 2017). The treated variable is a dummy variable that restricts the sample based on the distance to the nearest railway station. It takes the value of 1 if the distance between the region's centroid and the nearest railway station is lower than 20 km ($D < 20km$) and 10 km ($D < 10km$), and 0 in the opposite case.

The years analyzed coincide with the population data available in the *Tableau Général des Communes* (henceforth, TGdC). These statistics were published by the General Government of Algeria and provide information on population densities for both the civil and military territory.[29] Given that Algeria was very different to France (in particular, with regards to the presence of nomadic populations),[30] the applied census technique changed according to the surveyed population category. Namely, the populations in the civil territory and those located in the settlement centers in the military territory completed a family

[28] The *sénatus -consulte* was a law passed in 1863 that aimed to disintegrate tribal areas and divide them into territorial units known as *douars*. It delimited and registered indigenous properties and provided legal land titles in accordance to French law (Bellahsene, 2006, p. 169).

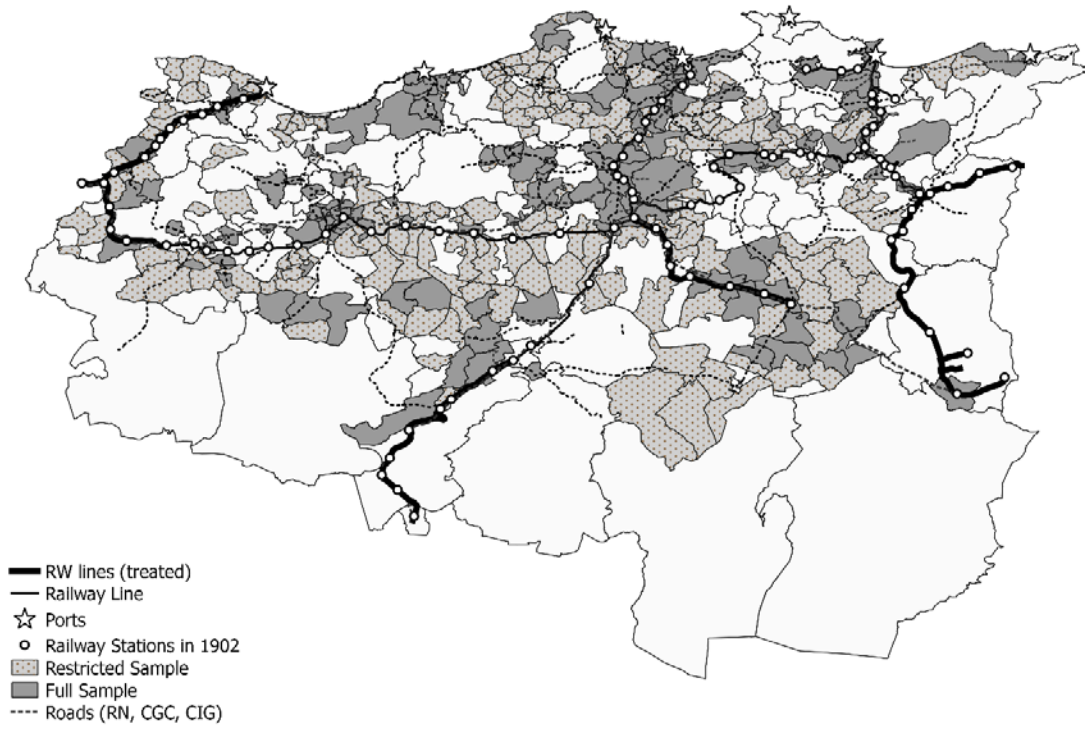
[29] The 1884 volume reflects the state of the population situation on the 30th of September while the 1892 and 1897 volumes refer to the 1st of January.

[30] Griffin (1976) explains that in Algeria, in addition to the sedentary indigenous rural owner or *fellah*, there were the semi-nomads and nomads. Among the semi-nomads, some moved continuously from one area to the next depending on the pasture, while others changed from a summer camp to a winter camp but they did not leave their tribal area. The nomads, however, did leave their tribal areas: in spring they moved from the southern regions in the Sahara to the north and returned in October.

questionnaire, while the populations in the tribal areas within the military territory were inferred by counting the number of tents (assuming that each tent hosted five to seven people). This study is limited to the civil territory, thus decreasing the potential problems resulting from differences in the data-collection procedure.

[Figure 3]

Figure 3. Selected regional sample: full and restricted, Const. between 1884 and 1897



Source: *Carte des Étapes de la province de Constantine* (1883), *Carte des voies de communication. Département de Constantine* (1902), and *Carte de la colonisation officielle, Algérie* (1902).

The data has a unique spatial detail and is subdivided into settlement centers (*centers*), plots of land (*fermes*), tribal areas or fractions, and *douars*. The territorial division

in French Algeria reflected a great regional variation regarding both administrative organization and population structure. The territory was divided into three departments that were then subdivided into municipalities or communes (i.e., *Communes de Plein Exercice* (CPE), *Communes Mixtes* (CM), and *Communes Indigènes* (CI)). This study exclusively covers the CPE and CM, as these were those under French control in the Northern part of Algeria during the years analyzed. Most settlers located in the CPE, whereas in the CM the population was mainly indigenous. The CM, which were projected to be future CPE, were mainly tribal areas and *douars*. They often included military posts, were endowed with none (or very few) settlers, and only a few of them were beginning to engage in commercial or industrial activities.

The diff-in-diff approach can be affected by sample selection bias if the selected groups differ significantly. In other words, regions that experienced railway access might differ from those that did not gain access in ways that can affect their trends in time without being uniquely explained by the impact of gaining railway. Empirically, this selection bias would require the assumption that experiencing railroad access was a random event; that is, in the absence of the treatment, both treated and control groups would have followed the same trends in time. In order to strengthen the evidence that the treatment is exogenous, the literature usually follows two strategies: 1) undertaking an instrumental variable approach and, 2) using pre-treatment data to test the parallel trends assumption (Atack and Margo, 2011; Atack et al., 2010). However, the data available do not allow implementing the pre-treatment test. Hence, based on previous studies (Atack et al., 2010; Banerjee et al., 2012; Berger and Enflo, 2017), this section first argues that gaining access was exogenous to population growth. It then demonstrates that the regions gaining railway were not significantly different in terms of population densities to those that did not gain access. Also,

it restricts the sample in various ways to improve the accuracy of comparison between both the treated and control groups, it applies a PSM technique, and it finally uses an instrumental variable approach.

In this study, the sample selection leads to a bias that helps to support the exogeneity assumption. The reasoning is as follows. First, matching the data between 1884 and 1897 requires omitting the areas that experienced boundary changes. Hence, the sample selection was biased towards relatively less settled regions given that those more attractive to settlement had a higher probability of being occupied and experiencing territorial modifications. Thus, one can infer that population pressure in these areas was not a major force explaining the construction of the railway. In fact, within all the regions matched between 1884 and 1897, only 30 percent had at least one settlement point (i.e., towns, villages, settlement centers, hamlets, and individual plots), thus suggesting that the regions that could potentially affect the probability of gaining rail access were a minority.[31]

Also, limiting the analysis to the regions that experienced a relatively later access to railways during the colonial years reinforces the exogeneity assumption. The *Dictionnaire des Communes* in 1878, together with secondary literature, provides insights as to why the railway would go to the end-points of the lines included in the analysis (that is, Batna–Biskra, Souk-Ahrás–Duvivier, Souk-Ahrás to Tunisia, Ouled-Rahmoun–Aïn Beïda, El Achir–Beni-Mançour, and Beni-Mançour–Bougie). Aïn Beïda had mineral resources (silver, lead, antimony, iron, and natural salt resources) and extensive forests. The El-Guerrah–Batna, Batna–Biskra, Ouhled-Rahmoun–Aïn-Beïda were "*lignes de pénétration* of only mediocre economic value" but were of strategical military interest. According to Belkacemi

[31] From a total of 336 areas only 101 had a settlement point.

(1984, p.318), the first two lines "penetrated the heart of the two most rebellious regions in the colony," whereas the Souk-Ahrás–Tébessa line was to "secure the border to Tunisia." Furthermore, the line concessions of Batna–Biskra, Souk-Ahrás–Tébessa, and Beni-Mançour–Bougie were built as a response to equality considerations aimed at achieving the same infrastructure distribution throughout Algeria. For instance, the construction of the Bougie and Beni-Mançour track, which was continuously delayed due to indigenous resistance in the Kabyle, was finally completed as a response to the pressure groups demanding equal terms.

To further overcome the potential bias caused by endogeneity – i.e., the railway was constructed in regions with a higher potential for population growth – the observations are classified into a "full" and "restricted" sample (see Figure 4). The full sample includes all the observations with available data for 1884, 1892, and 1897, whereas the restricted sample is limited to those regions that had few or no settlers in 1884. Thus, the restricted sample is restricted to the regions with less than 30 settlers, which were mostly used as caravan areas and remote military posts.[32] In fact, the *Dictionnaire des Communes* in 1884 shows that the regions with a colonial school or church had at least 50 settlers.[33] A 90 percent of the regions with less than 30 settlers were *douars* and tribal areas with restrictions to settlement. This reinforces the conclusion that the presence of settlers did not induce the construction of the railway.[34] In addition, based on the box plot in Figure 13 in the Appendix, an additional test is provided by limiting the full and restricted sample to the regions with an indigenous

[32] A total of 30 settlers is almost equivalent to 0.0002 per hectare in 1884 and 0.0005 in 1897.
[33] As an exception, only one hamlet with 66 settlers had a school.
[34] Moreover, given that these regions do not appear to be colonization centers in a 1902 official colonization map, it is reasonable to assume that they were not projected to be future settlements in 1884.

population density below the value of 2 to exclude extreme outside values in the control group and improve the accuracy of comparability between them.

Finally, it is necessary to test whether if, before the arrival of the railway, the regions that gained access were relatively advantaged (or disadvantaged) regarding population densities in comparison to those that were unconnected. The data shows very low correlations in 1884 between the indigenous and settler population density (in both the full and restricted samples) and the dummy variable for the treatment (equal to 1 if the region gained railway access and 0 if it did not).[35] Additionally, although Figure 10 in the Appendix shows that the average settler population of the unconnected regions in the restricted sample is higher than that of the connected areas, a t-test on their differences in 1884 is not statistically significant at a 1 percent confidence level.[36] About the indigenous population, although the density is higher in the treated sample, the differences in means are also insignificant (except for the restricted sample at a distance below 10 km).[37] Furthermore, Figure 12 displays the variations in the total settler population sample in 1884. The box plot shows that the median is approximately zero and that, despite numerous outliers, the highest value is of 2 settlers per hectare. Therefore, the correlations and mean differences suggest that the population pre-conditions do not affect the estimates.

[35] In the restricted sample, the correlation between the settler population density and the treated variable in 1884 at a distance below 20 km is -0.05 (N=206) and -0.02 (N=206) for the one below 10 km. For the full sample, the correlation is -0.02 (N=256) and 0.02 (N=256) respectively. The correlation in 1884 for the restricted sample between the indigenous population density and the treated dummy at less than 20 km is 0.08 (N=206) and 0.18 (N=206) for the distance below 10 km. In the full sample, the correlations are 0.002 (N=256) and 0.04 (N=256) respectively.

[36] This result is consistent for the distance below 20 km and 10 km.

[37] But it is not significant at a 1 percent confidence level.

Methodology and Results

The effect of gaining railway access on population density is estimated by applying a cross-section regression by OLS for the settler and indigenous populations. The equation is the following:

$$PopDens_{i,t} = \beta_0 Rail_i + \beta_1 Dyear_t + \beta_2 X_{i,t} + \beta_3 (DRail_{i,t} * Dyear_t) + s_{i,t}$$

The dependent variable is the population density for region *i* in year *t*, which can take the values 1884, 1892, and 1897. The equations are separately estimated for the settler and indigenous population. The variable $DRail_{i,t}$ is a dummy variable equal to 1 if the region *i* experienced railway access between 1884 and 1892 (or 1884 and 1897) and 0 if not. The municipalities are divided into two subsamples depending on the proximity from the region's centroid to the nearest railway station (in a straight line): below 20 kilometers and 10 kilometers. $Dyear_t$ is a dummy equal to 1 if the year is post-treatment (1892 or 1897) and 0 if pre-treatment (1884). As the number of observations is limited for both samples, particularly for the treatment group, the analysis restricts the number of control variables in vector *X* to the average elevation of the area under study. The average elevation is an exogenous variable that reflects the differences in the geographic location and distribution of both populations in French Algeria: while settlers tended to locate on the coastal plains, the indigenous populations were relatively concentrated in the hilly areas (particularly in the Kabyle region in the Western part of Constantine). Also, it could have affected the probability of gaining access to a station given that the gradient of land determined the speed of trains and increased fuel consumption and, thus, was taken into consideration when building a line (Belkacemi, 1984).

