INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

548

2020

Railroads, specialization, and population growth in small open economies: Evidence from the First Globalization

Andrés Forero, Francisco A. Gallego, Felipe González y Matías Tapia.

www.economia.puc.cl • ISSN (edición impresa) 0716-7334 • ISSN (edición electrónica) 0717-7593

Railroads, specialization, and population growth in small open economies: Evidence from the First Globalization*

Andrés Forero Francisco A. Gallego Felipe González Matías Tapia†

We explore how railroads affected population growth during the First Globalization (1865-1920) in Chile. We look at areas with strong comparative advantage in agriculture using novel data documenting sixty-years of railroad construction. Using instrumental variables, we present four main findings. First, railroads increased both urban and rural population growth. Second, the impact was stronger in areas with more potential for agricultural expansion. Third, railroads increased specialization in agriculture when combined with a high level of the real exchange rate. And fourth, railroads had little effects on human capital and fertility. These results suggest that the effects of transportation technologies depend on existing macroeconomic conditions.

JEL classification: F14, N16, N76, O1, R1, R4
Keywords: railroads, specialization, agriculture, population

*July 2020. We thank Oded Galor (the editor), two anonymous referees, José Díaz, Jeanne Lafortune, Rolf Lüders, José Tessada, Javier Turén, Alejandro Vicondoa, Gert Wagner, and seminar participants at PUC-Chile, SECHI meetings, and LACEA-LAMES meetings for valuable comments and suggestions. We would like to thank FONDECYT (Project 1170956) for financial support. Rodrigo Icarán, Antonia Paredes, José D. Salas, Alejandro Saenz, Felipe Vial, and Cristine Von Dessauer provided outstanding research assistance.

†Gallego and González: Pontificia Universidad Católica de Chile, Instituto de Economía. Tapia: Banco Central de Chile. Corresponding author: Francisco Gallego, e-mail `fgallego@uc.cl`

1 Introduction

Transportation technologies are one of the main drivers of lower transportation costs and increasing trade in the last centuries (Hummels, 2007). Countries and worldwide institutions still invest an impressive amount of resources in these technologies (World Bank, 2007; Faber, 2012; Donaldson, 2018). Among these, railroads have historically been one of the most important.[1] However, the availability of these transportation technologies can be the consequence of other economic outcomes and, therefore, it is hard to estimate their causal effects.[2] Moreover, their impacts might depend on economic factors and prevailing institutions. For instance, decreases in transportation costs might not have significant effects on trade if the cost of production is high relative to the price of the product. Similarly, although these technologies have the potential to create demographic transitions or structural transformations, the comparative advantage of places might preclude these processes to unfold (Galor and Mountford, 2008; Katz, 2018; Uribe-Castro, 2019).

This paper explores the effect of railroads during the First Globalization in the central and southern parts of Chile, a small open economy with a strong comparative advantage in agriculture. In particular, we provide estimates of the causal impact of railroads on population growth, its composition, specialization patterns, human capital, and fertility. We created a panel of departments observed in the 1865-1920 period to measure the construction of railroads that connected the capital with other cities in the central and southern parts of the country. The geographic area we study has a strong comparative advantage in agriculture in a period of expanded trade opportunities related to the First Globalization. The focus on this area allows us to test whether railroads can trigger a demographic transition or a process of structural transformation in the presence of a comparative advantage in agriculture.

Our empirical strategy uses two sources of exogenous variation for the construction of railroads from the capital to other cities. First, we exploit the speed of the available technology of construction to create an instrumental variable that measures the expected expansion of railroads. In particular, once we know the connecting cities, the first year of construction,

[1]Even today projects like the construction of the "Belt and Road Initiative" implemented by the Chinese government incorporates the construction of railroads (Brakman et al., 2019). Fogel (1962, 1964), Fishlow (1965), and Rostow (1967) were pioneers in studying railroads in economics. Similar technologies include the steamship (North, 1958) and the refrigerator car (Kujovich, 1970).

[2]Fishlow (1965) puts it simply: "whether railroads first set in motion the forces culminating in the economic development of the decade, or whether arising in response to profitable situations, they played a more massive role."

and the available technology – e.g. it is possible to construct 360 kilometers of railroads in 10 years – we construct the number of expected kilometers of railroads each department should have at different points in time. The motivation is that, conditional on department and time fixed effects, the timing of arrival should be exogenous. Second, we follow the existing literature and use "straight lines" variation in which some areas receive railroads because of their location between two connecting cities (Atack et al., 2010; Faber, 2012; Banerjee et al., 2020). We take this idea to panel data, where the instrument only uses straight-lines connecting the capital with the cities targeted by the policy.[3] We assume that the two sources of variation are valid to identify the effects of railroads, conditional on department and time fixed effects. Statistical tests suggest that both instruments contribute in the identification. Moreover, over-identification tests suggest that the instruments are valid.[4]

We present three main results. First, railroads increased both urban and rural population without affecting urbanization rates. Second, we find little evidence of railroads having direct effects on specialization patterns, human capital investments, fertility outcomes, and the share of foreigners. Third, we follow Katz (2018) and using a Duncan index we find that railroads seem to have decreased the skill content of jobs. These results can be interpreted along the lines of the literatures on demographic transition and structural transformation. In particular, this is consistent with the rationale in Galor and Mountford (2008) and with results in Uribe-Castro (2019) for Colombia: the new trade opportunities in a context of a comparative advantage related to agriculture increased the demand for low-skill occupations, which prevented human capital investments and the demographic transition to unfold.

