Does railway accessibility boost population growth? Evidence from unfinished historical roadways in France

Kakpo, Eliakim and Le Gallo, Julie and Grivault, Camille and Breuillé, Marie

Inra, Agrosup, Cesaer

October 2019

Does railway accessibility boost population growth? Evidence from unfinished historical roadways in France

Marie Breuillé, Camille Grivault, Eliakim Kakpo, and Julie Le Gallo

CESAER, AgroSup Dijon, INRA, Université de Bourgogne Franche-Comté

October 2019

Abstract

The railway revolution that swayed through Europe in the nineteenth century left a legacy of unexplored networks. In this paper, we explore a subset of unfinished railways to evaluate the impact of railroads on population growth. Using the random nature of the achieved portions, we compare municipalities located around the planned but not realized segment of the railways to those in the vicinity of the operated sections. Our results indicate that the railways boost population growth in the medium and long-run. However, the medium-run effects are only visible in municipalities with high pre-arrival population. The railroads also seem to have solved a coordination problem in the sense that treated municipalities were more likely to gain access to other transport infrastructures later.

Keywords: Urbanization, population growth, development, railways, transport
JEL Codes: N44, O18, R10

First draft: (Do not cite)

“Judging the utility of the railways through the lens of the capitalist or speculator, would lead infallibly to excessive and dangerous conclusions. Regarding this aspect, the public opinion has always showed a better sense of wisdom than most narrow-minded economists or short-sighted politicians. This consideration becomes even more obvious once one applies a similar logic to rivers, national, regional and local roads. What would our country be today without these modest means of communication?” Observations on the local and general interest railways. Jean-Baptiste Krantz (1875)

1. Introduction

The association between transportation infrastructure and economic growth is the concern of policy discussions around the world. Railroads could stimulate economic growth due to a reduction in trade costs. Yet, some scholars have defended that the benefits of transportation investments are uncertain (Fishlow 1965, Fogel 1962). The costs of such infrastructures could outweigh the benefits if the reduction in trade costs does not boost up local production enough, implying that the social saving is not sufficient to offset the costs.

The railway revolution that swayed through Europe from the early to late-nineteenth century left a legacy of networks that have shaped the contemporary distribution of economic activity and population in space. As many of these initiatives unfolded in a staggered fashion, accessibility and trade costs was progressively altered near locations at the periphery of the early urban centers. This could have impacted economic incentives and fostered growth. In France, the “Freycinet Plan” which was primarily guided by an administrative motive, took effect towards the later part of the nineteenth century leading to a remarkable expansion of the railroad network. In some cases, the railways were built in areas with limited economic potential especially in rural France. Interestingly, some of the roads did not fully materialize due to a number of hazards, featuring an exogenous variation that we leverage to explore the incidence of railway accessibility on economic expansion.

To do so, we study a set of unfinished railways. Building on the random nature of the achieved portion, we compare municipalities located around the planned but not realized segments of the projected railways to those in the vicinity of the operated sections. Our results indicate that in the medium-run, the infrastructure bears a modest positive effect on population growth in the municipalities with relatively high initial population. In the long-run, we notice that these effects are magnified specifically for municipalities that benefited earlier. These conclusions are also robust to post-treatment cofounding developments such as the road transportation which emerged as a less costly alternative to railroads in the interwar period.

The rich empirical literature on the legacy of the railroad revolution in the second half of the nineteenth century explored numerous aspects of the economy. The increased accessibility to areas in the periphery of early agglomerations has affected urbanization (Bogart et al. 2018, Erik Hornung 2012), population growth (Da Silveira et al. 2011, Thevenin et al. 2016), employment (Bogart et al. 2017), economic growth (Pereira et al. 2014, Haines and Margo 2006), firm location (Atack et al. 2008), agricultural improvement (Atack and Margo 2011), structural change (Berger and Enflo 2013), interregional trade (Donaldson 2018), school enrollment (Atack et al. 2012) or the emergence of entrepot cities (Montag and Xiong 2016) to name a few. Yet, the literature has devoted little attention to the incidence railroad accessibility bears on urbanization and economic outcomes long after the exploitation of the infrastructures. Except a few papers (Berger and Enflo 2013, Jedwab and Moradi 2011), most empirical works focus on short or medium-run effects of historical railways.

The long-run impacts of transportation infrastructure could vastly supersede the medium-run effects. This could be achieved in multiple ways. Proximity to rail transportation could spur the immediate emergence of agglomerations due to access to a larger market and trade. These economic clusters once established tend to thrive rapidly and last longer (Cronon 1998). Alternatively, even with modest medium-run effects, the railway networks could become a reference point for other public infrastructures (highways, landlines, interchange, etc.) that influence long-run productivity and economic outcomes.

This hypothesis lines up with a path-dependence argument which suggests that the trajectory of economic growth is determined by historical shocks. Economists have debated for a long time the importance of path-dependence in the spatial allocation of population and economic activity. Several contributions of the likes of Davis and Weinstein (2008), Bosker et al. (2007) and Miguel and Roland (2011) use the bombing of Japan, Germany and Vietnam during wartime, to explore whether these infrastructure shocks had long-term effects on the localities that were targeted. If these cities recover their population and industries in the long-run, then place specific features are probably fundamental to the spatial equilibrium of economic activity and local infrastructure investments (or destructions), such as railway constructions, might not be much relevant in the long-run. In contrast, other papers have defended that localized shocks can have permanent effects (Boskeret al., 2007, 2008; Bleakley and Lin, 2012; Dell, 2012).

Besides, there is a debate as to whether infrastructures that were built "ahead of traffic" (Fishlow 1965, Fogel 1962, Rostow 1960) would spill over other aspects of the local economy and ignite investments that spur prosperity in the medium-run. Fishlow (1965) argued that the railroads built in the American Ante-bellum in the second half of the nineteenth century were not built ahead of demand. He suggested that the economic success of the Midwestern railroads, owe to the fact that the railways crossed through densely populated areas, which intensified commercial agriculture in the affected localities. In the spirit of Fishlow's account, these railroads were successful in part because they ran through areas where the introduction of the railways lowered transportation costs to the extent that the social saving was substantial.

In this case, the railroads accelerated economic growth and urbanization because they opened up vast arrays of market opportunities for localities with a high growth potential. To reach this conclusion, he calculated the social returns of the railways by comparing the costs of transportation with the railroads to the second best alternative mode of transportation around at the time. All else equal, in densily populated areas, the social returns should get higher, because the new infrastructure reallocates resources in the productive sector leading up to an expansion of the local economy.

Many of the French railways under study were built for administrative and political motives and likely did not follow a local demand. Though, the railways could lower transportation costs in some areas, there is substantial evidence supporting the hypothesis that were not directed to places with a high economic potential. These railroads were not the initiative of private capital interests, but seem to have been imposed by the State to railroad construction firms in exchange for long-term concession grants of exploitation. The economic opportunity of the local railways was eloquently discussed by Jean-Baptiste Krantz (1875), engineer and member of the French parliament, who also happens to be a strong advocate of the expansion of rail transportation to the countryside.

> *"One could rightfully question what we could expect, as returns on investment, from the local railways (our railways belong to this network), when the general interest railways (many of which would become integral to the primary network) significantly more important, had only yielded, such modest returns, after years of exploitation" Observations on the local and general interest railways.* Jean-Baptiste Krantz (1875)

Expansion of the railroad network in France was at its height in the second half of the nineteenth century. The Freycinet Plan adopted in 1878 sets the goal of connecting by trains every "sousprefecture" i.e. subdivision of provinces. The original draft planned to build 8,700 km of railways by raising private capital and providing public grants. The goal was to connect the central government to the people especially in the countryside. However, several of the planned railways never fully materialized due to different hazards. Following the first world war and the Great depression, the government decided to nationalize the railroad network and conceded their exploitation to a new structure, the National Company of Railways (SNCF), which did not find many of the lines cost-effective.

In many cases, the railways were closed in the interwar period even though parts of the lines never opened. In some cases, significant chunks of the railway infrastructures have already been laid down. In general, the segments of the planned railways that were finalized were exploited while the remaining portions never opened. This represents a natural experiment that can be leveraged to investigate the impact of railroads on urbanization and population growth in rural precincts over time. As accessibility to the network reduced trade costs, it should open up new arrays of market opportunities to firms and could spur an inflow of population.

On the empirical front, identification heavily relies on comparing localities affected by the infrastructure to those that were not or a counterfactual of regions that could have been treated based on counterfactual trajectories of the routes or the construction technology available at the time. These strategies are not fully satisfying because of the non-random drawing of the trajectories and the omission of intangibles in the hypothetical linear routes between two points. The approach adopted in this paper consists of comparing municipalities treated in the sense that they were in the vicinity of constructed railways to those in the proximity of the planned but never achieved portions.

First, we observed that the infrastructure was beneficial to some municipalities in the medium-run. These are the municipalities with relatively high initial population. Second, we noted that these medium-run effects are magnified in the long-run. Similar to Berger and Enflo (2013), we hypothesize that the railways must have solved a coordination problem. Municipalities that gained access to the railway networks seem to have become reference points for other major transportation infrastructures (interchange, highways, etc.) later. To validate this story, we showed that municipalities near the railways were closer to highway interchanges and more likely to be on the path of a highway.

Plus, our findings indicate that the treatment effect depends on the distance of a municipality to the railway. The results discussed in this paper are robust to other cofounding factors at the regional level as well as other developments during and post-treatment such as the automobile revolution and other relevant railways. The remainder of the paper is organized as follows. In section two, we describe the historical policy background of the Freycinet Plan and the set of unfinished railways studied. Section three presents the conceptual framework. Section four exposes the identification strategy and empirical approach. Section five describes the data and variables used in the regressions. In section six, we discuss the results. Section seven provides some concluding remarks and explores avenues for future research on the topic.

2. Historical policy context: The Freycinet Plan

The railroad revolution that gained traction in Europe during the early twentieth century was not visible in France until the second half of the century. The first railways in France opened between Saint Etienne and Andrézieux, south west of Lyon, in 1828. Like early English railways it was initially seen as an adjunct to inland waterways. The first line to be built specifically for passenger traffic, which happened to be the first to serve the capital Paris, only opened in 1837 much later than its British pioneer which opened in 1830. However, railroad transportation remained very limited in France compared to Great Britain.

By 1840, France recorded around 469 km of railways, five times less than Great Britain which had already built up to 2,390 km. Most of the railway constructions in France at the time were undertaken by private local entrepreneurs seeking to promote their own business interests. The role of the State was limited at the dawn of the nineteenth century to the extent that the country had been weakened by the Napoleonic wars and simply represented a collection of regions competing for limited local interests.

This picture will soon be redefined by the "Freycinet Plan" which took effect in the last quarter of the century and set to connect the countryside to regional urban centers and Paris. Though the project included the construction of roads, ports, canals, bridges and the planning of inland waterways, railways remain by far the largest component of the initiative (about one third of the total costs). The railroads of the Freycinet Plan can be divided into a primary network, that would interrelate local agglomerations to Paris and a secondary network that would connect rural areas to regional centers. Apart from the economic benefits of the railroad expansion, the initiative ambitioned to connect all social forces of the French society.

The initial draft, anchored around the primary network, planned to build around 8,700 km of railways across the country. However, this figure will be enlarged to 19,000 km (including the secondary networks) due to pressure from provinces. The effects of the policy on the density of railways in France was immediate. By 1900, France had overtaken its pioneer neighbor with an estimated network of railroads of 38,109 km compared to 30,079 km in Great Britain. Table 1 reports the evolution of the railway networks in a selected set of European countries during the second half of the nineteenth century.

Table 1: Evolution of railway networks in Europe (km)

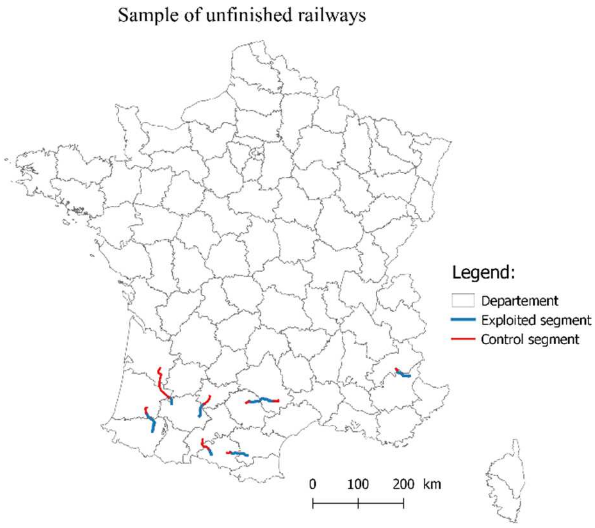
	1840	1860	1880	1900
Austria-Hungary	144	4,543	18,507	36,330
France	496	9,167	23,089	38,109
Germany	469	11,089	33,838	51,678
Great Britain	2,390	14,603	25,060	30,079
Netherlands	17	335	1,846	2,776
Russia	27	1,626	22,865	53,234
Sweden	-	527	5,876	11,303

Source: Railroads in 19th Century Europe, Korean Minjok Leadership Academy

Notwithstanding its early success, the Freycinet Plan did not fully materialize because of several hazards. First, as discussed earlier, the project was driven in part by an administrative and political motives. The Plan was amended due to vigorous debates in the French Parliament with each representative seeking to bring the network to its constituents. As a result, many of the secondary network lines specially in the countryside were not economically sustainable. Second, the railways were built and operated by private firms that were granted subsidized loans and the right of exploitation. The expectation was that the burden of the construction costs will be outweighed by the revenue stream that the lines will generate in the long-run. Unfortunately for railway companies, the anticipated success did not materialize everywhere resulting in operating losses, which were initially paid for by public subsidies.