As a first insight, the tables below display the baseline differences in means (and significance) accounting for both the restricted and full sample and with and without the termini points (population nuclei that were historical cities or endowed with natural resources). Tables 1 and 2 show that the share of observations in the treated sample for the settler population range from 10 to 25 percent. The differences in means are negative and insignificant for the years 1884 and 1892, but the t-values are not far from their critical values. The significance decreases for the period 1884 and 1897, and some values become positive. The results for the indigenous population displayed in Tables 3 and 4 present values that are consistently positive, insignificant for the period 1884 and 1892, and quasi-significant for the period 1884 and 1897. These results suggest that the effect of the railway was, in general, null for the populations located in the remote areas. However, the near-significance level of some estimates – the negative effect on settler density between 1884 and 1892 and the positive effect on the indigenous population between 1884 and 1897 – suggests that the results might become significant if the estimation methodology improves group comparability.

[Table 1]

Table 1. Base diff-in-diff: settler population density, Const. 1884 and 1892

[Table 2]

Table 2. Base diff-in-diff: settler population density, Const. 1884 and 1897

[Table 3]

Table 3: Base diff-in-diff: indigenous population density, Const. 1884 and 1892

[Table 4]

Table 4: Base diff-in-diff: indigenous population density, Const. 1884 and 1897

Therefore, this section implements a PSM to ensure the comparability between the treated and control groups as it balances them according to specific baseline characteristics (Stuart et al., 2014; Rosenbaum and Rubin, 1983; Stuart, 2010). This methodology estimates the probability of belonging to the treatment group in the baseline period using multinomial logistic regression. That is, it predicts the probability of gaining access to the railway given a set of covariates. It then assigns a weight to each observation based on the estimated probability, allowing the matching of the areas with similar probabilities. Hence, by taking advantage of GIS software, the average elevation (*elevation*) is included to estimate the propensity score.[38] This methodology should provide consistent estimates of the treatment effect if the covariate if biasing the estimation.

Table 9 in the Appendix shows that within the areas with no settlers in 1884, the proximity to a railway station had a negative but insignificant effect on the settler density levels. In the full sample, the value is null. Table 11 shows that the results change if we broaden the year range up to 1897, becoming positive as we reach the areas closest to the railway station. Overall, the results suggest that gaining access in these remote regions did not increase the settler population density levels.

Concerning the indigenous population density for the years 1882 and 1892, Table 10 demonstrates that in both samples the effect of the railway increases as the region gets nearer to the station, but it is not significant. The results in Table 12 for the years 1884 and 1897 display higher values and, although all of them are not far from the t-values, they are only

[38] This methodology requires two assumptions regarding *elevation*: 1) conditional independence (i.e., once you control for *elevation*, the potential density of population is independent of gaining railway access) and, 2) the common support (for each value of *elevation* the probability of being treated (or not) is positive).

significant in the full sample when the distance is below 10 km and if the termini points are included. These results suggest that the railway had a positive and significant effect on the indigenous population only in the regions that already had settlers and were closer to the railway station.

Furthermore, the areas endowed with extreme outside values of indigenous population density are dropped based on the box plots in the Appendix - Figures 12 and 13 (the limit is set above the value of 2). Doing this increases the similarity between the control and treatment groups. The main difference from the previous results is that the settler population, as displayed in Table 15, becomes significant in the full sample between the years 1884 and 1897. The indigenous population, as shown in Tables 14 and 16, is not affected and, thus, the prior results were likely determined by the extreme values.

In line with other authors, as a robustness check to the results, an instrumental variable approach is carried out (Atack et al., 2010; Banerjee et al., 2012; Berger and Enflo, 2017). Although the lines in the regions analyzed were built based on reasons exogenous to population levels, there still might be variables explaining the population density levels that could have also affected the probability of gaining access to a line. Therefore, the approach requires identifying an exogenous variable that predicts gaining access but that does not correlate to population density. By taking advantage from the 1878 *Dictionnaire des Communes, Villes & Villages de l'Algérie* – i.e., a report containing information on the administrative condition, the geographic situation, and facilities such as the number of schools, prisons, courts, gendarmeries, banks, etc.[39] – it is possible to create a dummy

[39] Occasionally, the information was complemented with additional material. For instance, in the case of Tébessa it was useful to look at the posterior 1903 DCVVA for natural resources or military strategy. Sometimes names of locations appeared duplicated; for example, El-Ghedir is reported as a hamlet on the route between Bône and Souk-Ahrás and as a *douar* in the CM El-Arrouch.

variable that is equal to 1 if the region has any natural resources and if it is a strategic location, and equal to 0 in the opposite case. The natural resources include silver mines, antimony, iron, salt, cedar, oak forests, lead, and 16 thousand palm trees. The variable is also equal to 1 if the region is of strategical relevance for colonization: that is, if it was a of military relevance (such as Batna and Sétif), if the region was officially projected to be settled, if it represented a geographically strategic position (for instance, Biskra was an on the way to the South and in the proximities of the *Caïdat* and an oasis), and if it had relevant Arab markets for commercial exchange.[40] As Figure 15 shows, straight lines were drawn linking each point to its nearest neighbor and the closest and most important port (i.e., Bône, Philippeville, and Bougie),[41] trying to avoid mountainous regions (thus, capturing the least costly and most feasible route).

As shown in the pair-wise correlation matrix in the Appendix, the correlation between the instrument and the population density level is null, while it is high (above 50 percent) and significant (at a 1 percent confidence level) with respect to the treatment variable (i.e., $D<20km$ and $D<10\ km$). The results using the instrumental variable displayed in Tables 17 and 18 provide similar insights: gaining railway access in remote areas was ineffective with regards to settlement expansion, particularly in the areas with no initial

[40] The following regions are also included: Sigus, as it was on the route between Constantine and Aïn Beïda and was being populated prior to the construction of the railway, Sidi Mesrich, as it was on the route between Bône and Constantine and being populated, Ports-de-Fer in Bibans, as it was a strategic passage key to the Alger-Constantine route between mountains, Tébessa, as it was a gateway to the south and positioned on the frontier with Tunisia, Souk Ahrás, as it appears on the national projected route from Constantine to Tunisia and is endowed with lead, copper, and zinc mines, and Bordj Bou Arreridj, as it has a military post but also a sandstone quarry and on the way between Sétif and Alger. Three Arab markets are included as they should capture important traditional exchange areas. Military posts in remote areas that are not on the way to any strategic location are excluded (such as Takitount or M'Sila). Tizi N'Bechar is excluded because, despite being populated, is too remote and does not go to any strategic region.

[41] The main port in Constantine was Bône, followed by Philippeville and Bougie. These were built during French occupation for military purposes and after the 1850s trade began to expand, particularly because of tariff changes. I do not include the tertiary ports Djidjelli, Collo, Herbillon and La Calle.

settlers in 1884. In contrast, the results show a positive and significant effect on the indigenous population density. The full sample, which includes the settlement centers (thus, capturing the regions with colonial agriculture) suggests that indigenous population responded to the infrastructure in the regions where demand for labor was potentially higher and where the construction of railway lines allowed the demand for indigenous labor to meet the supply more "easily." However, this conclusion is hypothetical due to the lack of sub-municipal data on agricultural production. Nonetheless, the correlation between the municipal data at a municipal level from the *Statistique Agricole* and the sub-municipal one used in this paper supports the "labor demand" hypothesis (Bennoune, 2002; Coontz, 1957; Ruedy, 2005).[42]

Finally, it is unclear whether the population changes brought in by the railway reflect increases in natural growth rates as a response to the creation of newer economic activity, or if they are only accounting for the displacement of population. If the redistribution of the population partly explains the effect due to changes in the economic activity across regions, then the coefficients obtained should be overestimating the impact of the railway on population. Hence, following a similar methodology as the one used by Berger (2016) and Redding and Turner (2014), a baseline regression was selected (i.e., the propensity score diff- in-diff model where the treated group are the regions below 10 km from the nearest railway station) and the nearby areas were sequentially dropped to see if the coefficients change significantly. If part of the explanation is a redistribution of population, then the coefficients would decrease. This is because dropping from the control group the regions

[42] The correlation, conditional to the regions that gained access to the railway, between indigenous population density (at a settlement center or *douar* level) and the indigenous wages per day (proxying for a high supply of labor) is negative and significant. At a distance lower than 10 km, the correlation is -0.29 with indigenous wages, at a distance below 20 km, the correlation is -0.23.

experiencing the outward movement of the population decreases the differential effect of the railway.[43] Tables 19 and 20 display the results. The size of the coefficients for the settler population shows a substantial decrease only when the regions below 20 km are dropped in the restricted sample.[44] The indigenous population density, on the other hand, decreases when the regions below 10 km are excluded, but the coefficients do not change significantly. These results support that the redistribution of the population into the nearby regions also explains the observed changes.

5. Discussion

The results from the diff-in-diff model demonstrate that gaining access between 1884 and 1892 had no significant effect on population densities in the regions that gained access during a "second wave" of railway expansion in the 1880s in Constantine. This is in line with Nouschi (1961), who states that the colony lacked sufficient economic life to render the network profitable.[45] Similarly, Yacono (1993) argues that the railways' impact was unfelt in specific areas due to their geographic remoteness and the inadequacy of the lines to trade flows and population movements. In particular, the effect of lines like El-Guerrah–Batna, Batna–Biskra, Souk-Ahrás–Tébessa, and Beni-Mançour–Bougie– was virtually null according to Belkacemi (1984).

[43] If there is redistribution, the difference between the treatment group (with higher levels as it receives the population inflows) and the control group (with lower levels as it experiences the outflows) is higher. But if we progressively drop from the control group the nearby regions that potentially experienced the outflows, then the difference should decrease, while if there is no redistribution, the coefficient is solely accounting for natural growth.

[44] This result must be regarded with extreme caution as the number of settler in the restricted sample is extremely low and thus are very vulnerable to changes.

[45] The author quotes Auguste Burdeau in the *Chambre d'Agriculture* in 1891.

However, the diff-in-diff methodology fails to provide a more detailed explanation concerning passengers and freights that can help explain the low impact in the analyzed regions. Numerous historians on colonial Algeria blame the tariff structure. Belkacemi (1984, p. 282) noted that there was a consensus, reflected in the 1877 and 1884 parliamentary reports, in the Chambers of Commerce, and the *Conseils Généraux*, that the tariff reductions in Algeria did "not stimulate commercial transactions."[46] Nouschi (1961) describes that by the 1880s the railway was extremely expensive and prices doubled those in France. The tariff by rail between Alger (port) and Constantine (hinterland) was 27 francs, while the one paid by sea through Philippeville (port) was 14.80 francs (out of which 60 percent of the cost was getting from Constantine to Philippeville by rail). As shown in Figures 5 and 6, the passenger and freight rates were significantly higher in comparison to those paid on average in the *Métropole*.[47] The rates were mainly detrimental to small-scale producers who were unable to take advantage of the infrastructure. Nouschi (1961, p. 603) notes that a cultivator paid 8.60 francs per ton at a low speed and 19.50 francs per ton at a fast speed, while large producers benefited from fixed price tariffs for volumes from eight to ten tons, paying 3.70 francs per ton and 3.37, respectively. Railways also resulted more expensive to those farmers who cultivated cheaper products like barley or wheat given that tariffs lacked product differentiation.[48] However, there is evidence suggesting that tariffs were also high for large producers. For example, Lützelschwab (2000, p. 190) describes that the *Compagnie genevoise des Colonies suisses*, a farming enterprise dedicated to cereal in

[46] Belkacemi (1984) relies mainly on the *Conseil Général d'Alger*, Oct. 1899; *Chambre des députés*, 1893; *Chambre de commerce de Bone*, 1883-84; and *Chambre de commerce d'Alger*, 1876-77.
[47] Only passenger rates in the line Mokta-el-Hadid were low, as it was solely devoted to the extraction of natural resources
[48] Small-scale farmers were also discouraged by the enhanced land values as a result of European speculators who demanded more land concessions around the projected areas for construction (Belkacemi, 1984)

the High Plains of Sétif, struggled to compete with foreign grains after the 1870s.[49] In addition, Figure 16 in the Appendix shows the prices for passengers by class between Philippeville and Constantine. If these prices are compared to the average wages from the agricultural statistics it is possible to see that the regional differences in wages were not enough to offset the cost of transport. For instance, the distance between Bizot and Constantine was of 12.9 km (aprox. 30 minutes). The total price per ticket in the third class was of 0.8 francs (0.062 fr/km). The difference in rural wages between the two regions was 0.087 francs for the indigenous population and 1.20 for the settler population.[50]

[Figure 5]

Figure 5: Average passenger rate per kilometer, French Algeria 1896-1905

[Figure 6]

Figure 6**:** Average freight rate per kilometer, French Algeria 1896-1905

Nevertheless, as Herranz-Loncan (2011) argues,[51] although tariffs were inadequate for agricultural growth, it might be that they are reflecting other forces such as the small size of the railway sector. For instance, a low quantity of freight can hinder scale economies and push companies to increase rates. Very broad estimates indicate that the size of Algeria's railway sector was unusually small. Although it is complicated to estimate the total contribution of railways to French Algeria's GDP due to the lack of data, Amin's (p. 101; 1996) GDP estimates for some benchmark years suggest that the total contribution of freight

[49] In addition, the company was negatively affected by the delay of the arrival of the railway (after the 1870s). The lack of communication infrastructure forced the company to rely on intermediaries/traders to trade its products

[50] These statistics are obtained from the *Statistique Agricole* collected by the *Gouvernement Général de l'Algérie*, 1913/14.