To further understand the mechanisms behind the results, we study the existence of heterogeneous effects. We present two additional results. First, the effect of railroads on population and specialization patterns are stronger in departments with more potential for agricultural expansion.[5] This result is consistent with the idea that railroads allowed areas with comparative advantage in agriculture to expand their production in that sector. Second, we find that the effects of railroads on urbanization and specialization depend upon the level of the real exchange rate. When the relative price of tradable to non-tradable

[3]Katz (2018) also applies the idea of "straight lines" in a panel data context for the U.S. using a different approach, which exploits the appearance of new big (connecting) cities.

[4]Our identification strategy is suited for the center and south parts of the country and it is not valid for the north, reason why we exclude it from the analysis in the paper.

[5]We construct the potential for agricultural expansion using data on agricultural suitability (FAO-GAEZ), also used by others in the literature (e.g., Alesina et al. 2013 and Galor and Ozak 2016).

goods is high, railways construction affect specialization through a substitution effect, with increases employment in agriculture and decreases it in manufacturing and services, which in turns affects urbanization.[6] Given that the economy of a department was open and mainly based on agriculture (a tradable good), this result highlights the role of price mechanisms in determining the impact of transportation technologies.[7] Moreover, we argue that internal migration was an important force behind these effects (Hurtado, 1966; Johnson, 1978).

Our paper makes four contributions. First, we contribute to the economic history of Latin America by measuring the causal impact railroads on several economic and social outcomes.[8] Second, we contribute to the long-standing literature on railroads and economic development (e.g. Fogel 1962; Fishlow 1965; Donaldson 2018; Banerjee et al. 2020). Specifically, our contribution is to analyze a small open economy with comparative advantage in agriculture that faces price shocks. In related work, Fajgelbaum and Redding (2018) study the interaction between railroads, trade openness, and internal geography using a trade model and cross-sectional variation. Our paper complements this literature by using a panel data approach and documenting heterogeneous effects of railroad construction on specialization depending on local endowments and the role of prices, conditional on characteristics that are constant at the department level and common shocks at the macro level.

Third, we contribute to a literature that estimates the effects of railroad construction on the process of urbanization, structural transformation, and the demographic transition. We show that railroads do not necessarily increase urbanization rates, improve human capital, and decrease fertility. We show that the effects of railroads depend upon patterns of comparative advantage, complementing previous literature documenting how trade patterns affect the process of structural transformation (Galor and Mountford, 2008; Uribe-Castro, 2019). Finally, this paper contributes to the literature on the effects of large infrastructure projects on economic outcomes by highlighting the role of time-varying macroeconomic conditions (Aschauer, 1989; Duflo and Pande, 2007; Michaels, 2008; Dinkelman, 2011).

[6]We argue that the real exchange rate was exogenous to the process we study and driven by mining activities in the north (Badia-Miro and Ducoing, 2015; Humud, 1974; Mamalakis, 1976; Meller, 1998).

[7]We implement a battery of exercises to confirm that the mechanism of relative prices is behind the estimated effects and not alternative channels related to other macroeconomic variables.

[8]Other papers studying the causal effects of railroads in Latin America are Miller (1976), Coatsworth (1979), Ramírez (2001), Summerhill (2005), Herranz-Loncán (2014), Zegarra (2011), and Perez (2018).

2 Historical background

This section briefly describes the railroad construction process in Chile between 1850 and 1920 and the associated changes in transportation costs. We also discuss the economic and social context of the country, with a particular focus on the macroeconomic environment, the role of mining and agriculture, and the importance of internal migration.

2.1 Railroad construction

Our focus is on the railroads of the central and south parts of the country.[9] The two largest cities in the central part, Santiago and Valparaiso, were connected through railroads. The idea of constructing a railroad between these two cities was thought by William Wheelwright around 1842. The idea materialized only when the public road between these locations became obsolete, and several interest groups pressed the government to build a railroad (Oppenheimer, 1976). Wheelwright failed in finding investors and farmers exerted pressured on the government to take advantage of the wheat boom in California. As a consequence, the government financed these railroads in a joint venture with private entrepreneurs. The construction began in 1852, it was supposed to take five years, but it was only finished in 1863. Santiago and Valparaiso were finally connected by a 187-kilometers railroad. In 1870 a branch of 49 kilometers extended this railroad to surrounding areas.