However, with the Great Depression and a thriving automobile transportation sector, many of the companies went bankrupt and the State was forced to nationalize the industry. This led to a closure of several lines, some of which were not completely achieved. The railway lines that we exploit in this paper belong to the second category. Most of these lines were closed and declassified following the nationalization and restructuring of the French railroads in the interwar period. These routes, included in the subsidiary secondary network of the Freycinet Plan, were predominantly rural and not economically sound. Figure 1 below describes a map of the railways.

Figure 1: Sample of unfinished railways studied



Source: See text

Interestingly, only portions of the referenced lines had been operational. The remaining sections which were expected to open later, were never exploited. This presents an interesting natural experiment that can be used to study the impacts these historical infrastructures have on economic expansion and population growth both in the medium-run and long after their exploitation. The random nature of the treatment in this context holds to factors related to the rising debt of railway companies and a growing competition with automobile transportation. These developments accelerated the nationalization of the railway companies and the declassification of the routes under study, all of which resulted in a spatial discontinuity of treatment across localities.

Overall, we surveyed seven railways that are all part of the Freycinet Plan and were never completely achieved. All but one of the roads were declassified in the interwar period with the nationalization of the French railway network. Declassification of a line implies a closure of all operating stations along the route. Table 2 describes the railways and the number of treated and control municipalities. The first three columns describe the denomination of the railway and the years when the railway line opened and closed. The next two columns describe the length of the exploited portion of the planned road and the duration over which the operated portion remained active. The last two columns of the table summarize the number of municipalities located in a ten-kilometer vicinity of (i) the operated section and (ii) the planned but never realized segment of the surveyed railways.

Table 2: Description of the sample of unfinished railways

(1) Railway denomination	(2) Opened	(3) Closed	(4) Length(*) (km)	(5) Duration[(a)]	(6) Treated[(b)]	(7) Control[(c)]
Gabarret-Eauze	1866	1938	21.0	72	77	19
Beaumont-Gimont	1904	1937	32.8	33	30	95
Hagetmau-Pau	1910	1938	14.9	28	30	129
Saint-Girons-Oust	1866	1969	19.3	103	87	20
Saint-Juéry-Saint-Affrique	1896	1939	68.0	43	36	61
Chorges-Barcelonette	1905	1935	42.0	30	23	9
Foix-Quillan	1884	1938	29.7	54	31	71
Total	-	-	227.9	-	300	418

(a) Duration of exploitation of the built section (in years) (b) Number of municipalities within 10 kilometers of the built portion
(c) Number of municipalities within 10 kilometers of the planned but not-built portion. (*) Length of the constructed section

3. Conceptual framework: economic impacts of railway accessibility

The empirical literature on railroad accessibility and economic growth builds on the conceptual contribution of Fogel (1965). The model considers that the benefit on the local economy is determined by the "social saving" of the transportation infrastructure. The author relates the social saving to the difference between the actual cost of shipping goods in that year and the alternative cost of shipping the same bundle of goods between exactly the same points without the railroads. A more formalized two-sector version of this model was described by Fogel in another paper in 1979. Below we present a similar dual version of the model.

In the spirit of Fogel (1979), consider a basic two-sector economy that produces two goods: (i) a transportation service (T) and (ii) an aggregate of all other outputs (A). Both goods are produced using capital (K) and labor (L) with respective prices *w* and *r*. Additionally the aggregate sector A uses transportation as an intermediate resource. Assume a zero elasticity of demand for transportation and that the transportation service can be produced using two different technologies F_h (advanced) and F_l (backward). Let the unit cost functions for the aggregate output sector (A) and both transportation technologies be defined as following:

$$c_A = c_A(w, r, c_T) \tag{1}$$

$$c_{T_h} = c_{T_h}(w, r) \tag{2}$$

$$c_{T_l} = c_{T_l}(w, r) \tag{3}$$

In addition to the traditional factor prices *r* and *w*, the unit cost of producing the aggregate good (c_A) includes the unit cost of the transportation service c_T. Likewise, c_{Th} and c_{Tl} represent the unit cost functions for the advanced and backward technologies respectively. As the more efficient railroad transportation service is introduced, companies save income by moving from the backward technology (water inland transportation, road transportation, etc.) to a more efficient mode of transportation. Under the assumption that the advanced transportation technology (T_h) is cheaper, it follows that the cost of producing a unit of the transportation service is lower with the more advanced railway transportation implying that, for any given combination of *w* and *r*:

$$c_{T_h}(w, r) < c_{T_l}(w, r) \tag{4}$$

Let Q_A, T_A, T_{Ah}, and T denote respectively the production of the aggregate output (A), the transportation input used in the aggregate sector, the net total transportation output and the total transportation output (including the intermediate part). The following identity relates all transportation outputs in the model.

$$\mathrm{T}=\mathrm{T}_{\mathrm{A}}+\mathrm{T}_{\mathrm{A_h}} \tag{5}$$

Similar to Fogel (1979), and for simplicity, we make the additional hypothesis that the introduction of the railroad infrastructure does not affect the demand for transportation service (inelastic demand of transportation services). As a result, the social saving of the railroad can be decomposed into two parts. First, as the unit cost of transportation is reduced, the unit cost for the production of the aggregate output will be reduced as well. This will boost local income assuming that all prices and levels of output remain unchanged. The saving income per unit of aggregate output can be expressed as following:

$$\Delta\mathrm{c}_{\mathrm{A}}=\mathrm{c}_{\mathrm{A}}(w,r,c_{T_l})-\mathrm{c}_{\mathrm{A}}(w,r,c_{T_h})\simeq\frac{\partial\mathrm{c}_{\mathrm{A}}}{\partial\mathrm{c}_T}(c_{T_l}-c_{T_h}) \tag{6}$$

For the transportation industry, the income saved equals the difference of the unit cost of transportation times the net transportation output:

$$\Delta\mathrm{S}_{\mathrm{A_h}}=(\mathrm{T}-\mathrm{T}_A)(c_{T_l}-c_{T_h})=\mathrm{T}_{A_h}(c_{T_l}-c_{T_h}) \tag{7}$$

The social saving of the upgraded transportation infrastructure equals the sum of the saving for the transportation sector and the all other aggregate output sector:

$$\Delta S=\Delta\mathrm{S}_{\mathrm{A_h}}+\mathrm{Q}_{\mathrm{A}}\Delta\mathrm{c}_{\mathrm{A}}\simeq\mathrm{T}_{\mathrm{A_h}}(c_{T_l}-c_{T_h})+\mathrm{Q}_{\mathrm{A}}\frac{\partial\mathrm{c}_{\mathrm{A}}}{\partial\mathrm{c}_T}(c_{T_l}-c_{T_h})=\Delta c_T\left[\mathrm{T}_{\mathrm{A_h}}+\mathrm{Q}_{\mathrm{A}}\frac{\partial\mathrm{c}_{\mathrm{A}}}{\partial\mathrm{c}_T}\right] \tag{8}$$

A more convenient formulation that highlights the main components of the social saving of the railway is expressed as following:

$$\Delta S\simeq\Delta c_T\left[\mathrm{T}_{\mathrm{A_h}}+\mathrm{Q}_{\mathrm{A}}\frac{\partial\mathrm{c}_{\mathrm{A}}}{\partial\mathrm{c}_T}\right]=\mathrm{T}_{\mathrm{A_h}}\Delta c_T\left[1+\frac{\mathrm{Q}_{\mathrm{A}}}{\mathrm{T}_{\mathrm{A_h}}}\frac{\partial\mathrm{c}_{\mathrm{A}}}{\partial\mathrm{c}_T}\right]=\Delta\mathrm{S}_{\mathrm{A_h}}\left[1+\frac{\mathrm{Q}_{\mathrm{A}}}{\mathrm{T}_{\mathrm{A_h}}}\frac{\partial\mathrm{c}_{\mathrm{A}}}{\partial\mathrm{c}_T}\right] \tag{9}$$

Equation (9) indicates that the social saving of the upgraded railway mode of transportation depends on (i) the size of the productive domestic sector, (ii) the cost saved for the transportation sector, and (iii) the responsiveness of a unit production cost with respect to a unit transportation cost in the productive sector. Specification (9) highlights that the size of the domestic economy positively affects the income saved with the arrival of the railways. Relatedly, the structure of the domestic productive sector will influence the social saving of the new mode of transportation. If the local economy is dominated by sectors that heavily rely on transportation as an input, then the social saving would be greater, and economic growth will likely follow the arrival of the railways.

However, the simplistic derivation described in equation (9) represents a lower bound of the true social saving of the infrastructure. The model abstracts from a demand elasticity for transportation services, and does not consider the reallocation of resources towards the sectors for which transportation represents a relatively more valuable input. Plus, the railroad would make local firms more productive by improving access to a wider variety of inputs and would enable entrepreneurs to cater to a larger market of consumers. The model also predicts that the social saving depends on how the cost of railway transportation compares to the cost of alternative modes of transportation in the region.

If the cost differential Δc_T and the aggregate output Q_A are non-negligible, local income will expand and the locality will attract new residents. This development should manifest itself through an acceleration of population growth. In the absence of a reliable data on municipal economic activity during this time period, many scholars have used population growth as a proxy for economic expansion. This paper follows the tradition and similar to Bogart et al. (2017), we capture accessibility to the railroad through the distance of the centroid of a municipality to the infrastructure.

4. Methodology: identification strategy and empirical approach

Measuring the treatment effects in this context presents some empirical challenges. First, we identify 718 municipalities within a ten-kilometer distance of the seven railways surveyed in for this paper. The sample is relatively balanced in the sense 300 of the municipalities are treated i.e. located within ten kilometers of the exploited segments of the railways, and the remaining 418 are considered as controls. Using historical data on the French population census, we observe these municipalities over the period 1850-2015 and adopt a lead-lag structure in the regression specifications. This method allows us to control for pre-treatment differences between the two groups.

Table 3 below describes the average growth in population between the two groups of municipalities before, during and after the exploitation of the active segment for each of the railways. The naive eyeball test seems to indicate that the difference in average population growth between the two groups increased during and after the exploitation of the active portion of the railway.

Table 3: Difference in population growth between treated and control groups

Period of comparaison	*before*	*during*	*after*
Railway denomination			
Gabarret-Eauze	0.0025	0.0015	0.0050
Beaumont-Gimont	0.0021	0.0014	0.0061
Hagetmau-Pau	0.0018	-0.0007	-0.0022
Saint-Girons-Oust	-0.0016	0.0058	0.0049
Saint-Juéry-Saint-Affrique	-0.0018	0.0011	0.0080
Chorges-Barcelonette	0.0015	0.0050	0.0050
Foix-Quillan	0.0015	-0.0016	0.0056
Total(*)	0.0060	0.0125	0.0324

(*) Population growth measured in decimal units (0.01 corresponds to one per cent)

The baseline estimated equation is the following:

$$\Delta\log(pop_{c,l,t})=\beta_0+\alpha_{l,t}+\beta_1 treat_{c,l}+\sum_{j=-2}^{5}\delta_j d_{t+20j}+\sum_{j=-2}^{5}\alpha_j treat_{c,l} d_{t+20j}+\eta X_{c,l,t}+u_{c,l,t} \quad (1)$$

The dependent variable $\Delta\log(pop_{c,l,t})$ measures the growth of population between two consecutive census years t-1 and t for municipality *c* located within a ten-kilomoter distance of the railway *l*. We choose to explain the growth instead of the level of population because the former does not usually feature a time trend implying that the difference in difference approach nicely lends itself to this context as it basically captures jumps in the outcome of interest. The inclusion of the railway-year effect $\alpha_{l,t}$ allows us to only compare treated and control municipalities that belong to the same railway line over the census year *t*. This term is the critical component of our regression approach.

The treatment dummy $treat_{c,l}$ equals one if municipality *c* belongs to the treatment side of the railway *l*. The variable d_t represents a time dummy for period *t* and has been broken down in twenty-year intervals. Given the quinquennial time-step of the dataset, using twenty-year windows around the opening of the lines enables us to use three to four census years on average to measure the difference in difference. Some specifications also include region-census year effects to capture time-varying local heterogeneity. The matrix $X_{c,l,t}$ includes several additional regressors such as the log of the land area.