[51] The author analyzes the case of railways in Uruguay between 1870 and 1913.

transport was low (changing from approximately a 0.59 percent in the 1880s to 1.61 percent in 1910).[52] The ratio between passenger and freight revenue collected from the railway's statistics and Amin's estimates shows that the contribution in 1880 and 1905 was below 0.02 percent.[53] The quantity of freight transported in relation to the population also reflects the small size of the sector. Figure 7 shows the amount of freight per capita between 1872 and 1924. The values reported were comparatively lower than other regions during the period. For example, in 1911 Algeria's 77.5 ton-km per capita was significantly smaller than those of Uruguay and Argentina in 1913 (263 ton-km and 1,201 ton-km respectively) and, in 1905, South Africa's 112 ton-km per capita was well above Algeria's 60 ton-km value.[54] Furthermore, Table 5 demonstrates that, after the 1900s, mileage per population and surface started to stagnate and lag behind regions such as Mexico, Brasil, Argentina, and Chile.

[Figure 7]

Figure 7: Passenger and freight transport, French Algeria 1863-1894

[Table 5]

Table 5: Railway density in French Algeria, 1881-1921

[52] These figures neglect passenger transport as it is included in the overall services category, with has a contribution to GDP below 20 percent for all years.

[53] Railway freight and passenger revenue for 1880 is obtained from the 1894 *Statistique des chemins defer francais, algériens et tunisiens* and the 1905 value is collected from the *Annuaire Statistique de l'Algérie* (1905). The GDP estimate for 1905 is approximated with Amin's (1966) nearest estimate which is that for 1910. Both GDP values are corrected for the CPI (base 1955=100) from Mitchell (1988).

[54] The value for South Africa was obtained from Herranz-Loncán (2011) and divided by the population value from Mitchell (1994). The population estimates prior to 1906 must be regarded with caution as they are inferred based on census observations and secondary literature. The official statistics point out that the data on Muslim population -in particular, prior to 1900- is far from reality. To build the population series, this paper relied on various *Annuaire Statistique de la France*, the 1948 Algerian census, and the work of Biraben (1969), Breil (1954), Fargues (1986), Good (1961), and Negadi et al. (1974).

The question is, then, why was the contribution of Algerian railways to the economy so low. One of the main reasons is that Algeria's share of fertile land limited the potential gains. French Algeria, which lasted from 1830 to 1962, was limited to the northern fertile regions; as Ruedy (2005, p. 5) explains, "the heart of historical Algeria is a band of valleys, mountains, and plains extending roughly three hundred kilometers inland from the Mediterranean." As an illustration, based on the World Bank's figures from 2015, the portion of arable land is of 3.1 percent (0.2 hectares per person), as opposed to South Africa's 10.3 percent (0.2 ha/person), Uruguay's 13.8 percent (0.7 ha/person), and Argentina's 14.3 percent (0.9 ha/person). Moreover, agricultural production was highly dependent upon climate. Cereal cultivation was particularly vulnerable before the introduction of new dry farming techniques in the 1900s. The years between 1890 and 1900 were critical due to bad harvests, the fall in international prices in the 1880s, and the competition with Tunisian wheat (which, after the 1890s, was free of duties in Marseille) (Nouschi, 1961). Thus, the potential gains from railways were limited as the lines analyzed in Constantine carried a significant share of cereal over the total freight: a 44 percent for Alger–Maison-Carrée, 78 percent for Ouled-Rahmoun–Aïn Beïda, 36 percent in Bougie–Beni-Mançour, and 13 percent between Batna and Biskra.[55]

However, Belkacemi (1984) explains that, even in times when agricultural production was high, the effects of the railway were restricted. The high seasonality of crops – notably, in Constantine where cereal cultivation was significant – often provoked shortage crises during high production seasons due to the lack of storage space in station facilities, the scarcity of specialized labor, and the insufficiency of rolling stock. The seasonality of crops

[55] *Statistique des Chemins de Fer Français* (1896).

complicated the ability of railway companies to calculate the optimal crop volume to be transported, often leading to an excess of transport capacity that increased expenses and pushed companies to lower operating costs.

Ultimately, the line concession and financing system did not provide the companies with the right incentives to decrease rates. On the one hand, there was a lack of route competition between the railway companies characterized by their geographical isolation and sharing of long-distance lines. The only competition encountered was based on redirecting the commercial flows to the ports by means of special tariffs. In Constantine, the competition between the CEA, the PLM, and the CBG with regards to certain commodity flows generated "complaints from disadvantaged regions that the benefits of reduced tariffs ought to be generalized throughout the colony" (Belkacemi, 1984, p. 285). Furthermore, the introduction of differential tariffs in the late 1880s "reduced the value of proximity" and benefited intra-termini, long-distance routes, at the expense of the regions relatively near the ports. They provided some producers with an "unfair advantage," whereas others, especially in the marginal regions, were negatively affected (Belkacemi, 1984, p. 283):

> They [the railway companies] were […] reluctant to extend the benefits of differential tariffs to undynamic economic regions where traffic was limited and operating costs high. The ability of the railway companies to engage in seemingly arbitrary tariff policies was constantly blamed on their monopoly position and the advantages offered by the guarantie d'intêret system.

On the other hand, with the *guarantie d'intêret*, the company's gains did not vary with its actual yearly expenses, so that the incentives were directed to gain profits by decreasing construction and exploitation costs. According to Belkacemi (1984), this system not only required an excessive share of the colonial budget (increasing from 10 percent

between 1872 and 1878 up to more than 26 percent after 1888), but it also forced concessionaries to minimize costs, benefiting shareholders and neglecting the economic development of the colony. Figure 8 shows that the CEA was the most expensive railroad network among the Algerian companies, which is consistent with the low impact of the railway on the growth of both population groups in the analyzed regions. Tariffs also highlighted regional inequalities in tariffs. For instance, Belkacemi (1984) describes that in 1886 the cost of transporting cereal in Constantine was 8 cents per ton higher as compared to that of Oran and Alger.

[Figure 8]

Figure 8. Francs per kilometer claimed annually by the railway companies to the state as *guarantie d'intêret* and insufficiency of exploitation, French Algeria 1896-1905

6. Conclusion

Extensive research demonstrates that gaining access to railways had a positive effect on economic development by allowing previously unconnected areas to gain competitive advantage and integrating with the market economy. However, few studies examine the regions that were unable to do so. This paper studies the effect of the railway infrastructure on population densities in regions that were not prioritized by the colonial administration and that gained access during a "second wave" of construction in the 1880s in the Constantine region in French Algeria. To do this, it uses a diff-in-diff model on two-time intervals at the end of the 1800s. To improve the comparability of the treated and control groups, it uses a PSM and restricts the sample in various ways. The results, which are additionally supported by an instrumental variable approach, show that the effect of gaining access to railways was insignificant. The settler population density displayed negative values and did not benefit

from the infrastructure. The indigenous population responded positively to the railway only in the regions nearest to the stations and endowed with settlers.

The results are consistent with the existing literature on Algerian railways. The negative effect concerning the settler population supports the crowding out of the small rural settler's hypothesis and the administration's failure to consolidate a small family-farm type of economy. Regarding the indigenous population, the positive effect is in line with authors such as Good (1961), who explains that, after the 1870s, the growth rates recovered and displayed continuous positive figures. According to Bennoune (2002) and Ruedy (2005), the persistent growth rates within an impoverished context were only feasible if the demand for labor was strong enough. In line with this, the results from the diff-in-diff support this view and show that indigenous population growth was positively (and significantly) affected by the railway only in the regions settled by Europeans and nearest to the railway stations.

A more detailed analysis of Algeria's railway lines suggests that the low impact on population densities in the later settled regions in Constantine resulted from geographical features that limited its potential gains. The lack of fertile land and the vulnerability of cereal to climate pushed companies to increase rates. Furthermore, the railway's line concession and financing system led to a lack of route competition that increased transport rates and highlighted regional inequalities.

Table 1: Base diff-in-diff: settler population density, Constantine 1884 and 1892

	RESTRICTED SAMPLE				FULL SAMPLE							
	No termini				With termini				No termini			
	N	1884	1892	1892-1884 (DD)	N	1884	1892	1892-1884 (DD)	N	1884	1892	1892-1884 (DD)
	Distance from station <20 km				Distance from station <20 km				Distance from station <20 km			
Treatment	49	0.13	0.115	-0.015	64	0.017	0.019	0.002	61	0.011	0.011	0.000
Control	157	0.243	1.086	0.843	192	0.023	0.033	0.010	191	0.018	0.024	0.006
T-C		-0.113	-0.97	-0.858		-0.006	-0.01	-0.008		-0.007	-0.013	-0.006
				(0.712)				(0.005)				(0.004)
	Distance from station <10 km				Distance from station <10 km				Distance from station <10 km			
Treatment	23	0.166	0.042	-0.124	37	0.027	0.030	0.003	34	0.018	0.018	0.000
Control	183	0.223	0.957	0.734	219	0.021	0.029	0.008	218	0.016	0.021	0.005
T-C		-0.057	-0.915	-0.858		0.006	0.001	-0.005		0.002	-0.003	-0.005
				(0.614)				(0.005)				(0.003)

Table 2: Base diff-in-diff: settler population density, Constantine 1884 and 1897

	RESTRICTED SAMPLE				FULL SAMPLE							
	No termini				With termini				No termini			
	N	1884	1897	1897-1884 (DD)	N	1884	1897	1897-1884 (DD)	N	1884	1897	1897-1884 (DD)
	Distance from station <20 km				Distance from station <20 km				Distance from station <20 km			
Treatment	49	0.13	0.769	0.639	64	0.017	0.028	0.011	61	0.011	0.019	0.008
Control	157	0.243	1.261	1.018	192	0.023	0.036	0.013	191	0.018	0.026	0.008
T-C		-0.113	-0.492	-0.379		-0.006	-0.01	-0.002		-0.007	-0.007	0.000
				(0.625)				(0.009)				(0.007)
	Distance from station <10 km				Distance from station <10 km				Distance from station <10 km			
Treatment	23	0.166	1.085	0.919	37	0.027	0.046	0.019	34	0.018	0.030	0.012
Control	183	0.223	1.152	0.929	219	0.021	0.032	0.011	218	0.016	0.023	0.007
T-C		-0.057	-0.067	-0.01		0.006	0.014	0.008		0.002	0.007	0.005
				(0.594)				(0.011)				(0.010)

Table 3: Base diff-in-diff: indigenous population density, Constantine 1884 and 1892

RESTRICTED SAMPLE					FULL SAMPLE							
No termini					With termini				No termini			
	N	1884	1897	1897-1884 (DD)	N	1884	1897	1897-1884 (DD)	N	1884	1897	1897-1884 (DD)
Distance from station <20 km					Distance from station <20 km				Distance from station <20 km			
Treatment	49	0.347	0.441	0.094	64	0.298	0.380	0.082	61	0.295	0.382	0.087
Control	157	0.286	0.350	0.064	192	0.295	0.361	0.066	191	0.287	0.350	0.063
T-C		0.061	0.091	0.030		0.003	0.019	0.016		0.008	0.032	0.024
				(0.029)				(0.025)				(0.025)
Distance from station <10 km					Distance from station <10 km				Distance from station <10 km			
Treatment	23	0.166	1.085	0.919	37	0.027	0.046	0.019	34	0.018	0.030	0.012
Control	183	0.223	1.152	0.929	219	0.021	0.032	0.011	218	0.016	0.023	0.007
T-C		-0.057	-0.067	-0.01		0.006	0.014	0.008		0.002	0.007	0.005
				(0.594)				(0.011)				(0.010)