The construction of railroads to the south began in 1855, reaching the southernmost city of the mainland in 1913. Railroads first integrated the city of Rancagua and Santiago with an 87-kilometers construction in 1859. A line of 52-kilometers was constructed three years later. The construction process reached the largest city in the south in 1870 by a 588-kilometers railroad. Between 1876 and 1887 the process took a pause due to the Pacific War (1879-1883) and an economic depression. Railroads reached the Arauco region in 1894 and several other branches were constructed in the late 1880s. All in all, Santiago became connected with the south of the country through a 1198 kilometers railroad, and numerous branches integrated the main line with cities by the mountains and the Pacific ocean.[10]

Figure 1 presents the evolution of the railroad network in census years. We can see in this

[9] The first railroads were constructed in the north to fulfill the demand of the mining industry. William Wheelwright, an American entrepreneur, led that process. The construction of these railroads started in 1850 and ended in 1851 and there were 243 kilometers constructed by 1871 (Alliende, 2006, p.14-19).

[10] See Alliende (2006, p.38-72), Thompson and Angerstein (1997, p.76-80), Gross (1998, p.2-9), and Humud (1974) for more details about the construction of railroads in Chile.

figure the halt of construction due to the Pacific War. The last panel depicts the railroad map in 1920. These maps show the differential exposure of areas to railroads in different points in time, an important source of variation we use in the following sections.

Carts and mules were the most common transportation before the arrival of railroads. The freight between Valparaiso and Santiago was done by 30 to 40 ox carts, with a capacity of 40 to 50 quintals, during a period of six days in the summer and 12 days in the winter. The cost of transporting one quintal fluctuated between $1 and $1.75. In contrast, using railroads the freight took only eight hours, with a cost that fluctuated between $0.44 and $0.55. Passenger traffic faced a similar situation: travel time decreased from 14-20 to 6 hours and travel costs decreased from $10-$20 to $2.50-$5 (see Oppenheimer 1976, p. 67-71; Alliende 2006, p. 20, 21, 33 and 38). Railroads also enhanced communication systems with telegraph systems and post offices and decreased other transport-related costs. The inability to transport goods through rivers differs from the case of the U.S.[11] In addition, the cold winter and flooding from rivers made transportation difficult, particularly in the south. Railroads were much less vulnerable to bad weather. Moreover, roads were of low quality and the threat of bandits was severe (Verniory, 2001). All in all, railroads made transportation cheaper, safer, and less likely to affected by weather shocks.

2.2 The economic and social context

Panel A of Table 1 presents the macroeconomic situation of Chile in the period of interest.[12] Annual per capita GDP growth was moderate with rates around 2% the first two censuses that go down to 0.9% in the 1907-1920 period.[13] Inflation was low and the economy became more open, with the ratio of exports to GDP increasing from 13% to 22%. In turn, government spending followed the mining sector, specially nitrates exploitation in the north. The terms of trade experienced an increase in the early twentieth century associated to the nitrate boom, and then a decrease in 1920 after the discovery of synthetic nitrate in Germany

[11]Coastal navigation between cities included in our analyses was rare (Veliz, 1961; Cariola and Sunkel, 1990). See Summerhill (2005) and Coatsworth (1979) for a similar argument in the case of Brazil and Mexico. See McGreevey (1989) for Colombia, a Latin American case closer to the U.S.

[12]This section is based on Humud (1974), Mamalakis (1976), Cariola and Sunkel (1990), Meller (1998), Badia-Miro and Ducoing (2015), and the data come from Diaz et al. (2016).

[13]We observe convergence with respect to the U.S. and a period of faster growth than other countries in Latin America, with the ratio going up from 1.33 to 1.61 (the value of this ratio is 0.43 for the history of Chile). Countries in Latin America include Argentina, Brazil, Mexico, and Peru due to data restrictions.

(Meller, 1998). The real exchange also followed the mining sector. Figure 2 presents the correlation between the real exchange rate and government spending. We argue that the real exchange rate induces movements between the tradable and non-tradable. Panel A of Table 1 also presents the evolution of the price of wheat: there was a decrease between 1885 and 1907 and a subsequent recovery in 1920.

Panel B of Table 1 shows an increase in years of schooling from 3.4 to 6.1 and high values of the Gini index. Regarding population growth, we observe slight declines in the second half of our sample. Behind these numbers, we observe small movements in birth rates (defined as the total number of newborns per 1,000 people) and a small increase in the death rate (defined as total deaths per 1,000 people) from 25 to 31. We also observe an increase in the share of foreigners which is consistent with government policies (González, 2019).

Panel C of Table 1 presents information about the economic structure. The share of GDP in agriculture decreased 18 p.p. between 1865 and 1920, while the share of mining increased by 20 p.p. Manufacturing decreased from 28 to 21% and the service sector increased by about 5 p.p. Despite the increase in mining, the labor share in this sector increased only from 3.4 to only 4.2%. This is a consequence of the low labor intensity of mining and its geographical location in the north (Badia-Miro and Diaz, 2017). In contrast, the decreased in the shares of agriculture and manufacturing in total employment is consistent with their decreases in GDP shares, while the service sector increased employment significantly (explained by increases in the trade, transportation, and personal service sectors). This panel also shows how agriculture went from explaining 30% of exports in 1865 to 10% in 1920.