In some specifications, we also control for the distance to the closest highway interchange, the distance to the closest train station, the distance to the closest major railway, the distance to the closest water stream and the log of the municipal population in 1850. The last two features are time-invariant. Notice that the presence of the time invariant variables controls for municipality specific fixed effects that are inherent to external market access factors or unobserved productivity and amenities. The idiosyncratic error term $u_{c,l,t}$ is clustered at the department year level to capture local shocks that affect urbanization patterns at the municipal level.

$$\Delta\log(pop_{c,l,t}) = \beta_0 + \alpha_{l,t} + \beta_1 treat_{c,l} + \sum_{j=-2}^{5}\delta_j d_{t+20j} + \sum_{j=-2}^{5}\alpha_j treat_{c,l} d_{t+20j} + \sum_{s=s_1}^{s_2}\sum_{dep}\theta_s d_s 1_{\{dep=1\}} + \eta X_{c,l,t} + u_{c,l,t} \quad (2)$$

Two other specifications are estimated. Equation (2) improves upon the baseline specification by including department-census year effects. This allows us to control for time-variant heterogeneity at the department level. Departments represent the second echelon of the French administrative architecture from the top. Various local policies are set at this level implying a spatial interdependence of policy shocks across municipalities in the same department.

$$\Delta\log(pop_{c,l,t}) = \beta_0 + \alpha_{l,t} + \beta_1 treat_{c,l} + \sum_{j=-2}^{5}\delta_j d_{t+20j} + \sum_{j=-2}^{5}\alpha_j treat_{c,l} d_{t+20j} + \sum_{j=-2}^{5}\mu_j treat_{c,l} d_{t+20j}.dist + \eta X_{c,l,t} + u_{c,l,t} \quad (3)$$

The last regression specification described in equation (3) explores the heterogeneity of treatment effects as a function of the intensity of treatment. Exposure to treatment which we measure through the distance to the exploited portion of the railway is interacted with the treatment dummies. It stands to reason that the degree of accessibility to the active segment of a railway varies with the distance to the line. The interaction term, if negative would reflect this consideration.

5. Samples, data and sources

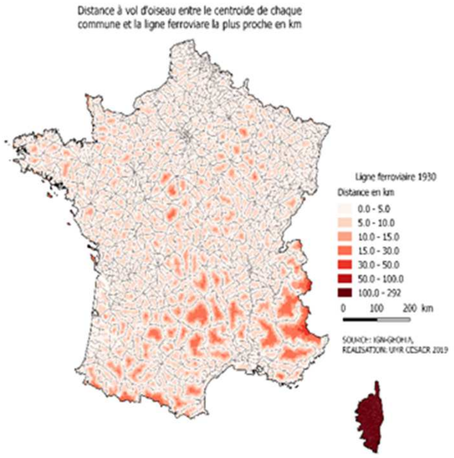
5.1 Sample of municipalities

Our identification approach consists of comparing treated and control spatial units around the exploitation of the surveyed railways. This approach rests critically on the assumption that the two sets of geographies are clearly defined. The treatment in this scenario is not easy to define. Following the empirical tradition, we consider the municipality as the spatial unit of interest and measure treatment through the "crow flies" distance from the centroid of a municipality to the exploited segment of a railway. This continuous notion of treatment does not lend itself nicely to a difference in difference identification strategy.

Therefore, we further refine the definition of treatment as a dummy variable which equals one if the municipality is located within ten kilometers of the exploited portion of a railway. Choosing ten kilometers as the vicinity for treatment status is not completely random and is the result of the distribution of distances from a municipality's centroid to its closest railway during the early nineteenth century in the southwestern region of France where most of our railways are located. At the peak of the Freycinet Plan, the average distance from a municipality to the closest train station was about seven kilometers in the region under study. The ten-kilometer window yields an average distance to the railway that matches the observed average in the region in the early decades of the twentieth century. Figure 2 below presents the map of the average distance from the centroid of a municipality to the closest railroad in 1930.

However, this choice is inconsequential for the qualitative implications of the results discussed in this paper as we consider alternative definitions for treatment based on a fifteen-kilometer, and a twenty-kilometer distance windows in the robustness checks. Alternatively, we interacted the treatment effects with the intensity of treatment measured through the distance to the exploited segments in some regressions.

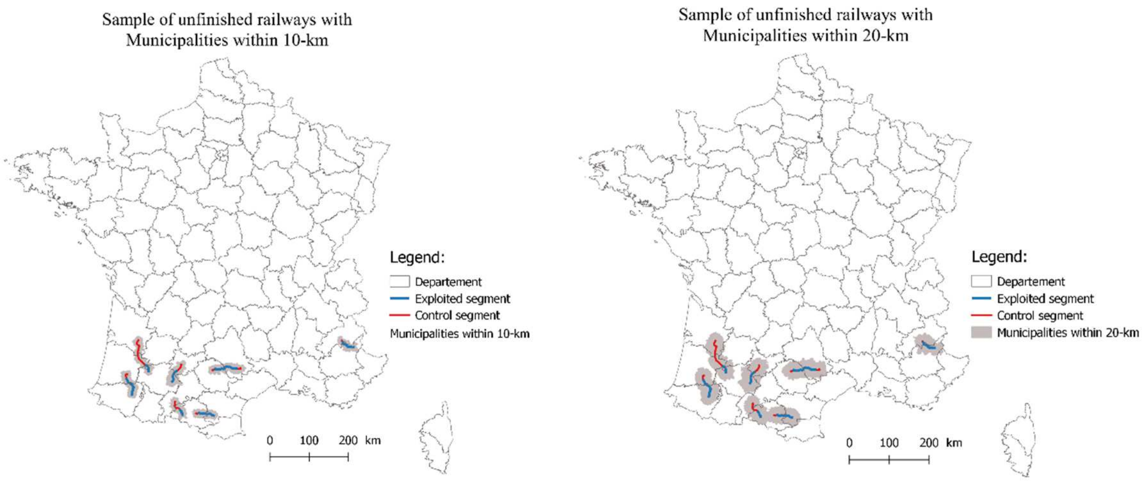
Figure 2: Distance of a municipality to the closest railway in 1930



One critical limitation that we could not really address relates to the fact that a few municipal boundaries have changed over time due to the fusion of administrative units. The municipality level is appropriate in this context, because we operate under the assumption that the aforementioned railways were rural and could only if anything have a local impact. Plus, the municipality is the lowest administrative division over which population data is available going back to the mid-nineteenth century. Unfortunately, the information necessary for the identification of boundary redefinitions is not available for the French municipalities. It also seems that the limits of French municipalities feature a remarkable historical stability over time (Thevenin et al. 2016).

In sum, the sample of analysis includes 1,634 municipalities located within a twenty-kilometer vicinity of the projected railways. The preferred specifications restrict the sample to 718 municipalities located within ten kilometers of the railways although we consider the intermediate sample of 1,165 municipalities in a fifteen-kilometer vicinity in the robustness checks. All municipalities are observed during census years (around every five years) over the period 1851-2015.

Figure 3: Map of the municipalities in our sample



5.2 Geo-localization of transport infrastructures

We use data on a set of railways that were planned but never achieved although portions were exploited. Data on these railways which are available online, have been put together from two sources. The website www.chemins.de.traverses.free.fr surveys historical archives on a number of rural railways thanks to contributions from volunteers and railroad historians. Plus, the website http://archeoferroviaire.free.fr provides valuable geographic information on many of the historical railways that have been declassified. The information is available as digitized image maps for the majority of the lines.

Thereafter, we geolocalized the maps using a GIS software. The dataset that comes out of this process describes, the total length, the opening and closing dates and the geographic location of each line. Next, we calculate the distance from every municipality's centroid to the railways and identify for each of the originally planned railways, the municipalities located in a ten-kilometer vicinity (based off a crow flies distance) of the line.

Information with regards to the location of highway interchanges is produced by the French Ministry of Ecological and solidary Transition and is available online via the website https://www.data.gouv.fr/fr/datasets/echangeurs-du-reseau-routier-national-concede/#_. The interchanges were localized relative to the highway transportation networks using the GIS dataset of French highways available https://www.data.gouv.fr/fr/datasets/liaisons-du-reseau-routier-national/. Finally, data on the historical geolocalization of the railways and all train stations which are available upon request was provided by Chritophe Mimeur of the Université Bourgogne Franche-Comté in (France), while the GIS shapefile of French waterways could be found on the National Institute of Geographic and Forestry Information (IGN) website: http://professionnels.ign.fr/donnees.

5.3 Historical population data

The historical data on the municipal population comes from a couple of sources. For the years prior to 1968, the data is produced by the French Laboratory of Historical Demography (LADEHIS) and are available online (http://cassini.ehess.fr/cassini/fr/html/index.htm). The estimates have been put together due to an extensive digitalization of successive French population censuses going back to 1793, year corresponding to the first exhaustive population census in France. After a second population census in 1800, all subsequent censuses took place on a five-year time step (years ending in 01 and 06) until 1946, except during wartime.

Thus, the population census of 1871 took place in 1872, the one that should have happened in 1916 was never done and the one in 1941 was never published. Since 1941, the interval between two consecutive censuses expanded over time resulting in censuses in 1954, 1962, 1968, 1975, 1982, 1990 and 1999. From 1968, the French Institute of Statistics, which is charged with the task of running census operations in the country, exhaustively counted individuals on the ground every five years. This approach lasted up until the census of 1999, and will only be amended from the following census.

From 2006, a new process established a distinction between small municipalities (less than 10,000 inhabitants) for which there is an exhaustive counting every five years and large municipalities (more than 10,000 inhabitants) where the Institute uses annual surveys and an adjustment factor to estimate the resident population. The time-step of our analysis follows the evolution of population censuses in France which occurs approximately every five-years. In addition, we include opening years of railways in order to measure changes relative to opening. If a railway opens during a census year, we measure the growth of population in the opening year as the average annual growth rate of population between the last two consecutive censuses. This is the case for the years 1866 and 1896 which correspond to census years. For other opening years, we estimate the average annual population growth by linear interpolation between the two censuses around the relevant period. Formally, the population of a municipality c in the year t in between two censuses is estimated as follows:

$$\mathrm{Pop}_{c,t} = \mathrm{Pop}_{c,1} + (\mathrm{Pop}_{c,2} - \mathrm{Pop}_{c,1}) \times \frac{\mathrm{Year}_t\text{-}\mathrm{Year}_1}{\mathrm{Year}_2\text{-}\mathrm{Year}_1}$$

$\mathrm{Pop}_{c,2}$ measures the population in municipality c during the census year 2 (upper end census), $\mathrm{Pop}_{c,1}$ refers to the same population for the census year 1 (lower end census) and Year_t refers to the year over which the municipal population is estimated.

6. Results and discussion

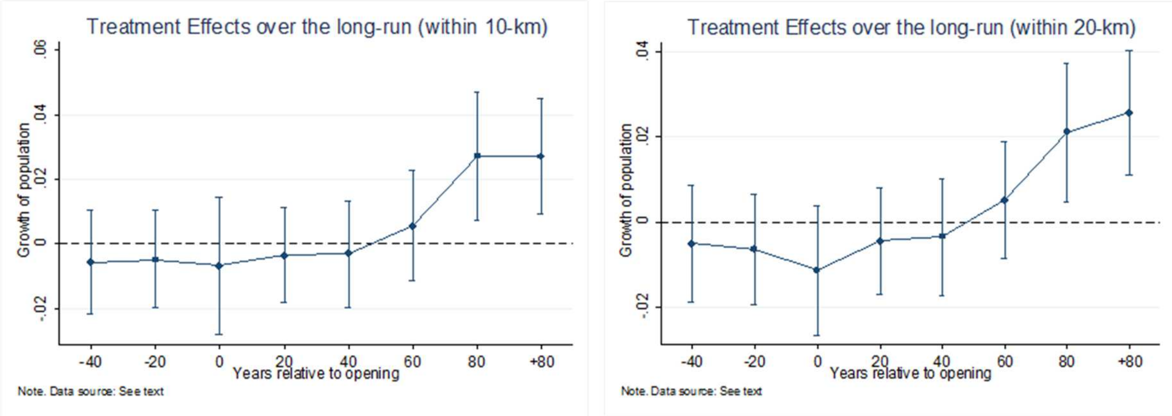
6.1 Evolution of the treatment effects over time

Table 4 in the appendix reports the results for our preferred difference in difference specifications using the restricted sample of municipalities located within ten kilometers of the railways. Column (1) which only includes railway-census year effects provides the baseline results. In column (2) we included railway fixed effects to capture attributes specific to the trajectory of a railway that could potentially affect local population growth and urbanization potentials (length, localities covered, etc.). Column (3) captures time-variant regional heterogeneity with the inclusion of department-census year effects. Last, column (4) includes all municipality specific fixed controls such as the log of the land area and the distance of the municipality to a major navigable water stream.