Table 4: Base diff-in-diff: indigenous population density, Constantine 1884 and 1897

RESTRICTED SAMPLE					FULL SAMPLE							
No termini					With termini				No termini			
	N	1884	1897	1897-1884 (DD)	N	1884	1897	1897-1884 (DD)	N	1884	1897	1897-1884 (DD)
Distance from station <20 km					Distance from station <20 km				Distance from station <20 km			
Treatment	49	0.347	0.478	0.131	64	0.298	0.420	0.122	61	0.295	0.412	0.117
Control	157	0.286	0.372	0.086	192	0.295	0.382	0.087	191	0.287	0.370	0.083
T-C		0.061	0.106	0.045		0.003	0.038	0.035		0.008	0.042	0.034
				(0.043)				(0.033)				(0.033)
Distance from station <10 km					Distance from station <10 km				Distance from station <10 km			
Treatment	23	0.457	0.616	0.159	37	0.336	0.471	0.135	34	0.336	0.463	0.127
Control	183	0.281	0.370	0.089	219	0.289	0.378	0.089	218	0.282	0.367	0.085
T-C		0.176	0.246	0.070		0.047	0.093	0.046		0.054	0.096	0.042
				(0.046)				(0.029)				(0.030)

Tables 1, 2, 3, and 4: *** $p<0.01$; ** $p<0.05$; * $p<0.1$. Means and standard errors are estimated by linear regression. Clustered standard errors in parenthesis. *Source*: TGdC (1884), TGdC (1892), and TGdC (1897).

Table 5: Railway density in French Algeria, 1881-1921

Year	Algeria	Tunisia	Morocco	Egypt	South Africa	Canada	USA	Mexico	Brasil	Argentina	Chile	Australia	New Zealand
	Mileage/Population (RW kilometer per 10.000 pop)												
1881	4.0			2.2		27.3	29.4	1.1	2.8	8.6	8.4	29.2	42.3
1891	6.8			1.8		46.1	43.3	8.3	7.0	26.0	10.5	50.3	47.3
1901	6.1			3.3	15.1	54.4	41.9	11.6	8.8	35.1	15.0	57.1	46.1
1911	5.9	8.8		3.6	20.0	56.7	43.6	16.3	9.2	39.0	17.2	63.4	44.0
1913	5.9	9.4	0.2	3.6	21.7	61.3	43.5	16.4	9.3	41.9	20.3	65.0	42.9
1921	6.2	9.7	2.5	3.3	22.1	72.1	39.4	18.9	9.3	40.6	21.9	76.4	59.7
	Mileage/surface area (RW km per 100 km2)												
1881	0.2	0.2		0.2	0.1	0.1	1.9	0.1	0.0	0.1	0.3	0.1	0.8
1891	0.5	0.2		0.2	0.3	0.2	3.9	0.5	0.1	0.3	0.4	0.2	1.1
1901	0.6	0.7		0.4	0.6	0.3	4.1	0.8	0.2	0.6	0.6	0.3	1.3
1911	0.6	1.4		0.5	1.0	0.4	5.2	1.2	0.3	0.9	0.8	0.4	1.7
1913	0.6	1.5	0.0	0.5	1.1	0.5	5.4	1.3	0.3	1.1	1.0	0.4	1.7
1921	0.6	1.6	0.3	0.5	1.3	0.6	5.4	1.3	0.3	1.2	1.1	0.6	1.8

Source: ASF 1926, p. 314*

Table 6: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Settler density	0.03	0.18	0	2.71	768
Indigenous density	0.35	0.47	0	5.89	768
Dummy< 10km	0.14	0.35	0	1	768
Dummy< 20km	0.25	0.43	0	1	768
Elevation	602.6	412.1	0	1,638	768
Instrumental variable	0.11	0.31	0	1	768

These values are for Constantine in the years 1904/05 and 1913/14 (full sample)

Table 7: Pair-wise correlation matrix

Variables	Settden	Indden	D< 10km	D< 20km	Elev	IV
Settler density	1.00					
Indigenous density	0.68	1.00				
	(0.00)					
Dummy< 10km	0.01	0.05	1.00			
	(0.70)	(0.15)				
Dummy< 20km	-0.02	0.02	0.71	1.00		
	(0.54)	(0.62)	(0.00)			
Elevation	-0.13	-0.20	0.11	0.20	1.00	
	(0.00)	(0.00)	(0.00)	(0.00)		
Instrumental variable	0.12	0.07	0.73	0.54	0.10	1.00
	(0.00)	(0.05)	(0.00)	(0.00)	(0.01)	

These values are for Constantine in the years 1904/05 and 1913/14 (full sample). Standard deviation in parentheses.

Table 8: Summary statistics of *elevation* and t-test mean difference between treated and control groups

	Restricted sample				Full Sample			
	<20 km		<10 km		<20 km		<10 km	
	N	Mean	N	Mean	N	Mean	N	Mean
Control	157	566.52	183	597.20	192	554.09	219	583.71
		(32.89)		(29.99)		(29.49)		(27.37)
Treated	49	800.84	23	821.61	64	748.20	37	714.54
		(57.27)		(102.18)		(49.10)		(73.06)
Mean Difference		-234.32		-224.41		-194.11		-130.83
		(66.05)***		(106.49)**		(57.28)***		(-130.83)*

*Significant at 10%; **significant at 5%; *** significant at 1%. Standard errors in parenthesis. The unit of observation are the sub-municipal areas such as settlement centers and *douars*. Results are shown for unequal variancesalthough significance level is robust under equal variance assumption. *Source*: CGIAR Shuttle Radar Topography Mission 3 from California Institute of Technology.

Table 9: Diff-in-diff regressions: effects of railway access on settler population density, Const. 1884-1892

	Dependent variable: Settler population density																	
	Restricted						Full Sample											
	<20km			<10km			<20km						<10km					
	No Termini						Term.			No Term.			Term.			No Term.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
VARIABLES	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.
-diff	-0.859	-0.584	-0.613	-0.858	-0.691	-0.721	-0.007	-0.003	-0.003	-0.006	-0.003	-0.003	-0.006	-0.004	-0.004	-0.005	-0.004	-0.004
	(0.712)	(0.481)	(0.487)	(0.614)	(0.487)	(0.480)	(0.005)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Observations	412	412	410	412	412	406	512	512	504	504	504	502	512	512	504	504	504	496
R-squared	0.011	0.011	0.011	0.008	0.011	0.012	0.001	0.000	0.001	0.001	0.000	0.000	0.001	0.002	0.003	0.000	0.000	0.001
Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
PS	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control t0	0.243	0.268	0.256	0.223	0.254	0.250	0.0232	0.0141	0.0134	0.0179	0.0113	0.0115	0.0206	0.0169	0.0163	0.0160	0.0133	0.0127
Mean treated t0	0.130	0.130	0.133	0.166	0.166	0.182	0.0165	0.0165	0.0171	0.0109	0.0109	0.0111	0.0268	0.0268	0.0284	0.0176	0.0176	0.0187
Diff t0	-0.113	-0.138	-0.123	-0.0570	-0.0883	-0.0684	-0.00662	0.00248	0.00365	-0.00698	-0.000387	-0.000383	0.00622	0.00992	0.0120	0.00164	0.00433	0.00602
Mean control t1	1.086	0.836	0.853	0.957	0.822	0.836	0.0329	0.0198	0.0189	0.0240	0.0150	0.0153	0.0292	0.0240	0.0231	0.0214	0.0178	0.0170
Mean treated t1	0.115	0.115	0.117	0.0423	0.0423	0.0463	0.0188	0.0188	0.0194	0.0114	0.0114	0.0116	0.0300	0.0300	0.0317	0.0177	0.0177	0.0188
Diff t1	-0.972	-0.721	-0.736	-0.915	-0.779	-0.790	-0.0140	-0.000978	0.000510	-0.0125	-0.00361	-0.00367	0.000715	0.00600	0.00854	-0.00373	-9.05e-05	0.00185

Table 10: Diff-in-diff regressions: effects of railway access on indigenous population density, Const. 1884-1892

	Dependent variable: Indigenous population density																	
	Restricted						Full Sample											
	<20km			<10km			<20km						<10km					
	No Termini						Term.			No Term.			Term.			No Term.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
VARIABLES	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.
-diff	0.030	0.028	0.029	0.047	0.046	0.054	0.017	0.023	0.026	0.024	0.028	0.029	0.022	0.026	0.030	0.033	0.036	0.041
	(0.029)	(0.027)	(0.028)	(0.036)	(0.035)	(0.038)	(0.025)	(0.023)	(0.023)	(0.025)	(0.023)	(0.023)	(0.025)	(0.024)	(0.025)	(0.025)	(0.024)	(0.025)
Observations	412	412	410	412	412	406	512	512	504	504	504	502	512	512	504	504	504	496
R-squared	0.022	0.032	0.030	0.049	0.083	0.061	0.006	0.012	0.011	0.007	0.014	0.013	0.008	0.015	0.011	0.009	0.018	0.013
Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
PS	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control t0	0.286	0.258	0.256	0.281	0.261	0.259	0.295	0.263	0.260	0.287	0.258	0.258	0.289	0.277	0.275	0.282	0.270	0.268
Mean treated t0	0.347	0.347	0.340	0.457	0.457	0.401	0.298	0.298	0.273	0.295	0.295	0.289	0.336	0.336	0.295	0.336	0.336	0.291
Diff t0	0.0607	0.0891	0.0841	0.176	0.196	0.141	0.00218	0.0347	0.0135	0.00841	0.0374	0.0311	0.0469	0.0593	0.0205	0.0538	0.0652	0.0225
Mean control t1	0.350	0.324	0.321	0.347	0.328	0.325	0.361	0.323	0.318	0.350	0.317	0.316	0.356	0.340	0.336	0.346	0.332	0.328
Mean treated t1	0.441	0.441	0.434	0.570	0.570	0.520	0.380	0.380	0.357	0.382	0.382	0.376	0.425	0.425	0.387	0.433	0.433	0.392
Diff t1	0.0908	0.117	0.113	0.223	0.241	0.195	0.0194	0.0577	0.0393	0.0326	0.0656	0.0597	0.0690	0.0850	0.0509	0.0866	0.101	0.0634

Tables 9 and 10: *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors in parentheses. Covariate is *elevation*. PS is Kernel-based Propensity Score Matching Differences-in-Differences and the PS is regressed on elev. The balancing property is always satisfied. The settler population in the restricted sample is per thousand. The termini points are not included. *Sources*: TGdC (1884), TGdC (1892), and TGdC (1897). See text for detail on sources.