The main products in agriculture were wheat, barley, pasture, and some fruits and vegetables.[14] Wheat production increased from 2,876,000 quintals in 1877-1878 to 6,207,000 quintals in 1917-18 (Cariola and Sunkel, 1990). Approximately 95% of wheat production was located in the center of the country in 1877-8 but this share decreased to about 65% in 1917-18 as areas located to the south of the Bio-Bio area increased their production.[15] Cariola and Sunkel (1990) document that the main part of wheat production was aimed at exports, initially concentrated in Peru, Australia, and the U.S., and then in a more diversified way, with the U.K. playing a key role. Mamalakis (1976) defines the Great Wheat

[14] Cariola and Sunkel (1990) documents that in 1917, 53% of the agriculture area corresponds to grains (of which 87% corresponds to wheat), 31% to forage, 8% to sylviculture, and the reminder to other crops.

[15] Only 4% of wheat was produced in the north in 1877-8 and about 1% in 1917-8. The south of the Bio-Bio region corresponds to what previous research calls the Frontier (Garcia-Jimeno and Robinson, 2011).

Trade period including the years between 1865 and 1926. In terms of internal consumption, a share of wheat production was also sent to the mining departments of the North but even at its peak it just represented 12% of total production (Cariola and Sunkel, 1990).

Finally, several authors have studied internal migration over this period (Hurtado, 1966; Johnson, 1978). Johnson (1978) documents that 20% of the adult population in 1854 was an internal migrant (i.e. born in other city) and 35% in 1920. This implies that the number of internal migrants increased from 145,912 to 652,791 people in this period. Behind this increase, she identifies three main phenomena. First, migration to areas with expanded agriculture production. Second, migration from small cities to the big cities. Third, internal migration to the mining areas in the north (but this involves fewer people, given the small effects on labor markets). Thus, internal migration was relevant during this period.

3 Data and descriptive evidence

3.1 Data construction

We combined information from historical sources to document the process of railroad construction in Chile. First, we use department level information available in historical population censuses.[16] We were able to construct a panel dataset of 34 departments observed in years 1865, 1875, 1885, 1895, 1907 and 1920 containing information on: (i) urban and rural population, (ii) people working in different economic sectors and occupations, (iii) foreign/national status, (iv) births per department, and (v) proxies of human capital and the supply of schools.[17] Crucially, for each department in our data we added the cumulative number of kilometers of railroads constructed at each census date. We used information and maps from Espinoza (1897) for the the period 1854–1897 and Vasallo and Matus (1943) for the period 1897–1920. We complemented and checked this information with Thompson and Angerstein (1997), Alliende (2006), and several web sources.

We complemented the data with the geographic area of departments, distances to Santiago and the closest railroad, the terrain ruggedness index of Nunn and Puga (2012) con-

[16] Censuses are available at the National Statistics Bureau (INE, web page `www.ine.cl`). We use *department* level data because it is the smallest administrative unit we can construct in panel data form. See Table A.1 for details, including the area of the departments (in km^2) and the populations in 1865 and 1920.

[17] We grouped occupations to make reliable comparisons. Table A.2 presents information for the 15 occupations and an occupational socioeconomic score following Duncan (1961) and Katz (2018).

structed with data from the U.S. Geological Survey, and an index of wheat suitability using data from the Global Agro-Ecological Zones (FAO-GAEZ). We also constructed an index of potential for agriculture expansion using the residual of a regression of the share of agriculture in 1865 in each department on wheat suitability, the share of literate people, urbanization, population density, and distance to Santiago. This variable measures the potential for agricultural expansion at the beginning of our sample and it is defined as the negative of the residual of the previous regression. Figure 3 presents the correlation between this measure and the share of agriculture in 1854; Table A.3 presents regression results. The figure shows both the heterogeneity in terms of agriculture share in the 1854 period and the opportunities for expansion, given the suitability for agriculture of different departments. Finally, we use macroeconomic variables from Diaz et al. (2016).

3.2 Descriptive statistics

Table 2 presents descriptive statistics for our sample. While Valparaiso was the first department to receive railways, by 1865 railways in our sample were just starting to be constructed and the average department had just 9.5 kilometers of railroads. It is easy to see the pause in construction during the Pacific War (1879-1883). The average urbanization rate increased from 23% to about 35% in 1920 and the average department more than doubled its population, from about 45,500 to more than 92,000 people. The Herfindahl index of concentration increases initially but then decreases in the twentieth century. This is a consequence of a decrease in the share of manufacturing and an increase in services without a significant change in the share of labor in agriculture. These trends are consistent with the macro evidence presented in Table 1.[18] In terms of human capital outcomes, literacy rates increased significantly and there is an expansion in schools per people. In contrast, the Duncan index remains similar because the sectoral composition does not change towards sectors which demand more human capital.

As a benchmark and preliminary evidence, Table 3 presents ordinary least squares (OLS) estimations of regressions of the following form:

$$y_{dt} = \alpha_d + \lambda_t + \beta \log r_{dt} + \varepsilon_{dt} \tag{1}$$

[18]Note that information for occupational variables is missing in the 1885 census.

where y_{dt} is a measure of the population living in department d at year t, r_{dt} is kilometers of railroads constructed until year t, and ε_{dt} is a random shock with a mean of zero.