For the full sample, we found a strong positive effect of accessibility to the railroad on quinquennial population growth but only long after the exploitation of the railways. The lead-lad regression structure allows us to compare population growth rates between the control and treated geographies prior to the opening of the operated sections. Twenty and forty years prior to the introduction of the railways and relative to the excluded previous years, we notice a negative but insignificant difference between the two groups. This negative difference persists over time but remains insignificant up until sixty years after the beginning of the exploitation of the active portions of the railways. However, a strong and significant positive effect is only observed between sixty and eighty years after the lines opened.

Specifically, we notice an average quinquennial growth differential of 3.9 percentage points between the municipalities that gained access to the railroads and the controls which did not in the long-run. This effect shrinks significantly to 2.9 percentage points once regional time-variant heterogeneity is accounted for through department-census year effects (see figure 4). The positive treatment effect in the long-run remains robust to the definition of treatment adopted. When accessibility is defined based on a window of fifteen or twenty kilometers (table 5 and table 6) from the railways, we estimate population growth differences of 2.7 and 2.6 percentage points respectively between treated and control municipalities. Besides, there appears to be a negative albeit very imprecise effect of the railway on population growth during the year of opening. This points to the hypothesis that workers involved in the construction of such massive infrastructures move away afterwards.

Figure 4: Estimated treatment effects of railway accessibility over the long-run

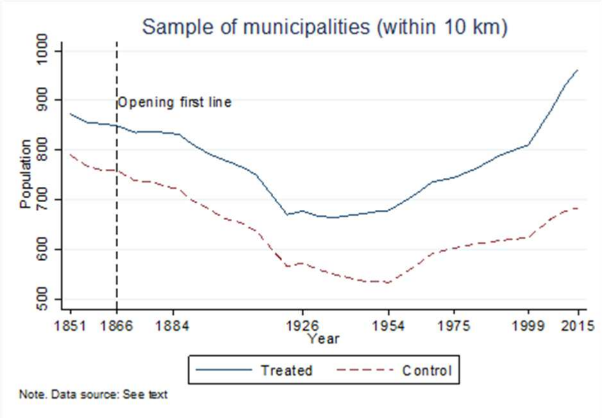


Plus, the location specific controls are significant with associated signs broadly consistent with expectations. A one percent higher population in 1850 entails a 0.5 percentage point increase in average annual population growth over the period of analysis, suggestive of increasing returns and agglomeration effects. Similarly, proximity to a major navigable water stream correlates with a higher average annual population growth. In contrast, and quite surprisingly the land area bears a negative association with population growth. This result could only be interpreted in conjunction with the estimate of the parameter of the log of initial population. If anything, the control of interest should be the density of population. Higher levels of density could spur more urbanization and population growth depending on the interaction between agglomerations benefits and congestion costs.

The finding that proximity to the operated sections of the railways did not have a positive effect on population in the medium-run might seem puzzling. But it simply confirms the presumption that several of the Freycinet railways, in particular those in the secondary network, were built in areas with limited potential for urbanization and growth. Plus, the periods of exploitation of all the railways cover the two world wars and the Great Depression, all of which were marked by a staggering economic stagnation and a loss of population in many parts of the country. Below, we plot on figure 5 below the series of population trends between treated and control municipalities over the period 1851-2015. Prior to the introduction of the railways, the average treated and control municipality was already on a downward path. This observation applies to most of the railways.

As evidenced by the graphs on figure 6 in the appendix, most of the studied railways ran through localities where population was on a decline before the inception of the transportation infrastructure. This should have been expected given that a significant portion of the Freycinet railroads were undertaken under political and administrative influences. In many cases, the lines were brought to the countryside following intense pressure from representatives in the French parliament. These findings combined with the graphical illustrations on figure 7 and figure 8 in the appendix reinforce the argument made by other scholars (Fishlow 1965) and Fogel (1962) notably that historical railways built ahead of demand do not necessarily translate into urbanization and economic development in the short or medium-run.

Figure 5: Population trends between treatment and control municipalities



This result becomes even more interesting once we factor in the fact that all the lines but two were active and operational for less than sixty years after they opened. Hence, the measured treatment effects are not necessarily associated with accessibility to railroad transportation and could be linked to other developments that correlate with both the growth of population in a municipality and treatment status (proximity to the active segment of the railroads). To explore this possibility, we consider the importance of motor transportation which emerged as a less costly alternative to railway transportation in the 1930s. Though, this mode of transportation appeared in the interwar period, it grew rapidly in the aftermath of the second world war along with the construction of highways.

Using a GIS software, we localize the introduction of major highway interchanges in France over time. If treated municipalities happen to be located near highway interchanges, therefore the observed positive population boost in the long-run could simply be reflecting the fact that accessibility to road transportation which emerged much later, correlates with the positioning of historical railways. The results in table 7 in the appendix consider the importance of accessibility to a highway. To do so, we include the interaction between two variables in the previous regressions. First, we construct a dummy variable that equals one over the period when a highway interchange appears in the vicinity of a municipality.

This indicator variable varies over time but takes on one for some municipalities starting from 1975, year corresponding to the introduction of the first highway interchange in the geographic where our municipalities are located. Next, we interact the highway presence dummy with the distance of the centroid of a municipality to the referenced highway interchange. In other words, the impact of the accessibility of a municipality to a highway interchange is only relevant for years over which the infrastructure was available.

Similarly, we control for the influence of other major railways that lived through the restructuring of the French railway networks, which occurred in the aftermath of the second world war. Following the introduction of the National Enterprise of the French Railways (SNCF), several other railways along with the ones under study were deemed unprofitable and shut down. The remaining network was restructured and will go on to be exploited by the SNCF. Interestingly, many of these lines still operate to the present. Under the assumption that conservation of a railway by the SNCF in the postwar era correlates with traffic and profitability along the line prior to the restructuring, we add the distance of a municipality to the closest of the referenced lines as a control. The assumption is that if treated municipalities are located close to other major lines that remain active to the present, we could be picking up the effects of this additional network in our measured treatment effects.

The results in table 7 indicate that both variables are significant with associated signs in line with expectations. Specifically, we observe that the distance to a highway interchange bears a negative association with the growth of population. More importantly, we observe that the long-run positive effect of railway accessibility shrinks from 2.9 to 1.8 percentage points once these post-treatment controls are included. This confirms that part of the long-run positive treatment effect estimate is inherent to the correlation between treatment status (proximity to an exploited railway segment) and future developments that affect access to highway transportation. The evidence points to the fact that the planning of transportation accessibility correlate with the presence of a railway in a locality in the past.

Table 9 in the appendix further illustrates this hypothesis. The table reports the results of the regressions of treatment status on the average distance to a highway interchange after controlling for other relevant factors. First, it is worth highlighting that the first highway interchange near the municipalities in our samples opened in 1975, long after the closing of all the studied railways. There is strong evidence in favor of the argument that highway interchanges are located closer to previously treated municipalities. Highway accessibility seems significantly higher near the treated municipalities. In that sense, the declassified railways may have solved a coordination problem for future infrastructures. The average distance from the centroid of a municipality to the closest road interchange is around eight kilometers shorter in the vicinity of the previously exploited segments of the railways.

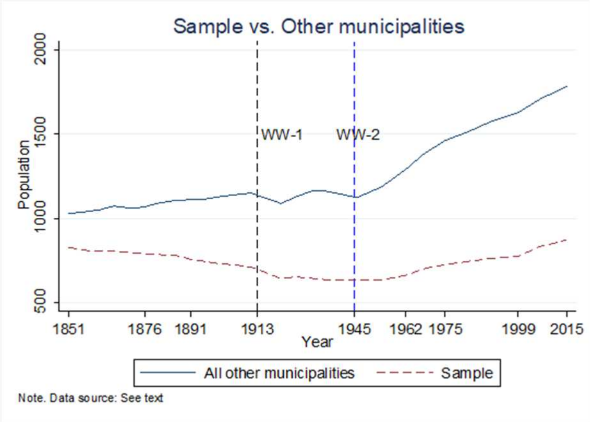
Similar to Berger and Enflo (2013), we hypothesize that the railways may have solved a coordination problem for other transportation networks that affect accessibility. Future infrastructures could follow pre-existing networks if the construction of transportation infrastructures involve fixed costs that could have been incurred in the past (sunk costs). For instance, the digging work and the complementary landlines that need to be set up prior to the exploitation of a highway could build on a pre-existing capital.

Yet, it also raises further questions with respect to the persistence of the treatment effects in the long-run even after controlling for other modes of accessibility. The access to the railways seem to have left a legacy that affected future urbanization and lived on long after the exploitation of the transportation infrastructure. The comparison of population trends between treated and control municipalities on figure 2 indicate that the divergence in growth paths began in the interwar period but accentuated starting from the mid-1970s. This observation along with the consideration that only one of the railways lasted through the second half of the twentieth century led us to explore how the exclusion of the referenced railway affects our results in the robustness checks.

6.2 Medium-run effects of the railways

The absence of a positive medium-run effect of the access to railways on population growth might seem odd at first, but is perfectly logical once we factor in the consideration that the railways were administrative in nature. Apart from the fact that the lines ran through rural areas, most of which were already on a downward population trend prior to the arrival of rail transportation, the period over which the infrastructures were exploited features some peculiarities. All active portions of the railways in our sample operated during the first world war and the Great depression. Figure 6 plots the evolution of municipal population in the municipalities in our sample relative to the rest of France. It is evident that population was not on an upward trajectory in France between the two wars. If anything, we would not expect the rural railways to be successful during an area that was also marked by other negative shocks to productivity.

Figure 6: Evolution of population in our sample vs. Other municipalities

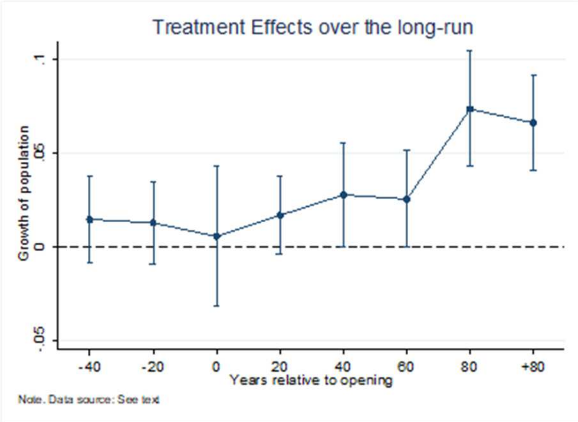


Nonetheless, the graphs described on figure 3 in the appendix also suggest a potential heterogeneity in the treatment effects by railway. For three of the railways namely “Beaumont-Gimont”, “Chorges-Barcelonnette” and “Saint-Juéry-Sainte-Affrique”, population trends after opening clearly indicate a jump around twenty years after opening. This time frame corresponds to the aftermath of the first world war. Given that these lines opened around the early 1900s and that the war broke out around 1913, it appears that the positive effects of access to railways are noticeable in the medium-run for some of the railways. This finding raises additional questions with respect to the heterogeneity in the dynamics of the treatment effects. The fact that the positive effect of access to rail transportation on population is perceptible very early in some places but only manifests itself in the long-run in other regions suggests that other factors are relevant in the medium-run.

The only distinctive feature of the aforementioned railways is that they relate to the most populous of the treated municipalities in our sample prior to the opening of these roads. Table 14 in the appendix describes the average municipal population for treated and control municipalities in 1896 for the three railways under scrutiny. The comparison indicates that the municipalities near the operated segment of the referenced railways, were the most populous of all the treated municipalities in our sample. The average population size in the municipalities belonging to this restricted sub-sample is 85 percent larger than the average population of the treated municipalities in the remaining railways, prior to the arrival of the infrastructure.

In the appendix, we present the results of regressions using this sample of railways. The positive treatment effects become significant around twenty and forty years after the lines opened. The medium-run positive effect of access to the railways on population growth is apparently only visible in densily populated municipalities where the structure of the productive sector yields a higher social saving. This also speaks to the importance of pre-existing agglomerations fostering more economic activity whenever alternative modes of transportation that reduce trading costs are introduced. Besides, we notice that the long-run effects are larger than the medium-run effects for these municipalities, and significantly larger than the long-run effects for other municipalities. Figure 7 below describes the evolution of the estimated treatment effects for this sub-sample of railways.

Figure 7: Estimated treatment effects for the three railways



6.3 Long-term legacy of the railroad

6.3.1 Population level

The divergence of population trends that emerged between treated and control municipalities in the postwar era had significant implications for population density in the long-run. Table 11 in the appendix reports the outcomes of the regressions of the log of population in 2006 and 2015 on treatment status and intensity of treatment. It appears that population is higher in municipalities near the exploited portion of the railways under study. For instance, for municipalities on the trajectory of the originally planned railway, those on the exploited section feature a population that was 50 percent and 58 percent larger in 1999 and 2015 respectively, with an associated population growth differential of 7.9 percent between 1999 and 2015. As expected, minor differences in annual growth of population could induce tremendous differences in the long-run. These results have been derived after controlling for the land area implying that the treated municipalities are denser on average.