Table 11: Diff-in-diff regressions: effects of railway access on settler population density, Const. 1884-1897

	Dependent variable: Settler population density																	
	Restricted						Full Sample											
	<20km			<10km			<20km						<10km					
	No Termini						Term.			No Term.			Term.			No Term.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
VARIABLES	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.
-diff	-0.379	-0.092	-0.111	-0.009	0.158	0.232	-0.001	0.004	0.005	-0.000	0.003	0.003	0.008	0.010	0.012	0.006	0.007	0.008
	(0.625)	(0.473)	(0.482)	(0.685)	(0.622)	(0.667)	(0.009)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)
Observations	412	412	410	412	412	406	512	512	504	504	504	502	512	512	504	504	504	496
R-squared	0.019	0.019	0.020	0.017	0.025	0.028	0.001	0.002	0.003	0.001	0.001	0.001	0.002	0.007	0.008	0.001	0.003	0.004
Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
PS	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control t0	0.243	0.268	0.256	0.223	0.254	0.250	0.0232	0.0141	0.0134	0.0179	0.0113	0.0115	0.0206	0.0169	0.0163	0.0160	0.0133	0.0127
Mean treated t0	0.130	0.130	0.133	0.166	0.166	0.182	0.0165	0.0165	0.0171	0.0109	0.0109	0.0111	0.0268	0.0268	0.0284	0.0176	0.0176	0.0187
Diff t0	-0.113	-0.138	-0.123	-0.0570	-0.0883	-0.0684	-0.00662	0.00248	0.00365	-0.00698	-0.000387	-0.000383	0.00622	0.00992	0.0120	0.00164	0.00433	0.00602
Mean control t1	1.261	0.999	1.019	1.152	1.016	1.025	0.0356	0.0214	0.0205	0.0256	0.0160	0.0163	0.0317	0.0259	0.0251	0.0229	0.0190	0.0181
Mean treated t1	0.769	0.769	0.785	1.085	1.085	1.189	0.0283	0.0283	0.0292	0.0185	0.0185	0.0188	0.0463	0.0463	0.0489	0.0304	0.0304	0.0323
Diff t1	-0.492	-0.230	-0.234	-0.0662	0.0696	0.164	-0.00733	0.00685	0.00869	-0.00709	0.00245	0.00250	0.0146	0.0204	0.0239	0.00749	0.0114	0.0142

Table 12: Diff-in-diff regressions: effects of railway access on indigenous population density, Const. 1884-1897

	Dependent variable: Indigenous population density																	
	Restricted						Full Sample											
	<20km			<10km			<20km						<10km					
	No Termini						Term.			No Term.			Term.			No Term.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
VARIABLES	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.
-diff	0.045	0.042	0.042	0.070	0.068	0.077	0.035	0.040	0.043	0.034	0.037	0.037	0.046	0.050*	0.054*	0.042	0.044	0.048
	(0.043)	(0.042)	(0.042)	(0.046)	(0.046)	(0.048)	(0.033)	(0.031)	(0.031)	(0.033)	(0.032)	(0.032)	(0.029)	(0.028)	(0.028)	(0.030)	(0.030)	(0.030)
Observations	412	412	410	412	412	406	512	512	504	504	504	502	512	512	504	504	504	496
R-squared	0.032	0.041	0.039	0.060	0.092	0.071	0.011	0.021	0.019	0.011	0.020	0.019	0.014	0.025	0.021	0.014	0.024	0.019
Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
PS	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control t0	0.286	0.258	0.256	0.281	0.261	0.259	0.295	0.263	0.260	0.287	0.258	0.258	0.289	0.277	0.275	0.282	0.270	0.268
Mean treated t0	0.347	0.347	0.340	0.457	0.457	0.401	0.298	0.298	0.273	0.295	0.295	0.289	0.336	0.336	0.295	0.336	0.336	0.291
Diff t0	0.0607	0.0891	0.0841	0.176	0.196	0.141	0.00218	0.0347	0.0135	0.00841	0.0374	0.0311	0.0469	0.0593	0.0205	0.0538	0.0652	0.0225
Mean control t1	0.372	0.347	0.344	0.370	0.352	0.348	0.382	0.345	0.340	0.370	0.338	0.337	0.378	0.362	0.358	0.367	0.353	0.350
Mean treated t1	0.478	0.478	0.470	0.616	0.616	0.566	0.420	0.420	0.396	0.412	0.412	0.405	0.471	0.471	0.433	0.463	0.463	0.420
Diff t1	0.106	0.131	0.126	0.246	0.264	0.218	0.0371	0.0747	0.0562	0.0422	0.0740	0.0676	0.0930	0.109	0.0746	0.0954	0.110	0.0707

Tables 9 and 10: *** p<0.01, ** p<0.05, * p<0.1. Cluster-robust standard errors in parentheses. Covariate is *elevation*. PS is Kernel-based Propensity Score Matching Differences-in-Differences and the PS is regressed on elev. The balancing property is always satisfted. The settler population in the restricted sample is per thousand. The termini points are not included. *Sources*: TGdC (1884, 1892, and 1897); see Appendix **??** for more detail on sources.

Table 13: Diff-in-diff regressions: effects of railway access on settler population density (*indden*<¡2), Const. 1884-1892

	Dependent variable: Settler population density																	
	Restricted						Full Sample											
	<20km			<10km			<20km						<10km					
	No Termini						Term.			No Term.			Term.			No Term.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
VARIABLES	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.
-diff	-0.862	-0.585	-0.608	-0.860	-0.693	-0.732	4.968	3.467	3.598	-2.270	-1.607	-1.636	5.936	5.388	5.867	-2.123	-1.880	-1.569
	(0.717)	(0.485)	(0.487)	(0.617)	(0.491)	(0.504)	(5.753)	(3.670)	(3.675)	(1.474)	(1.284)	(1.306)	(5.461)	(4.753)	(4.960)	(1.508)	(1.445)	(1.512)
Observations	409	409	407	409	409	407	506	506	498	499	499	497	506	506	498	499	499	487
R-squared	0.011	0.011	0.011	0.008	0.011	0.012	0.002	0.008	0.010	0.002	0.005	0.005	0.012	0.027	0.033	0.013	0.025	0.038
Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
PS	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control to	0.245	0.269	0.255	0.224	0.255	0.242	13.69	9.975	9.578	8.299	6.859	6.946	12.29	10.90	10.44	7.567	7.072	6.647
Mean treated to	0.130	0.130	0.133	0.166	0.166	0.173	16.54	16.54	17.07	10.89	10.89	11.07	26.82	26.82	28.35	17.59	17.59	19.28
Diff to	-0.115	-0.139	-0.122	-0.0582	-0.0893	-0.0691	2.841	6.560	7.491	2.588	4.028	4.123	14.52	15.92	17.91	10.03	10.52	12.63
Mean control t1	1.093	0.840	0.850	0.963	0.827	0.847	11.30	9.085	8.649	11.30	9.200	9.330	10.33	9.486	8.830	10.33	9.594	8.998
Mean treated t1	0.117	0.117	0.119	0.0442	0.0442	0.0463	19.11	19.11	19.74	11.62	11.62	11.82	30.79	30.79	32.60	18.23	18.23	20.06
Diff t1	-0.976	-0.723	-0.730	-0.918	-0.783	-0.801	7.809	10.03	11.09	0.318	2.421	2.487	20.46	21.31	23.77	7.902	8.640	11.06

Table 14: Diff-in-diff regressions: effects of railway access on indigenous population density (*indden*<¡2), Const. 1884-1892

	Dependent variable: Indigenous population density																	
	Restricted						Full Sample											
	<20km			<10km			<20km						<10km					
	Termini=No Termini						Term.			No Term.			Term.			No Term.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
VARIABLES	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.
-diff	-0.006	-0.007	-0.007	-0.022	-0.023	-0.025	0.006	0.004	0.005	0.001	0.002	0.001	-0.009	-0.009	-0.008	-0.012	-0.011	-0.009
	(0.036)	(0.035)	(0.035)	(0.078)	(0.077)	(0.080)	(0.029)	(0.028)	(0.028)	(0.029)	(0.028)	(0.028)	(0.049)	(0.049)	(0.052)	(0.053)	(0.052)	(0.057)
Observations	409	409	407	409	409	407	506	506	498	499	499	497	506	506	498	499	499	487
R-squared	0.024	0.030	0.028	0.050	0.076	0.066	0.012	0.016	0.012	0.016	0.018	0.016	0.016	0.020	0.011	0.020	0.023	0.013
Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
PS	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control to	0.273	0.248	0.245	0.269	0.251	0.252	0.262	0.241	0.239	0.253	0.237	0.236	0.260	0.251	0.250	0.252	0.245	0.244
Mean treated to	0.347	0.347	0.340	0.457	0.457	0.446	0.298	0.298	0.273	0.295	0.295	0.289	0.336	0.336	0.295	0.336	0.336	0.297
Diff to	0.0745	0.0994	0.0946	0.188	0.206	0.194	0.0358	0.0562	0.0346	0.0423	0.0588	0.0527	0.0765	0.0847	0.0458	0.0835	0.0910	0.0529
Mean control t1	0.339	0.315	0.312	0.337	0.320	0.321	0.312	0.294	0.290	0.312	0.295	0.294	0.313	0.304	0.301	0.313	0.304	0.302
Mean treated t1	0.408	0.408	0.400	0.503	0.503	0.489	0.354	0.354	0.330	0.355	0.355	0.348	0.380	0.380	0.338	0.384	0.384	0.346
Diff t1	0.0685	0.0921	0.0871	0.165	0.182	0.169	0.0421	0.0600	0.0397	0.0431	0.0604	0.0540	0.0670	0.0755	0.0374	0.0712	0.0803	0.0437

Note Tables 9 and 10: Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Covariate is *elevation*. PS is Kernel-based Propensity Score Matching Differences-in- Differences and the PS is regressed on elev. The balancing property is always satisfted. The settler population in the restricted sample is per thousand. The termini points are not included. *Sources*: TGdC (1884, 1892, and 1897); see Appendix **??** for more detail on sources.

Table 15: Diff-in-diff regressions: effects of railway access on settler population density (*indden*<¡2), Const. 1884-1897

	Dependent variable: Settler population density																	
	Restricted						Full Sample											
	<20km			<10km			<20km						<10km					
	No Termini						Term.			No Term.			Term.			No Term.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
VARIABLES	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.
-diff	-0.429	-0.139	-0.151	0.048	0.213	0.243	0.015*	0.013*	0.014*	0.005	0.005	0.006	0.022*	0.022*	0.023*	0.011	0.011	0.011
	(0.639)	(0.488)	(0.493)	(0.687)	(0.623)	(0.667)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.012)	(0.011)	(0.012)	(0.010)	(0.010)	(0.010)
Observations	408	408	406	408	408	402	505	505	493	498	498	496	505	505	499	498	498	496
R-squared	0.019	0.018	0.019	0.017	0.026	0.031	0.007	0.017	0.020	0.008	0.012	0.012	0.024	0.042	0.045	0.026	0.032	0.033
Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
PS	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control to	0.245	0.269	0.254	0.224	0.255	0.229	0.0137	0.00997	0.00971	0.00830	0.00686	0.00693	0.0123	0.0109	0.0115	0.00757	0.00707	0.00724
Mean treated to	0.130	0.130	0.133	0.166	0.166	0.154	0.0165	0.0165	0.0173	0.0109	0.0109	0.0111	0.0268	0.0268	0.0283	0.0176	0.0176	0.0181
Diff to	-0.115	-0.139	-0.121	-0.0582	-0.0893	-0.0757	0.00284	0.00656	0.00763	0.00259	0.00403	0.00414	0.0145	0.0159	0.0169	0.0100	0.0105	0.0109
Mean control t1	1.269	1.004	1.013	1.145	1.011	1.081	0.0116	0.00946	0.00916	0.0116	0.00956	0.00968	0.0106	0.00978	0.0102	0.0106	0.00988	0.0101
Mean treated t1	0.726	0.726	0.741	1.135	1.135	1.248	0.0291	0.0291	0.0306	0.0191	0.0191	0.0194	0.0476	0.0476	0.0504	0.0313	0.0313	0.0323
Diff t1	-0.544	-0.278	-0.272	-0.00989	0.123	0.167	0.0176	0.0197	0.0215	0.00752	0.00950	0.00972	0.0370	0.0378	0.0402	0.0207	0.0214	0.0221

Table 16: Diff-in-diff regressions: effects of railway access on indigenous population density (*indden*<¡2), Const. 1884-1897

	Dependent variable: Indigenous population density																	
	Restricted						Full Sample											
	<20km			<10km			<20km						<10km					
	No Termini						Term.			No Term.			Term.			No Term.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
VARIABLES	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.
-diff	-0.032	-0.035	-0.036	-0.000	-0.004	-0.003	-0.006	-0.010	-0.009	-0.023	-0.024	-0.025	0.018	0.017	0.017	-0.003	-0.002	-0.005
	(0.051)	(0.050)	(0.050)	(0.069)	(0.069)	(0.072)	(0.045)	(0.043)	(0.045)	(0.046)	(0.045)	(0.044)	(0.049)	(0.049)	(0.050)	(0.053)	(0.053)	(0.052)
Observations	408	408	406	408	408	402	505	505	493	498	498	496	505	505	499	498	498	496
R-squared	0.029	0.030	0.027	0.063	0.086	0.087	0.017	0.019	0.016	0.020	0.018	0.016	0.026	0.031	0.028	0.028	0.028	0.023
Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
PS	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control to	0.273	0.248	0.245	0.269	0.251	0.251	0.262	0.241	0.238	0.253	0.237	0.236	0.260	0.251	0.254	0.252	0.245	0.247
Mean treated to	0.347	0.347	0.340	0.457	0.457	0.463	0.298	0.298	0.276	0.295	0.295	0.289	0.336	0.336	0.333	0.336	0.336	0.325
Diff to	0.0745	0.0994	0.0946	0.188	0.206	0.212	0.0358	0.0562	0.0386	0.0423	0.0588	0.0527	0.0765	0.0847	0.0785	0.0835	0.0910	0.0777
Mean control t1	0.361	0.339	0.335	0.350	0.336	0.332	0.332	0.315	0.309	0.332	0.316	0.315	0.326	0.318	0.319	0.326	0.317	0.319
Mean treated t1	0.404	0.404	0.393	0.538	0.538	0.540	0.361	0.361	0.338	0.351	0.351	0.342	0.420	0.420	0.414	0.406	0.406	0.392
Diff t1	0.0427	0.0647	0.0583	0.187	0.201	0.208	0.0294	0.0460	0.0293	0.0189	0.0349	0.0273	0.0940	0.102	0.0958	0.0805	0.0891	0.0727

Note Tables 9 and 10: Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Covariate is *elevation*. PS is Kernel-based Propensity Score Matching Differences-in- Differences and the PS is regressed on elev. The balancing property is always satisfted. The settler population in the restricted sample is per thousand. The termini points are not included. *Sources*: TGdC (1884, 1892, and 1897); see Appendix **??** for more detail on sources.