Column 1 of Table 3 presents a strong, positive, and statistically significant correlation between the logarithm of railways constructed and the logarithm of the number of people living in urban areas, controlling for department and year fixed effects (α_d, λ_t). The estimated coefficients are elasticities. A coefficient of 0.225 implies that a 100% increase in railways (a 41 kilometer construction) is associated with an about 6,000 people living in urban areas (a 22.5% increase).[19] This coefficient decreases significantly to about 0.10 when we add year fixed effects in column 2. This means that 41 more kilometers of railways is associated with about 2,700 more people living in the department. This big decrease in the correlation, when adding time effects, implies that there are country-wide trends that explain a big part of the correlation between urban population and railroad construction. We then look at the log of rural population in columns 3-4. Results are similar to the impacts on urban population: the estimated coefficient of 0.104 compares with a coefficient of 0.099 for the urban population. Columns 5-6 document the correlation between railroads and the share of urban population. As expected, we find that this variable is weakly correlated with railroads. Finally, we find that total population is correlated with railroads with a coefficient of 0.088. In sum, these results suggest that railroads are positively correlated with population, without a significant differential effect on rural and urban population.

Table 4 presents robustness exercises. Columns 1 and 3 exclude the three largest departments in terms of population (Arauco, Valparaiso, and Santiago). The estimated coefficient for urban population does not change. In the case of rural population (column 3), the estimated coefficient decreases to 0.083 but it is still statistically significant. Columns 2 and 4 exclude the 1885 cross section and the correlation is similar.[20] In sum, these preliminary results suggest that there is a robust correlation between railroads and population size. The fact that adding time effects decreases the correlation suggest that there could be omitted variables both at the department and time level that explain the correlation we find. Next we propose an identification strategy to estimate the causal effects of railroads.

[19]The average department during the entire period had 26,618 people living in urban areas. The estimated coefficient of 0.225 implies an increase of $0.225\times25{,}618=5{,}975$.

[20]It is important check this for two reasons: (i) this year does not add much statistical information because railroad construction was delayed due to the Pacific War, and (ii) this is the only year for which we do not have information on occupations, so it is useful to present the correlation for later comparisons.

4 Empirical strategy

This section presents our strategy to estimate the causal effect of railroads. First, we present two sets of instruments and the corresponding identification assumptions associated to them. Then, we provide some evidence to support our econometric strategy.

4.1 Instrumental variables

The first instrumental variable we propose relies on the fact that, once it has been decided to construct a certain railroad from Santiago to other cities, we should expect the railroad to be constructed in a certain amount of years, given the technology available. Thus, the railroad should be finished in a known period of time. This time-variation in the railroad's arrival is plausibly exogenous, especially considering that we present regressions with department and year fixed effects. In contrast, *real* construction could vary endogenously. The information in Oppenheimer (1976) suggests that the available technology in the 1850-1875 period in the railroad industry allowed it to build approximately 36 kilometers annually. This means that we should expect a construction of 360 kilometers in a 10-year period. Then, using this idea and the fact that construction started from Santiago in 1852, we can construct the expected construction in each department in the subsequent periods.

Figure 4 presents the expected and actual construction for all department, ordered in the y-axis by their distance to Santiago. To construct expected construction we first identified departments with some railroads. Then, we notice that in order to get to departments located farthest from Santiago, railroads had to be constructed in departments closer to Santiago.[21] Thus, we start "constructing" railways in departments closer to Santiago and then we move to other departments at the available speed of construction.[22] The key identification assumption is that, although the construction of railroads in department i and not department j is indeed endogenous, once we control for fixed effects the predicted railroads constructed at time t given the technology in each department is plausibly exogenous and, therefore, affects population variables only through *effective* construction. Panel (a)

[21]For instance, looking at panels (a) and (b) in Figure 4, we note that in order to get to, for instance, the Concepcion department, railroads had to be constructed before in departments located closer to Santiago. Then, our predicted construction for Concepcion in 1865 is zero and only becomes positive in 1875.

[22]This idea implies that the timing of the treatment and our instrumental variable is correlated with distance to Santiago, which also applies to our second instrument below. However, Table A.9 shows that all results are robust to include an interaction of the treatment with distance to Santiago as control.

in Figure 4 shows how in 1865 several departments have already reached their expected construction. However, other departments were "late adopters" despite the fact that the available technology would have allowed them to have railroads. Panel (b) shows that in 1875 most of these "delayed" departments catch-up with expected construction, but other departments started new construction, and thus some others were delayed. By 1920 most departments have actual construction equal to expected construction.

The second instrument is based on previous literature that uses "straight lines" (Atack et al., 2010; Banerjee et al., 2020; Faber, 2012). This identification strategy exploits the fact that railroads usually attempt to connect large cities. Therefore, if place i happens to be in the middle of two big cities, when we draw a theoretical straight railroad line between the two big cities – the cost-effective construction – place i is likely to get railroads, although it was not the main purpose of the policy. Estimation then omits the largest cities from the data and estimate two-stage least squares using the straight line as instrument for effective railroad construction. To the best of our knowledge, Katz (2018) and this paper are the first to use "straight lines" in panel data. While Katz (2018) exploits the appearance of new big cities, we consider straight lines between Santiago and the three main cities to be connected according to the objectives of the policy.[23] We use these lines in panel data and, therefore, our instrument is the actual construction times a dummy taking a value of one if any of our straight lines passed through the department.[24]