5.2.2 Urbanization status

The significant difference in population level between treated and control municipalities in the long-run inexorably affected urbanization status. Using the urbanization definition of the French Institute of Statistics, which considers as urban a municipality of at least 2,000 inhabitants concentrated around its center, we classify the municipalities in our sample before the inception of the railways and long after their exploitation. Table 12 in the appendix describes the outputs of the regressions of the urbanization status on treatment status in 1851 and 2015 for different sub-samples. The results indicate that the probability of being urbanized has increased for treated municipalities regardless of the sub-sample considered, although these effects are significant only when we consider municipalities located within twenty kilometers of the railways (larger sample). In line with the results exposed in the previous section, there is also some evidence of a reversal in the probability of urbanization status between treated and control municipalities relative to the pre-railway year of 1851.

6.3.3 The density of the road network

The hypothesis that the delayed effects of the railways on population are inherent to the coordination of future transportation infrastructures cannot be directly tested. However, we observe the density and the length of the highway network in 2006 for all municipalities in France, corresponding to the earliest year for which this information is available. Using this GIS data, we compare treated and control municipalities along a projected railway. The results in table 13 describe the regressions of the density of the road network and the total length of the road network on treatment status with the inclusion of railway fixed effects and regional effects. The outputs in the table suggest that the municipalities with closer access to the active portion of the railways under study feature a higher density of the road network. Specifically, we measure a positive and significant difference of 0.07 kilometer of roads per square kilometer between treated and control municipalities.

This difference represents about 8.1 percent of the average density of the road networks in our sample. Relatedly, we noticed that the total length of the highway network is higher in treated municipalities. It also appears that these municipalities are 10.8 percent more likely to record the presence of a highway in their road network. These findings could be understood through two different lenses. First, the increased population near the active segments of the railways could affect the designing of future highways in a way that locate them near these denser locations. In this case, the positioning of highway road networks follow population. Alternatively, if the historical railways serve to solve the coordination problem discussed earlier, therefore other infrastructures could follow the railways leading to more population growth afterwards.

7. Robustness checks

7.1 Comparison of population trends prior to the arrival of the railroads

The eyeball interpretation of the graphical illustration described on figure 2 will lead us to conclude that prior to the inception of the railways under study, population trends are grossly similar between treatment and control municipalities. Yet, the railway specific graphs on figure 3 suggest that this conclusion might not be realistic for a couple of railways. To test this hypothesis in a robust fashion, we compare population levels and growth between 1851 and 1866 for the two sets of municipalities along a given projected railway. Table 11 in the appendix reports the regressions outputs for three separate regressions. The first column describes the regression of the log of population in 1851 on the treatment dummy, the distance to the projected railway and the interaction between these variables.

The results indicate that neither the treatment status nor the distance to the projected railroad had a significant impact on population in 1851, fifteen years prior to the opening of the first railway in our sample. The second column runs a similar regression for the year 1866, in reference of the opening of the first railway in the sample. Similar to 1851, we did not notice any significant difference between treated and control municipalities in 1861, regardless of the way treatment is defined (distance used as reference). The last column explores a potential difference in population trends between the two groups and explains the growth rate of population between 1851 and 1866 instead. However, we still do not observe any significant differences in the trends (growth rates) of population between the two sets prior to the arrival of the first railway in our sample.

All these results were derived with the inclusion of railway specific effects meaning that the comparison relates to municipalities located along a given railway. The evidence with respect to identical pre-treatment trends is stronger with the regressions displayed in table 11. Plus, the consideration that distance to the projected railroad does not correlate with population growth prior to the arrival of the transportation infrastructure further reinforces the presumption that the railroads in our sample were built ahead of demand and did not necessarily follow economic expansion.

7.2 Interaction of treatment status with distance to the railroad

In this paper, treatment status is defined through the distance of the centroid of a municipality to the planned railway. Though, we consider alternative windows of distance in our regressions, the discretion in the choice of 10-km as the baseline definition of treatment along with additional regressions using 15-km and 20-km windows, was solely justified on the basis of the average distance from the centroid of a municipality to a railroad in the region around 1930, at the height of rail transportation in France. However, if the intensity of treatment which we define as the distance of a municipality to the operated section of a railway matters for the treatment effects then, it might be appropriate to include the interaction between the treatment effects and the distance of to the active portion of a railway.

Table 8 in the appendix reports the regressions including the interaction between the treatment effects sixty and eighty years after treatment and the distance to the railway. We only included the interaction between distance to the railway and the treatment effects sixty and eighty years after the introduction of the railway because those were the only effects that were significant in the first place. It turns out that the treatment effect over both periods get weaker the farther away a municipality is located from the operational segment of the railway. This result is expected based on the findings when the distance window used to define the treatment status was relaxed to 15-km and 20-km respectively. We noticed in previous discussions that the average treatment effect was getting weaker as the distance used to define treatment was increasing.

7.3 Exclusion of the "Saint-Girons-Oust" railway

The railway "Saint-Girons-Oust" is peculiar in our sample. It was the first to open in 1866 and was closed in 1969 long after the other railways. Since, the effects of accessibility to the railroads on population growth were only visible sixty to eighty years after opening, it might be relevant to check if these delayed effects are driven by the only railway that was active for more than eighty years. Table 10 in the appendix reports the results with the exclusion of the "Saint-Girons-Oust" railway. This set of regressions yields conclusions similar to the ones made earlier, with stronger treatment effects sixty years after the lines opened. The results reinforce the hypothesis that the measured effects of railway accessibility emerge after the infrastructure's exploitation. This further points to the narrative that the treatment effects are the consequence of post-treatment developments that affect population growth. In the preceding sections, we found that treated municipalities were closer to highway interchanges and feature a higher density of the road network.

7.4 Spatial interactions

The identification strategy used in this paper is akin to a spatial discontinuity method. Treated and control municipalities are interrelated through space due to proximity. The positive effects that we measure in the medium-run for some of the railways could be upwardly biased due to a combination of factors. If the jumps in population growth on the treatment side of the roads are partly caused by a movement of population from the control to the treatment segment, then the estimated impacts would overestimate the economic benefits of the infrastructure. In this case, because there is a decrease of population on the control side, the difference in difference approach would capture both the positive growth effects and the loss of population along the control segment, which will overestimate the real growth accelerating impact of the railways.

To explore this possibility, we compare treated municipalities to an alternative set of control municipalities that are located between ten and twenty kilometers away from the control segments. The underlying assumption here is that if the spatial interactions described above are important, we should observe a treatment effect that is smaller in magnitude using this set of controls. As these municipalities are farther away from our treated municipalities, we don't expect the spatial interactions to be significant with this sample of controls. The results in table 15 suggest that the treatment effects get larger in fact with this alternative set of controls. This indicates that the positive effects that we measure for some of the railways are not driven by the spatial interactions between treated and control municipalities.

8. Discussion and limitations

Most of the railways under study opened during the early 1900s. This implies that the first estimated positive treatment effects which appear forty to sixty years after opening, coincide to the end of the second world war. For the railways running through the most populous regions, we notice that the positive effects appear much earlier. For the rest of the railways, the absence of a significant effect in the medium-run is indicative of the relevance of pre-arrival conditions. The conceptual framework described in section 3 explains how the size of the domestic economy before the arrival of the railways determines the positive incidence of the railroads on productivity and growth going forward.

For the estimated long-run effects, it is important to bear in mind that the reconstruction efforts that unfolded in the postwar era and the ensuing infrastructure investments could have benefited treated municipalities more due to the sunk costs and coordination problem discussed above. The fact that the long-run positive effects of the railways are stronger for municipalities that benefited from the infrastructure in the medium run, suggests that the sunk costs were probably larger in these places. Unfortunately, we are not able to gather additional informations on the presence of other public investments like schools, hospitals and the postal service in our sample of municipalities over the period when the railways were active.

Alternatively, for the other railways, the absence of a significant positive effect of access to railroad transportation on population growth in the medium-run could be explained by the fashion in which such projects are undertaken at the time. Many of the railways took a long time to be built. In some cases, construction on the control (unexploited) segment of the projected railway, carried on for a long time while the treatment portion was already being exploited. Knowing that these construction works generally involve several workers who settle temporarily nearby construction sites, it is possible that the positive effects of the railways on population growth are masked by the inflow of construction workers in the control municipalities.

Though, one can make the case that the internal validity of the results described in this paper are quite strong, the external validity remains opened to discussions. The heterogeneity of the medium-run effects highlight the necessity to consider pre-existing factors in the evaluation of the economic impact of railway transportation. The effect (or lack thereof) of the railways on population growth in some places features an interesting contrast. The positive effects in the medium-run in some places could be due to a higher cost saving per unit of transportation in these areas. The proximity of a major inland water stream or unobservable geographic characteristics that lower the opportunity cost of road transportation could reduce the benefits of the more advanced railway transportation. The implication from a policy standpoint is not clear-cut. Though, initial population seems to be the driving force for the economic benefits of the railways in our sample, this does not imply that railways are more likely to be successful in densily populated areas.

9. Concluding remarks

The economic benefit of transportation infrastructure remains the concern of policy discussions around the world. Estimating the potential welfare improvement of a new railway transportation infrastructure is challenging *a priori* and policymakers are forced to rely on empirical estimates. The impacts of historical railways on urbanization and economic growth has been widely documented in several advanced countries. Aside from identification concerns, many of the papers focus exclusively on short and medium-run effects. This paper observes that the long-run effects could greatly supersede the short and medium-run impacts.

First, we found that access to a new railway translates into prosperity and higher population growth in the medium-run but only in places with a relatively important initial population level. Second, these modest medium-run effects are magnified in the long-run leading up to an acceleration of population growth, long after the inception of the transportation infrastructure. We hypothesize that the railways must have solved a coordination problem for future infrastructure investments. To support this thesis, we show that the municipalities that gained access to the historical railways are located closer to future highway interchanges and are more likely to be on the trajectory of a highway. This can be explained by sunk investments which lower the opportunity costs of infrastructure projects in some areas.

We use a natural experiment entailed by the timing of railway constructions in France and a comprehensive dataset of French population censuses going back to 1850. Due to unexpected hazards, several of the French railways that were planned in the Freycinet Plan were never fully achieved, although significant portions of the construction works were laid down. We compare municipalities located near the achieved segments of these railways to a set of control municipalities located close to the portion that was planned but never realized. Two main results emerge from our analysis. First, we observe that accessibility to the railroad fostered population growth in the medium-run for a few railways, specifically those that ran through places with relatively important initial population.

Possibly, this means that the success of the infrastructure is conditioned on the social saving of the service, which is high in places with a significant productive sector. Second, we noted that the presence of the railway affects long-run population growth, and more so for the municipalities that benefited from the service earlier. The underlying mechanism relates to a coordination problem. The railways seemed to have become a reference point for highways and other roads long after their exploitation. To this effect, we showed that municipalities near the railroads were much closer to future highway interchanges and more likely to be crossed by a highway.

The effect (or lack thereof) of the railways on population growth in some places features an interesting contrast. The positive effects in the medium-run in some places could be due to a higher cost saving per unit of transportation in these areas. In contrast, the proximity of a major inland water stream or unobservable geographic characteristics that lower the opportunity cost of road transportation could reduce the benefits of the railway transportation. The implication from a policy standpoint is not clear-cut. Though, initial population seems to be the driving force for the economic benefits of the railways in our sample, this does not imply that railways are more likely to be successful in densily populated areas.

References

Alvarez, Eduard, Xavi Franch, and Jordi Martí-Henneberg. "Evolution of the territorial coverage of the railway network and its influence on population growth: The case of England and Wales, 1871-1931." Historical Methods: *A Journal of Quantitative and Interdisciplinary History 46.3 (2013): 175-191*

Atack, Jeremy, Fred Bateman, Michael Haines, and Robert Margo. 2010. "Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest 1850-60." *Social Science History*, 34: 171–197.

Atack, J., Haines, M., and Margo, R. 2008. Railroads and the Rise of the Factory: Evidence for the United States, 1850-1870. *NBER Working Paper 14410*

Atack, J. and Margo, R. A. 2011. The Impact of Access to Rail Transportation on Agricultural Improvement: The American Midwest as a Test Case, 1850-1860. *Journal of Transport and Land Use 4, 2*

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

Baum-Snow, Nathaniel. 2007. "Did Highways Cause Suburbanization?" *The Quarterly Journal of Economics*, 122(2): 775–805.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. 2012. Roads, Railroads and Decentralization of Chinese Cities. *Mimeo, University of Toronto*.

Bertrand, M., Duflo, E., and Mullainathan, S. 2004. How Much Should We Trust Differences-in-Differences Estimates? *The Quarterly Journal of Economics 119, 1, 249-275*.

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Portage and Path Dependence." *The Quarterly Journal of Economics*, 127(2): 587–644.

Bogart, D. and Chaudhary, L. 2013. Engines of growth: The productivity advance of Indian railways, 1874-1912. *The Journal of Economic History 73, 02, 339-370*.