Table 17: Diff-in-diff (IV) regressions: effects of railway access on population density, Constantine 1884-1892

	Dependent: Settler population density									Dependent: Indigenous population density								
	Restricted			Full						Restricted			Full					
	No Termini			Termini			No Termini			No termini			Termini			No Termini		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
VARIABLES	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.
-diff	-0.903	-0.620	-0.948	0.023	0.024	0.000	-0.005	-0.004	-0.002	0.042	0.042*	0.056**	0.037	0.041*	0.035*	0.027	0.031	0.046**
	(0.587)	(0.402)	(0.661)	(0.026)	(0.026)	(0.003)	(0.003)	(0.003)	(0.002)	(0.026)	(0.024)	(0.027)	(0.023)	(0.023)	(0.020)	(0.021)	(0.020)	(0.019)
Observations	412	412	310	512	512	402	504	504	374	412	412	310	512	512	402	504	504	374
R-squared	0.008	0.009	0.014	0.012	0.022	0.053	0.000	0.002	0.062	0.029	0.061	0.068	0.010	0.020	0.017	0.007	0.012	0.020
Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
PS	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control t0	0.218	0.246	0.216	0.0168	0.0139	0.00819	0.0156	0.0128	0.00596	0.290	0.270	0.261	0.288	0.275	0.249	0.288	0.275	0.244
Mean treated t0	0.199	0.199	0.185	0.0616	0.0616	0.0263	0.0218	0.0218	0.0238	0.436	0.436	0.442	0.362	0.362	0.295	0.303	0.303	0.294
Diff t0	-0.0189	-0.0471	-0.0320	0.0449	0.0478	0.0181	0.00622	0.00904	0.0178	0.146	0.165	0.181	0.0740	0.0872	0.0458	0.0155	0.0285	0.0507
Mean control t1	0.922	0.667	0.980	0.0222	0.0183	0.0106	0.0208	0.0170	0.00826	0.358	0.338	0.323	0.354	0.336	0.302	0.354	0.337	0.295
Mean treated t1	0	0	0	0.0899	0.0899	0.0289	0.0220	0.0220	0.0240	0.546	0.546	0.559	0.465	0.465	0.383	0.397	0.397	0.391
Diff t1	-0.922	-0.667	-0.980	0.0677	0.0716	0.0182	0.00123	0.00504	0.0158	0.188	0.208	0.236	0.111	0.128	0.0809	0.0428	0.0593	0.0963

Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Covariate is *elevation*. PS is Kernel-based Propensity Score Matching Differences-in-Differences and the PS is regressed on elev. The balancing property is always satisfted. The settler population in the restricted sample is per thousand. *Sources*: TGdC (1884, 1892, and 1897); see Appendix **??** for more detail on sources. The instrument is a dummy variable equal to one if the region touches the straight lines as shown in Figure 15 and explained in section **??**.

Table 18: Diff-in-diff (IV) regressions: effects of railway access on population density, Constantine 1884-1897

	Dependent: Settler population density									Dependent: Indigenous population density								
	Restricted			Full						Restricted			Full					
	No Termini			Termini			No Termini			No Termini			Termini			No Termini		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
VARIABLES	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.	No Cov.	Matched	C.Supp.
-diff	-0.492	-0.159	-0.314	0.047	0.048	0.020	0.011	0.012	0.016	0.063	0.058	0.072	0.071*	0.074**	0.057**	0.033	0.036	0.047*
	(0.675)	(0.575)	(0.755)	(0.035)	(0.034)	(0.013)	(0.014)	(0.013)	(0.014)	(0.050)	(0.049)	(0.055)	(0.037)	(0.036)	(0.027)	(0.027)	(0.026)	(0.026)
Observations	412	412	310	512	512	402	504	504	374	412	412	310	512	512	402	504	504	374
R-squared	0.018	0.014	0.018	0.017	0.031	0.061	0.001	0.008	0.057	0.039	0.066	0.072	0.017	0.030	0.028	0.011	0.017	0.024
Covariates	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
PS	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control t0	0.218	0.246	0.216	0.0168	0.0139	0.00819	0.0156	0.0128	0.00596	0.290	0.270	0.261	0.288	0.275	0.249	0.288	0.275	0.244
Mean treated t0	0.199	0.199	0.185	0.0616	0.0616	0.0263	0.0218	0.0218	0.0238	0.436	0.436	0.442	0.362	0.362	0.295	0.303	0.303	0.294
Diff t0	-0.0189	-0.0471	-0.0320	0.0449	0.0478	0.0181	0.00622	0.00904	0.0178	0.146	0.165	0.181	0.0740	0.0872	0.0458	0.0155	0.0285	0.0507
Mean control t1	1.181	0.876	1.119	0.0241	0.0199	0.0117	0.0222	0.0182	0.00871	0.382	0.367	0.350	0.376	0.360	0.327	0.376	0.360	0.318
Mean treated t1	0.671	0.671	0.774	0.116	0.116	0.0494	0.0394	0.0394	0.0430	0.590	0.590	0.602	0.521	0.521	0.429	0.424	0.424	0.416
Diff t1	-0.511	-0.206	-0.346	0.0920	0.0961	0.0376	0.0172	0.0212	0.0343	0.208	0.223	0.252	0.145	0.162	0.103	0.0483	0.0642	0.0978

Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Covariate is *elevation*. PS is Kernel-based Propensity Score Matching Differences-in-Differences and the PS is regressed on elev. The balancing property is always satisfted. The settler population in the restricted sample is per thousand. *Sources*: TGdC (1884, 1892, and 1897); see Appendix **??** for more detail on sources. The instrument is a dummy variable equal to one if the region touches the straight lines as shown in Figure 15 and explained in section **??**.

Table 19: Diff-in-diff: effect of railway access on settler and indigenous population density, economic activity redistribution, Constantine 1884-1892

Drop variable:	Dependent variable: Settler population density									Dependent variable: Indigenous population density								
	Restricted			Full						Restricted			Full					
	Baseline	10 to 20 km	20 to 30 km	Baseline		10 to 20 km		20 to 30 km		Baseline	10 to 20 km	20 to 30 km	Baseline		10 to 20 km		20 to 30 km	
	No Termini			Term.	No Term.	Term	No Term.	Term.	No Term.	No Termini			Term.	No Term.	Term.	No Term.	Term.	No Term.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
diff	-0.721	-1.015	-0.528	-0.004	-0.004	-0.007	-0.006*	-0.005	-0.005	0.054	0.050	0.054	0.030	0.041	0.024	0.035	0.028	0.034
	(0.480)	(0.712)	(0.318)	(0.004)	(0.003)	(0.006)	(0.004)	(0.005)	(0.003)	(0.038)	(0.039)	(0.039)	(0.025)	(0.025)	(0.027)	(0.026)	(0.026)	(0.026)
Observations	406	296	352	504	496	384	376	440	434	406	296	352	504	496	384	376	440	434
R-squared	0.012	0.016	0.013	0.003	0.001	0.001	0.001	0.002	0.000	0.061	0.060	0.078	0.011	0.013	0.010	0.012	0.010	0.013
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control to	0.250	0.254	0.241	0.0163	0.0127	0.0233	0.0179	0.0186	0.0154	0.259	0.254	0.271	0.275	0.268	0.289	0.281	0.290	0.289
Mean treated to	0.182	0.182	0.154	0.0284	0.0187	0.0276	0.0181	0.0292	0.0187	0.401	0.401	0.463	0.295	0.291	0.326	0.325	0.301	0.332
Diff to	-0.0684	-0.0722	-0.0871	0.0120	0.00602	0.00427	0.000268	0.0106	0.00330	0.141	0.146	0.191	0.0205	0.0225	0.0374	0.0434	0.0110	0.0425
Mean control t1	0.836	1.134	0.661	0.0231	0.0170	0.0336	0.0243	0.0268	0.0208	0.325	0.325	0.335	0.336	0.328	0.354	0.344	0.356	0.354
Mean treated t1	0.0463	0.0463	0.0463	0.0317	0.0188	0.0308	0.0182	0.0326	0.0188	0.520	0.520	0.580	0.387	0.392	0.415	0.422	0.394	0.431
Diff t1	-0.790	-1.087	-0.615	0.00854	0.00185	-0.00279	-0.00607	0.00585	-0.00204	0.195	0.196	0.245	0.0509	0.0634	0.0610	0.0780	0.0387	0.0767

Table 20: Diff-in-diff: effect of railway access on settler and indigenous population density, economic activity redistribution, Constantine 1884-1897

Drop variable:	Dependent variable: Settler population density									Dependent variable: Indigenous population density								
	Restricted			Full						Restricted			Full					
	Baseline	10 to 20 km	20 to 30 km	Baseline		10 to 20 km		20 to 30 km		Baseline	10 to 20 km	20 to 30 km	Baseline		10 to 20 km		20 to 30 km	
	No Termini			Term.	No Term.	Term.	No Term.	Term.	No Term.	Termini=No Termini			Term.	No Term.	Term.	No Term.	Term.	No Term.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
diff	-0.314	-0.727	0.064	0.020	0.016	0.054	0.015	0.055	0.016	0.077	0.066	0.077	0.054*	0.048	0.039	0.035	0.052*	0.041
	(0.755)	(0.973)	(0.612)	(0.013)	(0.014)	(0.037)	(0.014)	(0.037)	(0.014)	(0.048)	(0.048)	(0.048)	(0.028)	(0.030)	(0.030)	(0.030)	(0.029)	(0.030)
Observations	310	228	274	402	374	308	282	356	334	406	296	352	504	496	384	376	440	434
R-squared	0.018	0.022	0.018	0.061	0.057	0.042	0.046	0.045	0.059	0.071	0.069	0.087	0.021	0.019	0.018	0.017	0.019	0.018
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean control to	0.216	0.237	0.225	0.00819	0.00596	0.0108	0.00743	0.00823	0.00531	0.259	0.254	0.271	0.275	0.268	0.289	0.281	0.290	0.289
Mean treated to	0.185	0.213	0.200	0.0263	0.0238	0.0665	0.0227	0.0665	0.0238	0.401	0.401	0.463	0.295	0.291	0.326	0.325	0.301	0.332
Diff to	-0.0320	-0.0240	-0.0249	0.0181	0.0178	0.0558	0.0153	0.0583	0.0184	0.141	0.146	0.191	0.0205	0.0225	0.0374	0.0434	0.0110	0.0425
Mean control t1	1.119	1.470	0.800	0.0117	0.00871	0.0157	0.0112	0.0124	0.00837	0.348	0.354	0.357	0.358	0.350	0.383	0.372	0.379	0.376
Mean treated t1	0.774	0.718	0.838	0.0494	0.0430	0.125	0.0411	0.125	0.0430	0.566	0.566	0.626	0.433	0.420	0.460	0.451	0.441	0.460
Diff t1	-0.346	-0.751	0.0386	0.0376	0.0343	0.110	0.0299	0.113	0.0346	0.218	0.212	0.269	0.0746	0.0707	0.0769	0.0784	0.0626	0.0835

Table 19 and 20: Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Covariate is *elevation*. The treated variable are the regions below 10 km from nearest railway station. With ftxed effects the coefficients and standard errors are almost identical. All regressions are PS Kernel-based Propensity Score Matching Differences-in-Differences and the PS is regressed on elev. The balancing property is satisfted and all samples are restricted to common support. The settler population in the restricted sample is per thousand. *Sources*: TGdC (1884, 1892, and 1897)

References

Ageron, C. R. (1968). *Les musulmans Algériens et la France, 1870-1919*. Paris: Presses Universitaires de France.