Figure 5 presents a map with the straight lines, the centroid of each department, and the shortest line between them. There is significant variation in terms of treatment associated with being located closest to the lines connecting the big cities defined by the policy. The key identification assumption is that, although the construction of railroads in department d at time t is probably endogenous, the *interaction* between actual construction and the dummy for being located on the straight line, once we control for fixed effects, is exogenous. More precisely, we exploit the same variation as the one used in Banerjee et al. (2020) and Atack et al. (2010), in the sense that at any moment having or not railroads depends on

[23] In particular, we consider one straight line from Santiago to Valparaiso and another from Santiago to Puerto Montt. Then, using the line from Santiago to Puerto Montt, we add another segment, defined as the shortest straight line between Concepcion and the straight line between Santiago and Puerto Montt.

[24] As a robustness check, we also use a measure of the distance from the centroid of the department to the closest straight line and normalize the distance to make the distance to the straight line comparable to the dummy, with a value of one when the centroid of the department is on the straight line and a value of zero when the distance is equal to the maximum distance observed in the data.

being in the straight line across connecting cities. The only difference is that in our case, the timing of the construction also matters. This may produce a secondary threat to the identification if the intensity of construction at time t in departments located in the straight lines is correlated with time-varying shocks in those periods in these departments. Although we cannot test this assumption, over-identification tests on the two instruments make us confident that this potential threat is unlikely to be relevant.

4.2 Empirical validation

Table 5 presents the relationships between railroad construction and the two instruments, first stages in our instrumental variables approach. We present results using the two instruments separately (in columns 1 and 2) and jointly (in column 3). These results are important because it could be the case that both instruments capture the same type variation and, thus, we do not add new information to the estimates. Importantly, coefficients are interpreted as elasticities from the instruments to the endogenous variable.

Both instruments are highly predicted of railroads. The theoretical construction of railroads is positively correlated with actual construction, with a coefficient of 0.73 (column 1, t-stat of 10.02). Similarly, straight lines are also positively correlated with a coefficient of 0.85 (column2 , t-stat of 21.62). Column 3 combines both instruments and the two coefficients decrease, but they still are individually and jointly significant with an F-stat of 121.06. These results are important because they suggest that both instruments, while positively correlated, capture different sources of variation. Columns 4-6 present robustness exercises. Columns 4-5 exclude departments located in connecting and big cities respectively and results remain similar. Column 6 defines the dependent variable considering just the intensive margin, i.e. a dummy that takes a value of one if the department has some railroads. Results indicate that both variables have statistical power. We conclude that the instruments are strongly correlated with the endogenous variable.

We also implemented two falsification exercises to test whether the instruments are correlated with variables that could affect outcomes *before* the railroads were constructed. We focus on population density in 1854 and our index of wheat suitability, both closely correlated with outcomes that we study. We define the cumulative kilometers of railroads constructed and the two instruments for the entire period. Then we estimate cross-sectional regressions of the two outcomes on our instruments. Figure 6 presents value added plots for each instrument and dependent variable, and Table A.4 presents regression results. Reassuringly the instruments are uncorrelated with these important initial conditions. These results provide some support to our identification assumption. We complement the analyses with over-identification tests.

5 Railroads and population growth

This section discusses results using our instrumental variables approach. Table 6 uses urban population, rural population, the share of urban population, and total population as dependent variables and present estimates for each instrumental variable separately and then both together. We also present the Kleibergen-Paap under-identification test and the Sargan over-identification test. The latter is particularly important to study the validity of our strategy when we using the two instruments together. We adjust standard errors adjusted for spatial autocorrelation using a 200 kilometers cut-off in distance (Conley, 1999).[25]

5.1 Urban and rural population

Columns 1-3 in Table 6 studies urban population. Column 1 uses straight lines as instrument. The coefficient on railroads is 0.14 (s.e. 0.04). The magnitude is somewhat bigger than the one presented in Table 3. Column 2 uses predicted railroads as and the coefficient barely changes to an estimate of 0.16 (s.e. 0.04). When we combine both instruments, the estimated coefficient is 0.14 (s.e. 0.04). The Sargan test provides support for the validity of instruments and the Kleibergen-Paap test comfortably rejects the null hypothesis of under-identification. We conclude that there is a significant causal effect of railroad construction on urban population. The estimated coefficient of implies that a 100% increase in railways, i.e. a 41 kilometer construction, causes a 14% increase in urban population, which represents about a 3,700 people increase in urban areas.

Columns 4-6 present the same exercises but using rural population as the left-hand side variable. In this case, we also document a significant positive effect of railroads with some small differences for different specifications. We again pass the under-identification and over-identification tests comfortably. Taking the results of the over-identified model in column 6,

[25] Results are also robust to using other distance cutoffs. Given that all results are robust to the use of spatial correlation, Table 6 is the only one reporting Conley standard errors.

the estimated coefficient is 0.15 (s.e. 0.03). This implies that a 100% increase in railways, i.e. a 41 kilometer construction, causes a 15% increase in urban population, which represents about a 5,800 increase in people living in rural areas.