Bosker, Maarten, Steven Brakman, Harry Garretsen, and Marc Schramm. 2007. "Looking for multiple equilibria when geography matters: German city growth and the WWII shock." *Journal of Urban Economics*, 61(1): 152–169.

Bosker, Maarten, Steven Brakman, Harry Garretsen, and Marc Schramm. 2008. "A century of shocks: The evolution of the German city size distribution 1925-1999." *Regional Science and Urban Economics, 38(4): 330–347*

Casson, Mark. "The determinants of local population growth: A study of Oxfordshire in the nineteenth century." *Explorations in Economic History 50.1 (2013): 28-45*

Davis, Donald R., and David E. Weinstein. 2002. "Bones, Bombs, and Break Points: The Geography of Economic Activity." *American Economic Review, 92(5): 1269–1289*.

Davis, D. R. and Weinstein, D. E. 2008. A Search for Multiple Equilibria in Urban Industrial Structure. *Journal of Regional Science 48, 1, 29-65*.

Donaldson, D. (2017). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. forthcoming The American Economic Review.

Donaldson, D. and Hornbeck, R. (2016). Railroads and American Economic Growth: A \Market Access" Approach. The Quarterly Journal of Economics, 131 799{858.

Duranton, G. and Puga, D. 2004. Micro-Foundations of Urban Agglomeration Economies. *Handbook of regional and urban economics 4, 2063-2117.*

Duranton, G. and Turner, M. A. 2012. Urban Growth and Transportation. The Review of Economic Studies 79, 4, 1407-1440.

Fishlow, A. (1965). American railroads and the transformation of the antebellum economy. *Cambridge, Harvard University Press.*

Fogel, R. W. (1964). Railroads and American economic growth: essays in econometric history. *Baltimore, Johns Hopkins Press.*

Goodrich, C. (1961). “Canals and American economic development, y Carter Goodrich [and others]”. *New York, Columbia University Press.*

Haines, M. and Margo, R. 2006. “Railroads and Local Economic Development: The United States in the 1850s”. *NBER Working Paper 12381*

Haites, Erik; Mak, James; and Gary Walton. Western River Transportation: The Era of Internal Development, 1810-1860. *Baltimore, MD: Johns Hopkins University Press.*

Herranz-Loncan, A. 2006. “Railroad impact in backward economies: Spain, 1850-1913”. *The Journal of Economic History 66, 04, 853-881.*

Jedwab, R. and Moradi, A. 2011. “Transportation Infrastructure and Development in Ghana”. *Mimeo.*

Jenks, L. H. (1944). "Railroads as an Economic Force in American Development." *The Journal of Economic History* 4(1): 1-20

Miguel, E. and Roland, G. 2011. “The Long-Run Impact of Bombing Vietnam”. *Journal of Development Economics 96, 1,* 1-15.

Ransom, Roger. 1970. “Social Returns from Public Transport Investment: A Case Study of the Ohio Canal” *Journal of Political Economy* 78: 1041-1064

Stover, John F. 1978. “Iron Roads to the West: American Railroads in the 1850s”. *New York: Columbia University Press.*

T. Thevenin, C. Mimeur, R. Schwartz, and L. Sapet (2016). “Measuring one century of railway accessibility and population change in France. A historical GIS approach” *Journal of Transport Geography, Vol. 56, pp 62-76.*

Table 4: Difference in population growth for municipalities within 10-km of the planned railways

	(1)	(2)	(3)	(4)
Difference in difference	ΔLog(Pop.)	ΔLog(Pop.)	ΔLog(Pop.)	ΔLog(Pop.)
Treatment dummy	0.00268	0.00268	0.00538	0.00568
	[0.55]	[0.55]	[0.79]	[0.83]
Period-Treatment interaction				
40 to 20 Years Prior to Opening ([-40,-20[)	0.000227	0.000227	-0.00591	-0.00551
	[0.04]	[0.04]	[-0.72]	[-0.67]
20 to 10 Years Prior to Opening ([-20, 0[)	-0.00159	-0.00159	-0.00478	-0.00467
	[-0.28]	[-0.28]	[-0.62]	[-0.61]
Year of Opening (Year 0)	-0.00555	-0.00555	-0.00648	-0.00657
	[-0.68]	[-0.68]	[-0.59]	[-0.60]
20 Years After Opening ([0, 20[)	-0.00224	-0.00224	-0.00335	-0.00340
	[-0.41]	[-0.41]	[-0.45]	[-0.46]
20 to 40 Years After Opening ([20, 40[)	-0.000501	-0.000501	-0.00269	-0.00292
	[-0.08]	[-0.08]	[-0.32]	[-0.35]
40 to 60 Years After Opening ([40, 60[)	0.00355	0.00355	0.00597	0.00585
	[0.56]	[0.56]	[0.69]	[0.67]
60 to 80 Years After Opening ([60, 80[)	0.0144*	0.0144*	0.0284***	0.0281***
	[1.82]	[1.82]	[2.80]	[2.78]
More than 80 Years After Opening (> 80)	0.0391***	0.0391***	0.0292***	0.0289***
	[6.03]	[6.03]	[3.23]	[3.19]
Log (Population 1850)	-0.00130	-0.00130	0.000223	0.00566***
	[-1.15]	[-1.15]	[0.19]	[2.83]
Presence of a train station	0.0160***	0.0160***	0.0162***	0.0156***
	[8.36]	[8.36]	[8.48]	[8.17]
Distance to a water stream				-0.000690***
				[-3.19]
Log(Land area)				-0.00794***
				[-3.93]
Intercept	-0.00106	-0.00106	-0.0198	0.00948
	[-0.05]	[-0.05]	[-1.32]	[0.61]
Railway-Census Year Effect	Yes	Yes	Yes	Yes
Railway Fixed Effect	No	Yes	Yes	Yes
Region-Census Year Effect	No	No	Yes	Yes
Location controls	No	No	No	Yes
Number of Municipalities	22400	22400	22400	22400
R- squared	0.190	0.190	0.228	0.230

Notes: OLS regressions using a panel of 1,634 municipalities over the period 1851-2015 and located within ten kilometers of the originally planned railway segments. t-statistics based on robust standard errors in brackets. Estimates of the period dummies are not reported for the sake of space. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable represents the growth of population in municipality *c* between two consecutive population censuses. This table shows the difference in difference estimate of the treatment effect of the railway exploitation on population growth in a lead-lag regression structure. In column (1), we include a railway-year effect to compare treated and control municipalities in a given census year. In column (2), we add railway fixed effects to control for the location and the attributes of each rail route (length, trajectory, etc.). In column (3), we add department-year effect to control for policy and other time variant developments at the department level. Last, column (4) includes municipality specific geographic controls such as the distance to an important inland water stream (See text for explanation) and the land area of the municipality. Other controls featured in all regressions include the log of the municipal population in 1850 and the presence of a train station in municipality *c* during the census year *t*.

Table 5: Difference in population growth for municipalities within 15-km of the planned railways

	(1)	**(2)**	**(3)**	**(4)**
Difference in difference	ΔLog(Pop.)	ΔLog(Pop.)	ΔLog(Pop.)	ΔLog(Pop.)
Treatment dummy	0.00724	0.00724	0.00942	0.00907
	[1.61]	[1.61]	[1.53]	[1.48]
Period-Treatment interaction				
40 to 20 Years Prior to Opening ([-40,-20[)	-0.00596	-0.00596	-0.0100	-0.00938
	[-1.13]	[-1.13]	[-1.36]	[-1.27]
20 to 10 Years Prior to Opening ([-20, 0[)	-0.00518	-0.00518	-0.0101	-0.00956
	[-1.05]	[-1.05]	[-1.51]	[-1.43]
Year of Opening (Year 0)	-0.00828	-0.00828	-0.0113	-0.0108
	[-1.32]	[-1.32]	[-1.34]	[-1.28]
20 Years After Opening ([0, 20[)	-0.00564	-0.00564	-0.00658	-0.00602
	[-1.16]	[-1.16]	[-1.01]	[-0.92]
20 to 40 Years After Opening ([20, 40[)	-0.00341	-0.00341	-0.00675	-0.00625
	[-0.61]	[-0.61]	[-0.89]	[-0.83]
40 to 60 Years After Opening ([40, 60[)	0.00204	0.00204	0.00360	0.00405
	[0.37]	[0.37]	[0.48]	[0.54]
60 to 80 Years After Opening ([60, 80[)	0.00747	0.00747	0.0213^{**}	0.0217^{**}
	[1.08]	[1.08]	[2.31]	[2.35]
More than 80 Years After Opening (> 80)	0.0345^{***}	0.0345^{***}	0.0271^{***}	0.0276^{***}
	[6.03]	[6.03]	[3.39]	[3.46]
Lop (Population 1850)	-0.000703	-0.000703	0.000902	0.00426^{***}
	[-0.83]	[-0.83]	[1.06]	[2.88]
Presence of a train station	0.0135^{***}	0.0135^{***}	0.0143^{***}	0.0136^{***}
	[8.71]	[8.71]	[9.26]	[8.82]
Distance to a water stream				-0.00107^{***}
				[-6.49]
Log(Land area)				-0.00485^{***}
				[-3.15]
Intercept	-0.0107	-0.0107	-0.0196^{*}	0.000376
	[-0.78]	[-0.78]	[-1.70]	[0.03]
Railway-Census Year Effect	Yes	Yes	Yes	Yes
Railway Fixed Effect	No	Yes	Yes	Yes
Region-Census Year Effect	No	No	Yes	Yes
Location controls	No	No	No	Yes
Number of Municipalities	36218	36218	36218	36218
R- squared	0.174	0.174	0.201	0.204

Notes: OLS regressions using a panel of 1,165 municipalities over the period 1851-2015 and located within fifteen kilometers of the originally planned railway segments. t-statistics based on robust standard errors in brackets. Estimates of the period dummies are not reported for the sake of space. *p<0.10, **p< 0.05, ***p< 0.01. The dependent variable represents the growth of population in municipality c between two consecutive population censuses. This table shows the difference in difference estimate of the treatment effect of the railway exploitation on population growth in a lead-lag regression structure. In column (1), we include a railway-year effect to compare treated and control municipalities in a given census year. In column (2), we add railway fixed effects to control for the location and the attributes of each rail route (length, trajectory, etc.). In column (3), we add department-year effect to control for policy and other time variant developments at the department level. Last, column (4) includes municipality specific geographic controls such as the distance to an important inland water stream (See text for explanation) and the land area of the municipality. Other controls featured in all regressions include the log of the municipal population in 1850 and the presence of train station in municipality *c* during the census year *t*.

Table 6: Difference in population growth for municipalities within 20-km of the planned railways

	(1)	(2)	(3)	(4)
Difference in difference	ΔLog(Pop.)	ΔLog(Pop.)	ΔLog(Pop.)	ΔLog(Pop.)
Treatment dummy	0.00901**	0.00901**	0.00813	0.00705
	[2.15]	[2.15]	[1.34]	[1.16]
Period-Treatment interaction				
40 to 20 Years Prior to Opening ([-40,-20[)	-0.00534	-0.00534	-0.00642	-0.00512
	[-1.12]	[-1.12]	[-0.92]	[-0.73]
20 to 10 Years Prior to Opening ([-20, 0[)	-0.00594	-0.00594	-0.00790	-0.00647
	[-1.31]	[-1.31]	[-1.19]	[-0.97]
Year of Opening (Year 0)	-0.0117**	-0.0117**	-0.0129*	-0.0113
	[-2.14]	[-2.14]	[-1.67]	[-1.46]
20 Years After Opening ([0, 20[)	-0.00842*	-0.00842*	-0.00627	-0.00444
	[-1.88]	[-1.88]	[-0.98]	[-0.70]
20 to 40 Years After Opening ([20, 40[)	-0.00327	-0.00327	-0.00526	-0.00342
	[-0.67]	[-0.67]	[-0.75]	[-0.49]
40 to 60 Years After Opening ([40, 60[)	-0.000220	-0.000220	0.00360	0.00513
	[-0.04]	[-0.04]	[0.51]	[0.73]
60 to 80 Years After Opening ([60, 80[)	0.00580	0.00580	0.0204**	0.0220***
	[0.97]	[0.97]	[2.45]	[2.64]
More than 80 Years After Opening (> 80)	0.0302***	0.0302***	0.0250***	0.0269***
	[5.84]	[5.84]	[3.35]	[3.60]
Lop (Population 1850)	-0.00115*	-0.00115*	0.000453	0.00311***
	[-1.68]	[-1.68]	[0.66]	[2.75]
Presence of a train station	0.0162***	0.0162***	0.0163***	0.0154***
	[12.92]	[12.92]	[13.00]	[12.37]
Distance to a water stream				-0.00129***
				[-9.70]
Log(Land area)				-0.00415***
				[-3.52]
Intercept	-0.0140	-0.0140	-0.0241**	-0.00397
	[-1.33]	[-1.33]	[-2.47]	[-0.40]
Railway-Census Year Effect	Yes	Yes	Yes	Yes
Railway Fixed Effect	No	Yes	Yes	Yes
Region-Census Year Effect	No	No	Yes	Yes
Location controls	No	No	No	Yes
Number of Municipalities	50896	50896	50896	50896
R- squared	0.170	0.170	0.195	0.198

Notes: OLS regressions using a panel of 1,634 municipalities over the period 1851-2015 and located within twenty kilometers of the originally planned railway segments. t-statistics based on robust standard errors in brackets. Estimates of the period dummies are not reported for the sake of space. $^{*}p< 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The dependent variable represents the growth of population in municipality c between two consecutive population censuses. This table shows the difference in difference estimate of the treatment effect of the railway exploitation on population growth in a lead-lag regression structure. In column (1), we include a railway-year effect to compare treated and control municipalities in a given census year. In column (2), we add railway fixed effects to control for the location and the attributes of each rail route (length, trajectory, etc.). In column (3), we add department-year effect to control for policy and other time variant developments at the department level. Last, column (4) includes municipality specific geographic controls such as the distance to an important inland water stream (See text for explanation) and the land area of the municipality. Other controls featured in all regressions include the log of the municipal population in 1850 and the presence of train station in municipality *c* during the census year *t*.