Amin, S. (1966). *L'Économie du Maghreb*, volume 1. Éditions de Minuit.

Atack, J., Bateman, F., Haines, M., and Margo, R. A. (2010). Did railroads induce or follow economic growth? *Social Science History*, 34(02):171–197.

Atack, J. and Margo, R. A. (2011). The impact of access to rail transportation on agricul- tural improvement: the American Midwest as a test case, 1850-1860. *Journal of Transport and Land Use*, 4(2).

Banerjee, A., Duflo, E., and Qian, N. (2012). On the road: access to transportation infras- tructure and economic growth in China. Working Paper No. 17897, National Bureau of Economic Research.

Belkacemi, B. (1984). *French railways in Algeria, 1850-1900: a contribution to the study of colonial history.* PhD thesis, University of East Anglia.

Bellahsene, T. (2006). *La colonisation en Algérie: processus et procédures de création des centres de peuplement: institutions, intervenants et outils*. PhD thesis, University of Paris 8.

Bennoune, M. (2002). *The making of contemporary Algeria, 1830-1987*. Cambridge Uni- versity Press, 2 edition.

Berger, T. (2016). *Engines of growth: essays in Swedish economic history*. PhD thesis, Lund University.

Berger, T. and Enflo, K. (2017). Locomotives of local growth: the short-and long-term impact of railroads in Sweden. *Journal of Urban Economics*.

Bertazzini, M. C. (2018). The long-term impact of Italian colonial roads in the horn of Africa. 1935-2000. Economic history working paper No. 272/2018, LSE.

Bobrie, F. (1976). *Finances publiques et conquête coloniale: le coût budgétaire del'expansion francaise entre 1850 et 1913. Annales. Éonomies, Sociétés, Civilisations*, 31(6):1225–1244.

Bogart, D., Chaudhary, L., and Herranz-Loncán, A. (2015). The growth contribution

of colonial Indian railways in comparative perspective. Working paper, Available at SSRN 257689

Coatsworth, J. H. (1979). Indispensable railroads in a backward economy: the case of Mexico. *The Journal of Economic History*, 39(04):939–960.

Coontz, S. H. (1957). *Population and the economic interpretation*. London: Routledge.

Donaldson, D. (2010). Railroads of the Raj: Estimating the impact of transportation infrastructure, forthcoming. *American Economic Review*.

Donaldson, D. and Hornbeck, R. (2012). Railroads and American economic growth: new data and theory. *NBER Working Paper*, 19213.

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

Fourie, J. and Herranz-Loncán, A. (2004). Growth (and segregation) by rail: how the rail- ways shaped colonial South Africa. Working Paper no 539, Economic Research Southern Africa (ERSA).

Good, D. (1961). Notes on the demography of Algeria. *Population Index*, 27(01):3–32.

Gregory, I. N. and Henneberg, J. M. (2010). The railways, urbanization, and local demography in England and Wales, 1825–1911. *Social Science History*, 34(2):199–228.

Griffin, K. B. (1976). *Land concentration and rural poverty*. Holmes & Meier.

Harter, J. (2005). *World railways of the nineteenth century: a pictorial history in Victorian engravings*. JHU Press.

Herranz-Loncán, A. (2011). The role of railways in export-led growth: The case of Uruguay, 1870–1913. *Economic History of Developing Regions*, 26(2):1–32.

Herranz-Loncán, A. and Fourie, J. (2017). "For the public benefit"? railways in the British Cape colony. *European Review of Economic History*, pages 1–28.

Hornung, E. (2013). Railroads and micro-regional growth in Prussia. Ifo Working Paper 127, Leibniz Institute for Economic Research at the University of Munich.

Jedwab, R., Kerby, E., and Moradi, A. (2017). History, path dependence and

development: Evidence from colonial railways, settlers and cities in Kenya. *The Economic Journal*, 127(603):1467–1494.

Jedwab, R. and Moradi, A. (2016). The permanent effects of transportation revolutions in poor countries: evidence from Africa. *Review of Economics and Statistics*, 98(2):268–284.

Lützelschwab, C. (2000). *La Compagnie genevoise des Colonies suisses. De Sétif (Algérie) et les innovations agricoles de son directeur Gottlieb Ryf (1884-1903), révélatrices des mutations sociales de l'Algérie coloniale*. *Revue Française d'Histoire d'Outre-Mer*, 87(328):185– 207.

Maddison, A. (2006). *The World Economy: Historical statistics*, volume 2. Paris, France: Development Centre Studies, Organisation for Economic Co-operation and Development (OECD).

Mitchell, B. R. (1988). *International historical statistics: Europe 1750-1988*. New York: Stockton Press.

Mitchell, B. R. (1994). *International historical statistics: Africa, Asia, and Oceania 1750- 1988*. New York: Stockton Press.

Nouschi, A. (1961). *Enquête sur le niveau de vie des populations rurales Constantinoises: de la conquête jusqu'en 1919: essai d'histoire économique et sociale*. Paris: Presses Universitaires de France.

Prados de la Escosura, L. (2012). Output per head in pre-independence Africa: quantitative conjectures. *Economic History of Developing Regions*, 27(2):1–36.

Prochaska, D. (2004). *Making Algeria French: colonialism in Bône, 1870-1920*. Cambridge University Press.

Redding, S. J. and Turner, M. A. (2014). Transportation costs and the spatial organization of economic activity. Working Paper no 20235, National Bureau of Economic Research.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Ruedy, J. D. (2005). *Modern Algeria: the origins and development of a nation*. Georgetown University Press.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward.*Statistical Science: a Review Journal of the Institute of Mathematical Statistics*, 25(1):1.

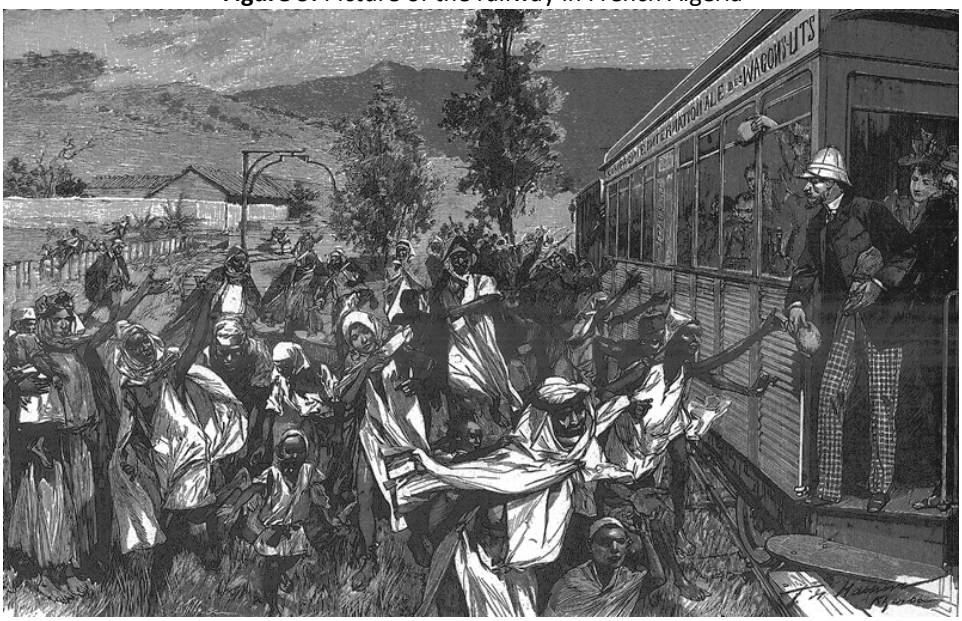
Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M. E., and Barry, C. L. (2014). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Services and Outcomes Research Methodology*, 14(4):166–182.

Tang, J. P. (2014). Railroad expansion and industrialization: evidence from Meiji Japan.*The Journal of Economic History*, 74(03):863–886.

Yacono, X. (1993). *Histoire de l'Algérie: de la fin de la Régence turque à l'insurrection de 1954*. Atlanthrope.

Appendix

Figure 9: Picture of the railway in French Algeria



Source: Harter (2005, p. 243)

Table 6: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Settler density	0.03	0.18	0	2.71	768
Indigenous density	0.35	0.47	0	5.89	768
Dummy< 10km	0.14	0.35	0	1	768
Dummy< 20km	0.25	0.43	0	1	768
Elevation	602.6	412.1	0	1,638	768
Instrumental variable	0.11	0.31	0	1	768

These values are for Constantine in the years 1904/05 and 1913/14 (full sample).

Table 7: Pair-wise correlation matrix

Variables	Settden	Indden	D< 10km	D< 20km	Elev	IV
Settler density	1.00					
Indigenous density	0.68	1.00				
	(0.00)					
Dummy< 10km	0.01	0.05	1.00			
	(0.70)	(0.15)				
Dummy< 20km	-0.02	0.02	0.71	1.00		
	(0.54)	(0.62)	(0.00)			
Elevation	-0.13	-0.20	0.11	0.20	1.00	
	(0.00)	(0.00)	(0.00)	(0.00)		
Instrumental variable	0.12	0.07	0.73	0.54	0.10	1.00
	(0.00)	(0.05)	(0.00)	(0.00)	(0.01)	

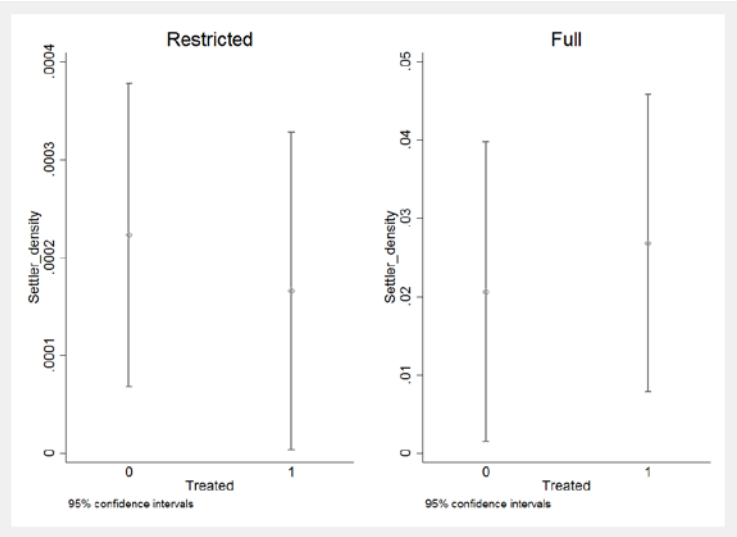
These values are for Constantine in the years 1904/05 and 1913/14 (full sample). Standard deviation in parentheses.

Table 8: Summary statistics of *elevation* and t-test mean difference between treated and control groups

	Restricted sample				Full Sample			
	<20 km		<10 km		<20 km		<10 km	
	N	Mean	N	Mean	N	Mean	N	Mean
Control	157	566.52	183	597.20	192	554.09	219	583.71
		(32.89)		(29.99)		(29.49)		(27.37)
Treated	49	800.84	23	821.61	64	748.20	37	714.54
		(57.27)		(102.18)		(49.10)		(73.06)
Mean Difference		-234.32		-224.41		-194.11		-130.83
		(66.05)***		(106.49)**		(57.28)***		(-130.83)*

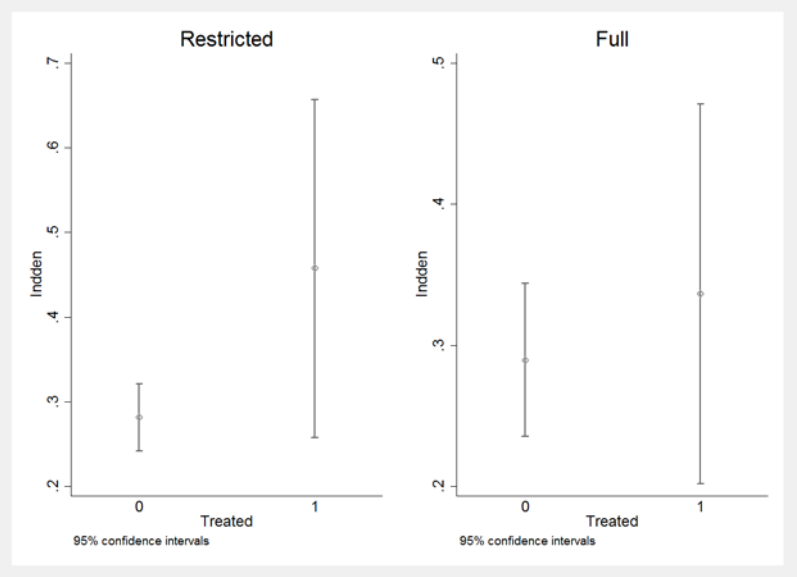
*Significant at 10%; **significant at 5%; *** significant at 1%. Standard errors in parenthesis. The unit of observation are the sub-municipal areas such as settlement centers and *douars*. Results are shown for unequal variancesalthough significance level is robust under equal variance assumption. *Source*: CGIAR Shuttle Radar Topography Mission 3 from California Institute of Technology.