The remaining columns present estimates for urbanization rates (columns 7-9) and total population (columns 10-12). As expected, we find a null effect on urbanization rates and a positive effect on total population. If we take results from column 12, we find an estimate of 0.13 (s.e. 0.03). In sum, results in Table 6 suggest a positive causal effect of railroads on both rural and urban population, without a significant change in urbanization rates.

Table A.5 presents results of three different robustness checks to our estimates: excluding the three largest departments, excluding departments with connecting cities, and using an alternative straight lines instrument which uses distance to the closest straight line instead of a dummy variable. These are important checks and, moreover, excluding connecting cities is also relevant for our second instrument (Banerjee et al., 2020). The estimates are robust to all of these changes. Moreover, both under- and over-identification tests are again passed.[26]

In order to examine the robustness of the results, Figure 7 displays the distribution of point estimates for the effect of railroads on urban, rural, and total population from 34 different regressions for each population variable with a different department excluded from the sample. The estimated coefficients on railroads are robust to excluding single departments from the sample. The coefficients for the regressions of total population vary from 0.100 to 0.143, urban population from 0.117 to 0.155, and rural population from 0.117 to 0.161. All these changes are within the standard errors estimated in our main regressions. Thus, this evidence suggests that our main results are not explained by individual departments.

In sum, results in this section suggest that there is a positive and economically significant effect of railroads on population, implying similar effects on both urban and rural areas of the departments. The evidence of railroads increasing *both* urban and rural population is interesting and novel in the literature. These results imply that railroads do not necessarily increase urbanization rates. We argue that a change in urbanization depends on the context and the comparative advantage of the affected areas.

[26]Table A.6 studies whether the effects of railroads are driven by the extensive or intensive margin. Using the two instruments we construct "extensive margin" instruments and run regressions of population on our main measure of railroads penetration in a department and a dummy of whether the department has any railroads using four instrumental variables. Results suggest that the intensive margin is more important.

5.2 The role of time-varying factors

The impact of railroads might differ depending on (i) macroeconomic conditions that affect production, and (ii) in the short- versus the long-run. This section estimates additional panel data models to study whether effects are likely to depend on time-varying factors: random effects, first-differences, and long-differences. Results are presented in Table 7.

Columns 1, 4, 7, and 10 present results of random effects models. Results are similar to the ones using fixed effects models in Table 6, suggesting that measurement-error is probably unlikely to be relevant. Next, columns 2, 5, 8, and 11 display first-differences estimates. Estimates are still statistically significant but smaller in magnitude than our fixed- and random-effects estimators. The remaining columns present long-difference estimates. These estimates are slightly bigger than our fixed effects estimates, suggesting that that the long-run effects are similar or even larger than the short-term effects. Taken together, long- and first-differences estimates differ from fixed effects estimates, which suggests that the effects of railroads are likely to vary over time.

In sum, results in this subsection suggest that railroads might have different effects on population growth depending on time-varying factors. The next section studies the role of macroeconomic conditions as measured by the exchange rate.

6 Specialization, human capital, and fertility

This section explores potential mechanisms underlying our estimated impacts of railroad construction on population size. The motivation comes from two literatures. First, the trade literature in which railroad construction operates as a reduction in transport costs in the context of a small open economy. This reduction reduces trade barriers and allows regions to specialize in goods in which they have a comparative advantage, with labor mobility playing a crucial role. This mechanism implies that the effect of railroads depends on the real exchange rate. In particular, on relative prices at the time in which railroads arrived. If the price of tradable goods is relatively high, relatively more workers will move to agriculture.[27] Second, the literatures on demographic transitions and structural transformations, which suggest

[27] These ideas follow from models in Desmet and Rossi-Hansberg (2014), Rossi-Hansberg and Wright (2007), Fajgelbaum and Redding (2018), Adamopoulos (2011), Vandenbroucke (2008), Herrendorf et al. (2008), Caselli and Coleman (2001), Donaldson (2018), and Donaldson and Hornbeck (2016), among others.

that the effect of additional trade opportunities on human capital accumulation and fertility outcomes depend on the comparative advantage of different places (Galor and Mountford, 2008; Katz, 2018; Uribe-Castro, 2019). This mechanism implies that railroads should affect population as well as human capital accumulation and fertility patterns.

We extend the previous analyses in two dimensions to study potential mechanisms. First, we study heterogeneous effects of railroads on all outcomes along two dimensions: potential for agriculture expansion and the the real exchange rate level in year t. Second, we add dependent variables to our population and urbanization variables. Some of them are related to economic mechanisms such as the share of labor in agriculture, manufacturing, and services, and a Herfindahl index of labor concentration. Others are related to fertility rates, human capital levels and accumulation, and the share of the foreign composition of the population.

6.1 The importance of comparative advantage

Motivated by the trade literature we begin by studying potential heterogeneous effects of railroads. Our argument implies that areas with more potential for agricultural expansion should be more affected by railroads. In turn, changes in the real exchange rate should be particularly important. Figure 2 plots fluctuations in the real exchange rate from 1865 to 1920. If a department receives railroads in 1865 with a high real exchange rate, the effects should be significantly different on a number of economic outcomes than if the department receives railroads in 1875 when the real exchange rate was low.