Table 7: Regressions with the control of highway interchange accessibility

	(1)	(2)	(3)	(4)
Difference in difference	ΔLog(Pop.)	ΔLog(Pop.)	ΔLog(Pop.)	ΔLog(Pop.)
Treatment dummy	0.00207	0.00207	-0.00399	-0.00323
	[0.43]	[0.43]	[-0.57]	[-0.46]
Period-Treatment interaction				
40 to 20 Years Prior to Opening ([-40,-20[)	-0.00459	-0.00459	-0.0102	-0.00987
	[-0.75]	[-0.75]	[-1.23]	[-1.19]
20 to 10 Years Prior to Opening ([-20, 0[)	-0.00934*	-0.00934*	-0.00994	-0.00963
	[-1.67]	[-1.67]	[-1.29]	[-1.25]
Year of Opening (Year 0)	-0.0168**	-0.0168**	-0.0124	-0.0121
	[-2.04]	[-2.04]	[-1.12]	[-1.10]
20 Years After Opening ([0, 20[)	-0.0149***	-0.0149***	-0.0102	-0.00988
	[-2.73]	[-2.73]	[-1.36]	[-1.31]
20 to 40 Years After Opening ([20, 40[)	-0.0146**	-0.0146**	-0.00935	-0.00911
	[-2.25]	[-2.25]	[-1.11]	[-1.08]
40 to 60 Years After Opening ([40, 60[)	-0.00770	-0.00770	0.000600	0.000818
	[-1.22]	[-1.22]	[0.07]	[0.09]
60 to 80 Years After Opening ([60, 80[)	0.00324	0.00324	0.0219**	0.0220**
	[0.41]	[0.41]	[2.15]	[2.17]
More than 80 Years After Opening (> 80)	0.0168**	0.0168**	0.0180**	0.0182**
	[2.49]	[2.49]	[1.96]	[1.97]
Lop (Population 1850)	-0.000385	-0.000385	-0.000143	0.00417**
	[-0.34]	[-0.34]	[-0.13]	[2.09]
Presence of a train station	0.0138***	0.0138***	0.0123***	0.0123***
	[7.29]	[7.29]	[6.41]	[6.41]
Distance to closest major railway	-0.00129***	-0.00129***	-0.00156***	-0.00145***
	[-17.50]	[-17.50]	[-14.29]	[-12.89]
Presence*distance to highway interchange	-0.000853***	-0.000853***	-0.000780***	-0.000794***
	[-5.13]	[-5.13]	[-4.05]	[-4.11]
Distance to a water stream				-0.000123
				[-0.57]
Log(Land Area)				-0.00608***
				[-3.01]
Intercept	0.00695	0.00695	-0.00435	0.0144
	[0.34]	[0.34]	[-0.29]	[0.93]
Railway-Census Year Effect	Yes	Yes	Yes	Yes
Railway Fixed Effect	No	Yes	Yes	Yes
Region-Census Year Effect	No	No	Yes	Yes
Location controls	No	No	No	Yes
Number of Municipalities	22400	22400	22400	22400
R- squared	0.202	0.202	0.236	0.236

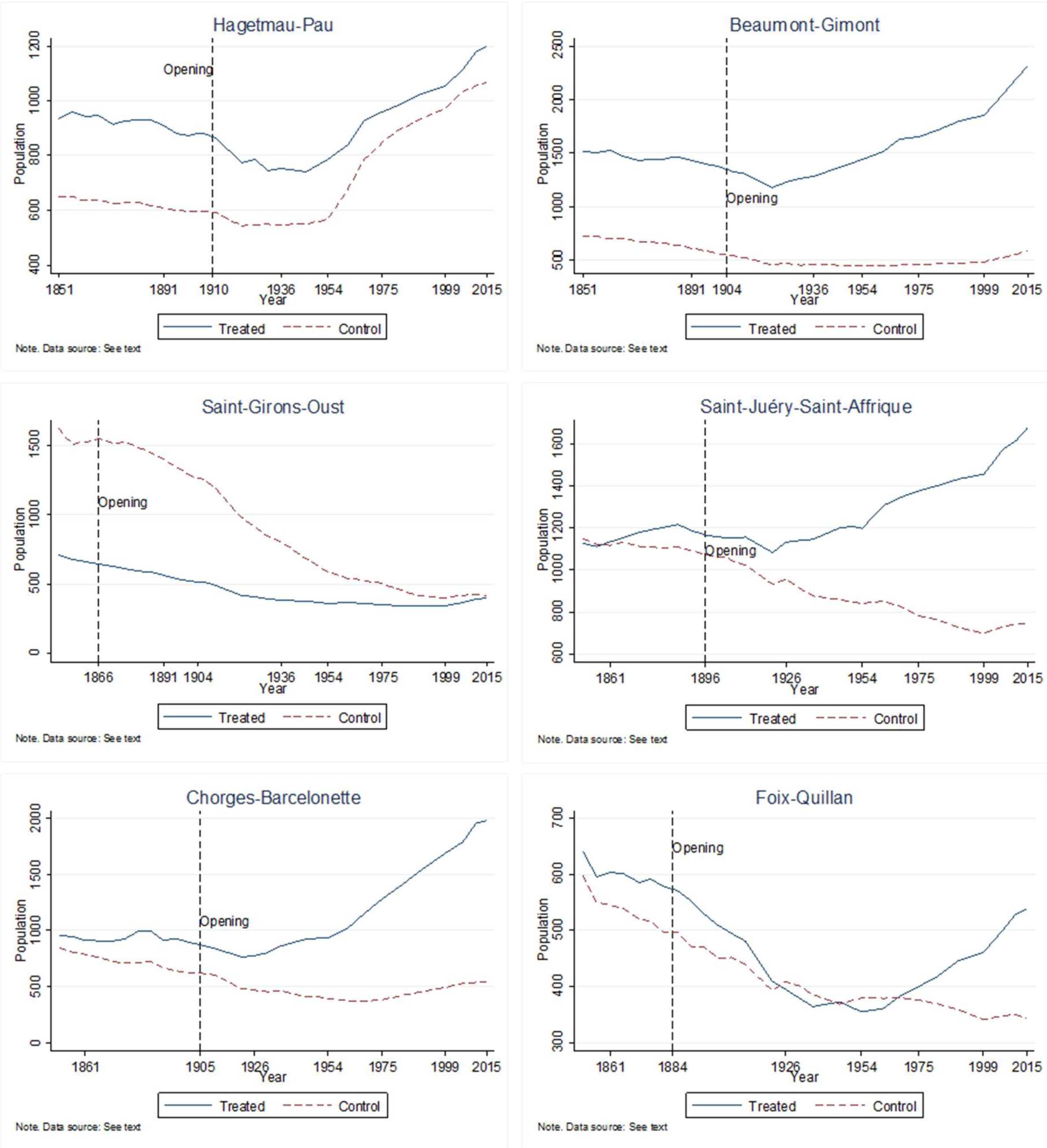
Notes: OLS regressions using a panel of 718 municipalities over the period 1851-2015 and located within ten kilometers of the originally planned railway segments. t-statistics based on robust standard errors in brackets. Estimates of the period dummies are not reported for the sake of space. *p < 0.10, **p < 0.05, ***p < 0.01. The dependent variable represents the growth of population in municipality *c* between two consecutive population censuses. This table shows the difference in difference estimate of the treatment effect of the railway exploitation on population growth in a lead-lag regression structure. In column (1), we include a railway-year effect to compare treated and control municipalities in a given census year. In column (2), we add railway fixed effects to control for the location and the attributes of each rail route (length, trajectory, etc.). In column (3), we add department-year effect to control for policy and other time variant developments at the department level. Last, column (4) includes municipality specific geographic controls such as the distance to an important inland water stream (See text for explanation) and the land area of the municipality. Other controls featured in all regressions include the log of the municipal population in 1850, the presence of a train station in municipality *c* during the census year *t*, the presence of a highway interchange within 50 kilometers and the interaction of this variable with the distance to the highway interchange.

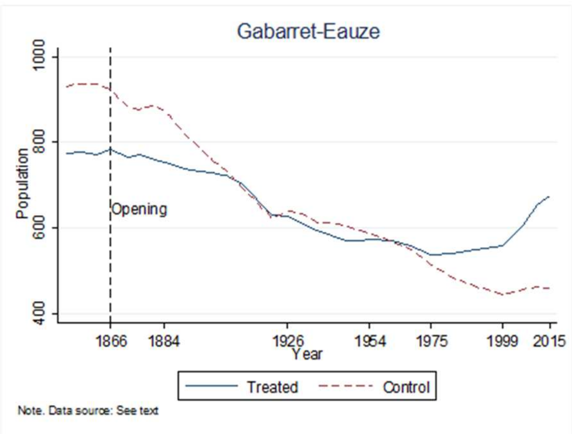
Table 8: Results with the interaction of the treatment effects and the distance to the railway

	(1)	(2)	(3)	(4)
Difference in difference	ΔLog(Pop.)	ΔLog(Pop.)	ΔLog(Pop.)	ΔLog(Pop.)
Treatment dummy	0.00905**	0.00905**	0.00816	0.00710
	[2.16]	[2.16]	[1.34]	[1.16]
Period-Treatment interaction				
40 to 20 Years Prior to Opening ([-40,-20[)	-0.00536	-0.00536	-0.00643	-0.00515
	[-1.12]	[-1.12]	[-0.92]	[-0.73]
20 to 10 Years Prior to Opening ([-20, 0[)	-0.00596	-0.00596	-0.00791	-0.00651
	[-1.32]	[-1.32]	[-1.19]	[-0.98]
Year of Opening (Year 0)	-0.0117**	-0.0117**	-0.0130*	-0.0114
	[-2.15]	[-2.15]	[-1.67]	[-1.46]
20 Years After Opening ([0, 20[)	-0.00846*	-0.00846*	-0.00629	-0.00450
	[-1.89]	[-1.89]	[-0.99]	[-0.70]
20 to 40 Years After Opening ([20, 40[)	-0.00331	-0.00331	-0.00528	-0.00347
	[-0.67]	[-0.67]	[-0.76]	[-0.50]
40 to 60 Years After Opening ([40, 60[)	-0.000251	-0.000251	0.00361	0.00512
	[-0.05]	[-0.05]	[0.51]	[0.73]
60 to 80 Years After Opening ([60, 80[)	0.0113	0.0113	0.0292***	0.0298***
	[1.50]	[1.50]	[3.13]	[3.20]
More than 80 Years After Opening (> 80)	0.0453***	0.0453***	0.0413***	0.0419***
	[7.40]	[7.40]	[5.07]	[5.14]
60 to 80 Years After Opening (> 80)∗distance	-0.000493	-0.000493	-0.000790*	-0.000700
	[-1.07]	[-1.07]	[-1.73]	[-1.55]
More than 80 Years After Opening (> 80)∗distance	-0.00135***	-0.00135***	-0.00147***	-0.00136***
	[-4.64]	[-4.64]	[-5.03]	[-4.67]
Lop (Population 1850)	-0.00119*	-0.00119*	0.000424	0.00305***
	[-1.75]	[-1.75]	[0.62]	[2.69]
Presence of a train station	0.0160***	0.0160***	0.0159***	0.0151***
	[12.78]	[12.78]	[12.76]	[12.16]
Distance to a water stream				-0.00126***
				[-9.56]
Log(Land area)				-0.00410***
				[-3.48]
Intercept	-0.0137	-0.0137	-0.0239**	-0.00407
	[-1.30]	[-1.30]	[-2.45]	[-0.41]
Railway-Census Year Effect	Yes	Yes	Yes	Yes
Railway Fixed Effect	No	Yes	Yes	Yes
Region-Census Year Effect	No	No	Yes	Yes
Location controls	No	No	No	Yes
Number of Municipalities	50896	50896	50896	50896
R- squared	0.171	0.171	0.195	0.198

Notes: OLS regressions using a panel of 1,634 municipalities over the period 1851-2015 and located within twenty kilometers of the originally planned railway segments. t-statistics based on robust standard errors in brackets. Estimates of the period dummies are not reported for the sake of space. *p < 0.10, **p < 0.05, ***p < 0.01. The dependent variable represents the growth of population in municipality *c* between two consecutive population censuses. This table shows the difference in difference estimate of the treatment effect of the railway exploitation on population growth in a lead-lag regression structure. In column (1), we include a railway-year effect to compare treated and control municipalities in a given census year. In column (2), we add railway fixed effects to control for the location and the attributes of each rail route (length, trajectory, etc.). In column (3), we add department-year effect to control for policy and other time variant developments at the department level. Last, column (4) includes municipality specific geographic controls such as the distance to an important inland water stream (See text for explanation) and the land area of the municipality. Other controls featured in all regressions include the log of the municipal population in 1850 and the presence of a train station in municipality *c* during the census year *t*. The treatment effect is interacted with the intensity of treatment measured through the distance of a treated municipality to the railway.