F**igure 10:** Mean and confidence intervals of settler population density (by treated and control groups) in year 1884



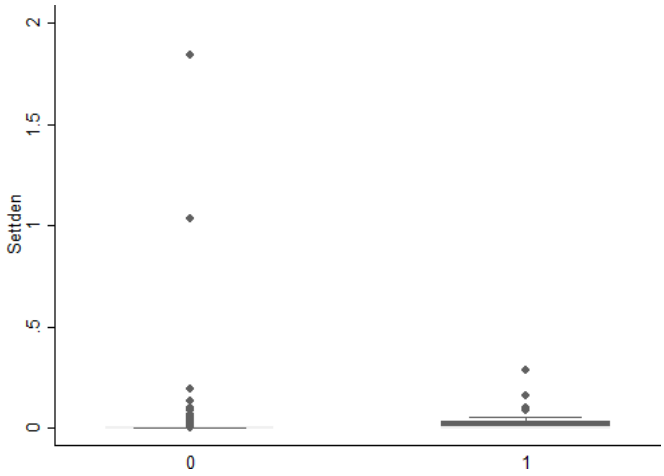
Source: Tableau Général des Communes (1884)

Figure 11: Mean and confidence intervals of indigenous population density (by treated and control groups) in year 1884



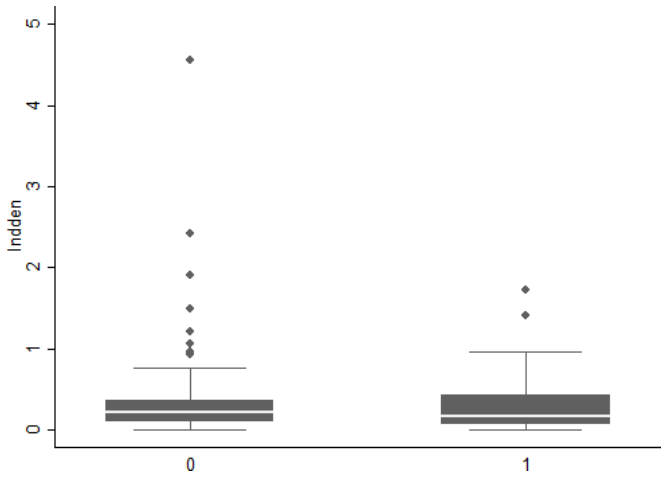
Source: Tableau Général des Communes (1884).

Figure 12: Box plot of settler population density (by treated and control groups) in year 1884



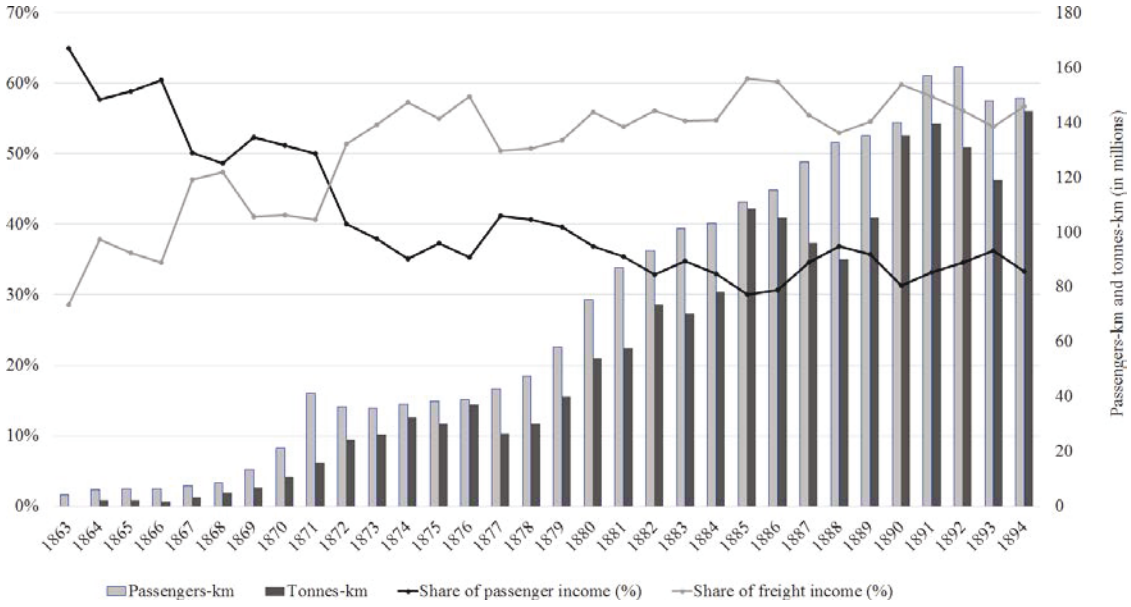
X-axis: control group is labeled 0 and treated group is 1. *Source*: TGdC (1884).

Figure 13: Box plot of indigenous population density (by treated and control groups) in year 1884



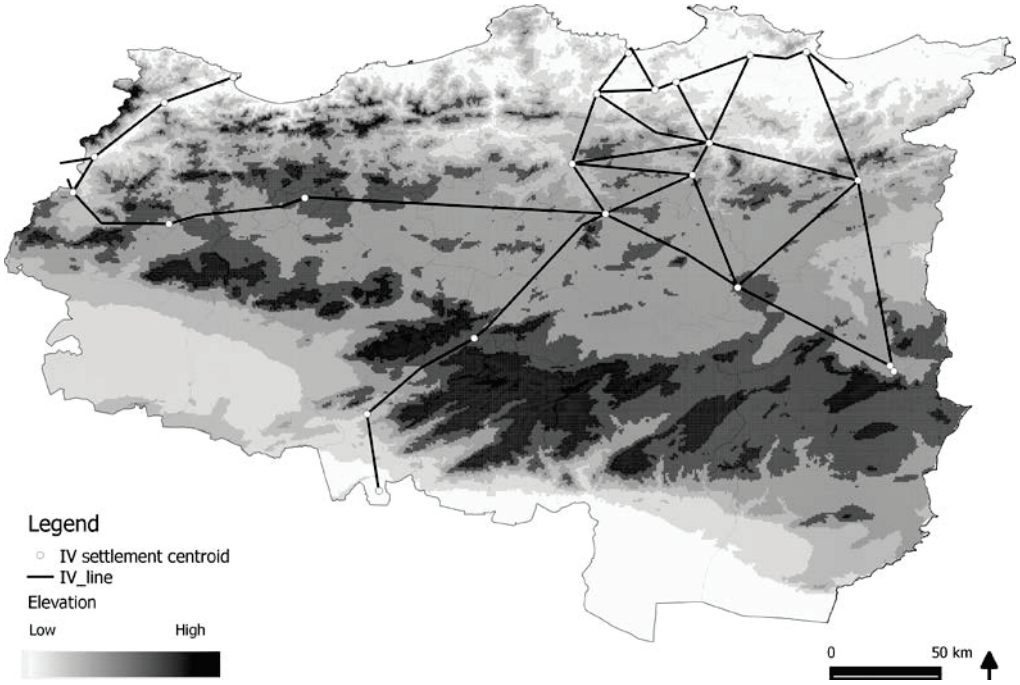
X-axis: control group is labeled 0 and treated group is 1. *Source*: TGdC (1884).

Figure 14: Passenger and freight transport, French Algeria 1863-1894



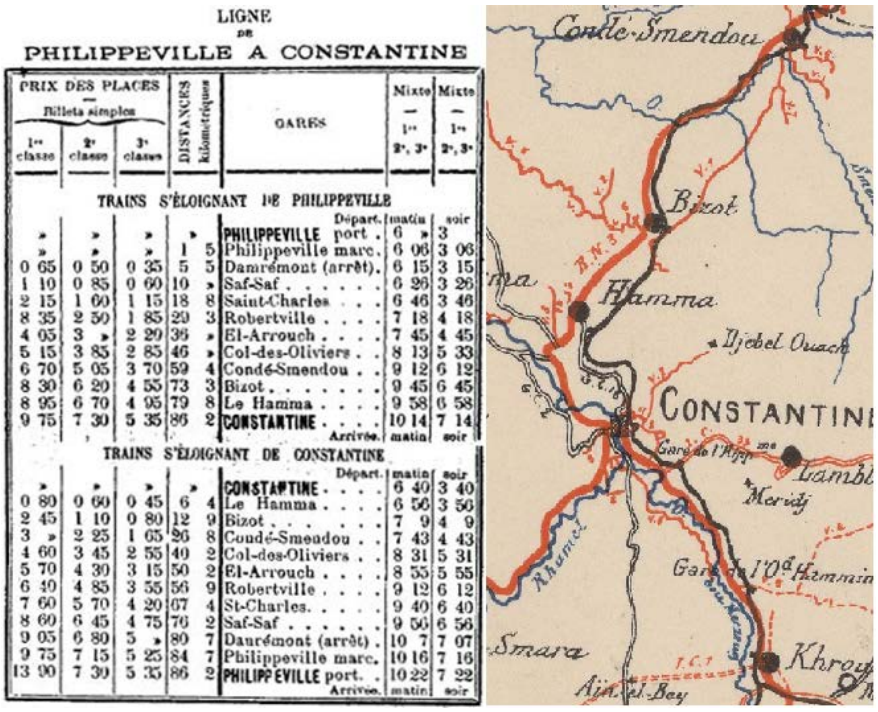
Source: ASA 1905

Figure 15: Map of railway instrumental lines



Source: For detail on construction of the instrumental variable see the methodology section description in paper

Figure 16: Railway prices and map Philippeville-Constantine



LIGNE
DE
PHILIPPEVILLE A CONSTANTINE

PRIX DES PLACES — Billets simples 1re classe	2e classe	3e classe	DISTANCES kilométriques	GARES	Mixte — 1re 2e, 3e	Mixte — 1re 2e, 3e
TRAINS S'ÉLOIGNANT DE PHILIPPEVILLE						
					Départ. matin	soir
»	»	»	»	PHILIPPEVILLE port.	6 »	3
»	»	»	1	5 Philippeville marc.	6 06	3 06
0 65	0 50	0 35	5	5 Damrémont (arrêt).	6 15	3 15
1 10	0 85	0 60	10	» Saf-Saf	6 26	3 26
2 15	1 60	1 15	18	8 Saint-Charles	6 46	3 46
8 35	2 50	1 85	29	3 Robertville	7 18	4 18
4 05	3 »	2 29	36	» El-Arrouch	7 45	4 45
5 15	3 85	2 85	46	» Col-des-Oliviers	8 13	5 33
6 70	5 05	3 70	59	4 Condé-Smendou	9 12	6 12
8 30	6 20	4 55	73	3 Bizot	9 45	6 45
8 95	6 70	4 95	79	8 Le Hamma	9 58	6 58
9 75	7 30	5 35	86	2 CONSTANTINE	10 14	7 14
				Arrivée.	matin	soir
TRAINS S'ÉLOIGNANT DE CONSTANTINE						
					Départ. matin	soir
»	»	»	»	CONSTANTINE	6 40	3 40
0 80	0 60	0 45	6	4 Le Hamma	6 56	3 56
2 45	1 10	0 80	12	9 Bizot	7 9	4 9
3 »	2 25	1 65	26	8 Condé-Smendou	7 43	4 43
4 60	3 45	2 55	40	2 Col-des-Oliviers	8 31	5 31
5 70	4 30	3 15	50	2 El-Arrouch	8 55	5 55
6 40	4 85	3 55	56	9 Robertville	9 12	6 12
7 60	5 70	4 20	67	4 St-Charles	9 40	6 40
8 60	6 45	4 75	76	2 Saf-Saf	9 56	6 56
9 05	6 80	5 »	80	7 Damrémont (arrêt)	10 7	7 07
9 75	7 15	5 25	84	7 Philippeville marc.	10 16	7 16
13 90	7 30	5 35	86	2 PHILIPPEVILLE port.	10 22	7 22
				Arrivée.	matin	soir

Source: Algérie Chemin de Fer Philippeville-Constantine (- 1830-1962 ENCYCLOPEDIE de L'AFN)