Figure 3: Long-run average population trends for treated and control municipalities by railway





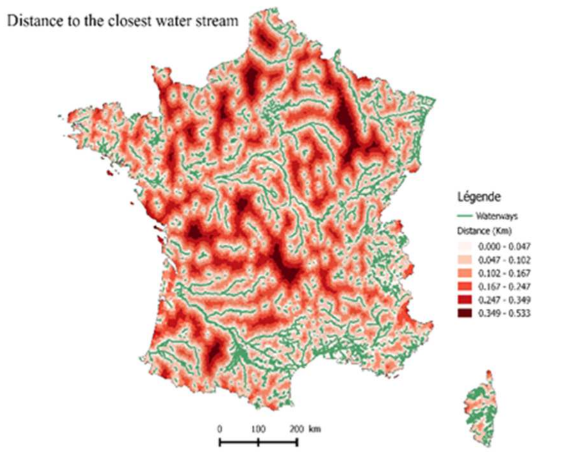


Table 10: Population trends between treated and control municipalities before opening

	(1)	(2)	(3)
	Log (Pop. 1851)	Log (Pop. 1866)	ΔLog (Pop 1851-1866)
Treatment Dummy	0.0831	0.0517	-0.0141
	[1.52]	[0.91]	[-0.32]
Distance to the Railway	-0.00304	-0.00413	0.000735
	[-1.10]	[-1.43]	[0.33]
Treatment Dummy∗Distance	0.000596	0.00339	0.0000599
	[0.14]	[0.75]	[0.02]
Log (Land Area)	0.867^{***}	0.887^{***}	0.0405^{***}
	[40.94]	[40.61]	[2.86]
Distance to a water stream	-0.0438^{***}	-0.0447^{***}	0.00809
	[-13.40]	[-13.47]	[1.44]
Intercept	0.432^{**}	0.299^{*}	-0.264^{*}
	[2.49]	[1.69]	[-1.90]
Region Fixed effects	Yes	Yes	Yes
Railway Effects	Yes	Yes	Yes
Number of Observations	1598	1606	1634
R-squared	0.664	0.668	0.127

Notes: OLS regressions using the full sample of 1,634 municipalities located within twenty kilometers of the originally planned railway segments. t-statistics based on robust standard errors in brackets. $^{*}p< 0.10$, $^{**}p<0.05$, $^{***}p<0.01$. The regressions in this table compare population levels and trends between treated and control municipalities prior to the opening of the railways. The years of comparison (1851 and 1866) correspond to the first census period in our sample and the last census period before the first line opened. The inclusion of railway effects is meant to compare treated and control municipalities along a planned railway. In column (1), we compare average population levels between the two sets in 1851. The distance to the (not built at the time) railway is included to check if proximity to the future railways is correlated with pre-exploitation population levels. In column (2), we ran the same regression for the census year 1866. In the last column (3), we ran a similar regression with the growth of population between 1851 and 1866 as the dependent variable.

Table 11: Population trends of treated and control municipalities long after exploitation

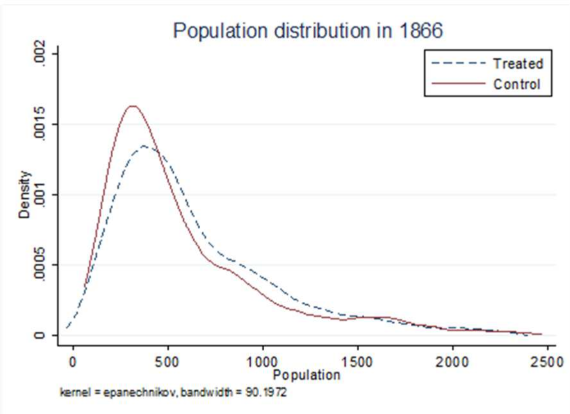
	(1)	(2)	(3)
	Log (Pop. 1999)	Log (Pop. 2015)	ΔLog (Pop 1999-2015)
Treatment Dummy	0.507^{***}	0.586^{***}	0.0793^{***}
	[4.64]	[5.40]	[3.91]
Distance to the Railway	-0.0141^{**}	-0.0154^{**}	-0.00129
	[-2.28]	[-2.50]	[-1.17]
Treatment Dummy∗Distance	-0.00368	-0.00339	0.000288
	[-0.43]	[-0.40]	[0.17]
Log (Land Area)	0.912^{***}	0.880^{***}	-0.0322^{***}
	[23.56]	[22.74]	[-4.60]
Distance to a water stream	-0.0923^{***}	-0.0927^{***}	-0.000427
	[-16.97]	[-16.81]	[-0.37]
Intercept	-0.583^{*}	-0.351	0.232^{***}
	[-1.80]	[-1.08]	[3.60]
Region Fixed effects	Yes	Yes	Yes
Railway Effects	Yes	Yes	Yes
Number of Observations	1634	1634	1634
R-squared	0.447	0.445	0.174

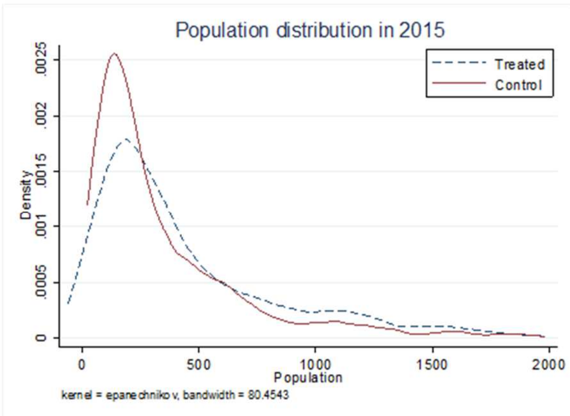
Notes: OLS regressions using the full sample of 1,634 municipalities located within twenty kilometers of the originally planned railway segments. t-statistics based on robust standard errors in brackets. $^{*}p< 0.10$, $^{**}p<0.05$, $^{***}p<0.01$. The regressions in this table compare population levels and trends between treated and control municipalities long after the exploitation of the railways. The inclusion of railway effects is meant to compare treated and control municipalities along a planned railway. In column (1), we compare average population levels between the two sets for the census year 1999. The distance to the (previously exploited and closed) railway is included to capture the effect of the intensity of treatment on population growth. In column (2), we ran the same regression for the census year 2015. In the last column (3), we ran a similar regression with the growth of population between 1999 and 2015 as the dependent variable.

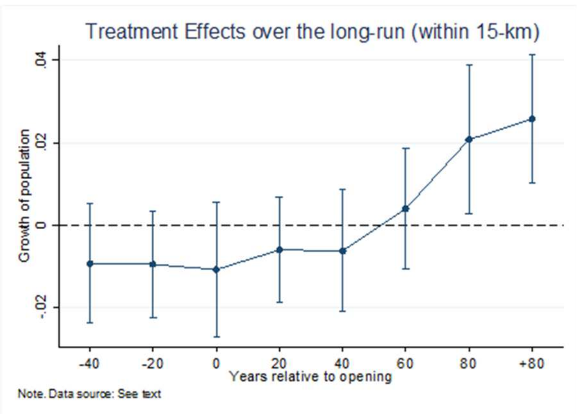
Table 12: Urbanization status before and long after treatment

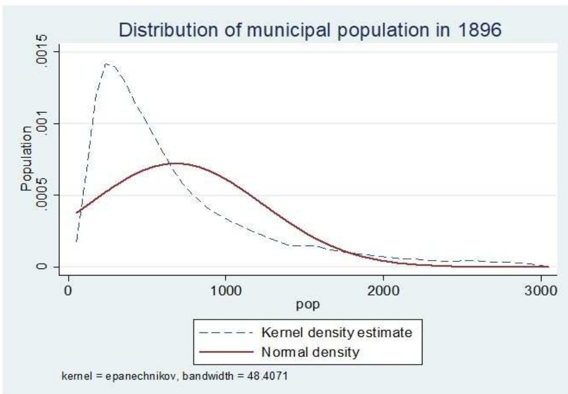
Sample	**within 10-km**		**within 15-km**		**within 20-km**	
	Urbanized (in 1851)	Urbanized (in 2015)	Urbanized (in 1851)	Urbanized (in 2015)	Urbanized (in 1851)	Urbanized (in 2015)
Treatment dummy	-0.033	0.0389	-0.0283	0.0284	-0.0263	0.0406**
	[-1.30]	[1.48]	[-1.43]	[1.45]	[-1.53]	[2.38]
Intercept	-0.805***	-0.730***	-0.874***	-0.563***	-0.821***	-0.649***
	[-5.30]	[-4.84]	[-7.74]	[-5.25]	[-8.74]	[-7.04]
Railway Effects	Yes	Yes	Yes	Yes	Yes	Yes
Department Effects	Yes	Yes	Yes	Yes	Yes	Yes
Location controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	718	718	1165	1165	1634	1634
R-squared	0.207	0.119	0.192	0.101	0.179	0.114

Notes: OLS regressions of a total sample of 1,634 municipalities. t-statistics based on robust standard errors in brackets. t-statistics in brackets. *p <0.10, **p <0.05, ***p <0.01. The dependent variables represent the urbanization status of a municipality in 1851 and 2015 respectively (dummy variable equals one if municipality is urban. The regressions consider respectively the sample of municipalities within 10-km of the planned railways (columns 1 and 2), within 15-km (columns 3 and 4) and within 20-km (columns 5 and 6). Following the norm adopted by the French Statistics Institute (INSEE) going back to 1846, we consider as urbanized, municipalities with a population of more than 2,000 inhabitants. The location controls include the land area and the distance to a major water stream.









Appendix: Selected quotes from newspapers and online sources

Beaumont-Gimont

…The railway Moissac-Cahors declared as being of public interest in 1879 will have its construction interrupted in 1934. Studies and the building of the railways lasted fifty-five years but the rails have never been laid down between Beaumont and Gimont…

Source: https://www.ladepeche.fr/article/2019/01/04/2934738-moissac-cahors-une-ligne-avortee.html

Gabarret-Eauze

…The platform of the section between Gabarret and Eauze was built but the rails have never been laid down...

Source: https://fr.wikipedia.org/wiki/Ligne_de_Langon_%C3%A0_Gabarret

Saint-Girons-Oust

…On August 22, 1881, the extension of the railroad to Oust was declared as being of public interest... The platform and related accessory buildings were finished around 1920 but the rails have never been laid down…

Source: *https://fr.wikipedia.org/wiki/Ligne_de_Boussens_à_Saint-Girons*

Hagetmau-Pau

…An extension of the railroad from Hagetmau to Pau was adjudicated in 1908…the construction begins after the first world war, but have never been achieved…

Source: *https://fr.wikipedia.org/wiki/Ligne_de_Saint-Sever_à_Hagetmau*

Chorges-Barcelonnette

…The construction began in 1909. The first world war slowed down the construction work which ultimately involved German war prisoners. After the war, construction works restarted…however in 1935, the construction was stopped even though part of the platform was already built up to Martinet, 15 km from Barcelonnette…

Source: *https://fr.wikipedia.org/wiki/Ligne_de_Chorges_à_Barcelonnette*

Saint-Juéry-Sainte-Affrique

…Between St-Juéry and St-Affrique, though the construction works progressed slowly, the advent of the first world war would hinder the course of the infrastructure, leading up to the hiring of a foreign workforce, along with German war prisoners between 1915 and 1917…On April 1st 1932, the State adjudicated the project to a company which found it unattractive and got an authorization from the National Economic Council to stop the superstructure work at its charge…

Source: *http://chemins.de.traverses.free.fr/Albi_Ste-Affrique/Saint-Juery_Ste-Affrique.htm*

Foix-Quillan

…The railway Foix-Lavelanet-Quillan has never been built. It was the subject of numerous studies which projected several trajectories with associated regional, local, financial and political controversies. In any case, the delay incurred by the technical studies was never made up and the onset of the war in 1914 precipitated the end of this ambitious endeavor…

Source: *http://chemins.de.traverses.free.fr/lavelanet/Foix_Lavelanet.htm*