Table 8 presents estimates using both instruments.[28] Columns 1-2 present the effects on urban and rural population. The direct impacts of railroads are positive and economically significant, consistent with our previous results. The interactions of railroads with potential for agriculture expansion are also significant and positive: railroads increased population more in regions with more potential for agriculture expansion. The interaction of the real exchange rate with railroads is negative for urban population and zero for rural population.

Figure 8 helps to clarify the economic significance of the effects for different departments and different real exchange rates. We present the effects for values of the percentiles 25, 50, and 75 of both agricultural expansion and the real exchange rate. Areas with potential

[28] Note that while our variable for agriculture expansion has a mean of zero (as it is constructed using the residual of a regression), we demeaned the log of the real exchange rate in the regressions and, therefore, the main effect of railroads represents the impact of this variable at the average real exchange rate.

for expansion located in the percentile 75 of the distribution experienced increases in population in both rural and urban areas that were much bigger than the average effects. In turn, the interaction effects for the real exchange rate imply that while the effects of railroads on rural population is similar by the level of the exchange rate, the effect of railroads on the urban population are affected by the real exchange rate, as the urban population increases significantly less when the railroads arrival takes place in periods when they real exchange is relatively high. These patterns imply the results presented in column 3 of Table 8 which shows little interaction effects for the potential of agriculture expansion but a negative interaction effect of railroads and the real exchange rate.

Columns 4-8 in Table 8 help to understand these results. Column 4 presents estimates for a Herfindahl-Hirschman index of employment concentration. While the main effect is not different from zero, results show positive and statistically significant effects of the interactions of railroads with the potential for agriculture expansion and with the real exchange rate, implying an increase in specialization. The higher specialization is driven by a significant and positive interaction effects on the share of agriculture (for both agriculture expansion and the real exchange rate), which is compensated by negative interaction effects of the real exchange rate with railroad construction in the manufacturing and services shares. This is expected given the insights of the trade models discussed above. These results help to understand the interaction effects on urban and rural population, as periods with high exchange rates affect movements of urban population.

These estimates are subject to some challenges. The most relevant is related to the interaction effects of the real exchange rate and railroad construction. As we discussed before, the real exchange rate is probably a consequence of the macroeconomic equilibrium of the Chilean economy in this period. Therefore, the interaction effects may be a consequence of other macroeconomic shocks. Table A.7 explores the robustness of results to other macroeconomic channels. We consider the potential influence of GDP growth, GDP growth in mining, and government expenditure as a percentage of GDP. In all cases we add an additional interaction of railroads and these other macroeconomic variables and compare estimates for the interaction of railroads and the real exchange rate with the ones reported in Table 8. While in some cases the new interactions are statistically significant (e.g. column 1 in panel A) and some of the point estimates change, the main conclusions derived from Table 8 remain: the interaction of the real exchange rate and railroads is negative for urban population, negative for urbanization, positive for the Herfindahl index, positive for the share of labor in agriculture, negative for the share of labor in industry, and negative for services.[29]

Finally, to further understand results, we considered two additional dimensions that may be relevant: distance to Santiago and terrain ruggedness.[30] Results are presented in Table A.9. First, as expected departments located far from Santiago benefited more in terms of population growth without a discernible effect on urbanization. Interestingly, it seems that the arrival of the railroad helped more the development of manufacturing than agriculture, probably because these departments were part of the frontier and had low levels of development. Second, regarding terrain ruggedness, we do not observe that there are significant heterogeneous effects of this variable with railroads, probably capturing the fact that in Chile the departments with higher values in the ruggedness index are located in average in the central areas, which were more populated and developed initially.

In sum, results in this section suggest that the effect of railroads on the economy are mediated by comparative advantage considerations, related to both endowments and to changes in price conditions, as suggested by a simple rationale derived from trade models. This explains why we find, in contrast to the previous literature, that railroads do not necessarily increase urbanization and even the opposite during periods of favorable price conditions. Our results also suggest a key role for the expansion of the size of the urban population and internal migration within the country. Unfortunately we lack data to test the role of internal migration but Johnson (1978) documents a significant role of movements within the country to explain population growth in different departments.

6.2 Human capital and fertility

As suggested by our theoretical rationale, all previous results should have implications for patterns of human capital accumulation and fertility outcomes. Table 9 studies whether this is the case. We again estimate the direct effect of railroads and heterogeneous effects related to agricultural expansion and the real exchange rate. Column 1 presents results for the Duncan index, a proxy for the skill intensity in different departments. We observe

[29]Table A.8 presents results using the interaction of the real price of wheat instead of the real exchange rate. The sign and magnitude of the interactions is consistent with results in Table 8, confirming that the heterogeneous effects of railroad with the real exchange rate are related to production incentives.

[30]Our proxy for the potential of agricultural growth is orthogonal to distance to Santiago because it is the negative of a residual of a regression including distance to Santiago. The same applies to initial population.

little relationship with railroads but a negative interaction with the potential for agriculture development and with the real exchange rate. This is consistent with our previous results in Table 8. Column 2 shows small direct and heterogeneous effects of railroads on literacy rates. Together, these results suggest that railroads decreased the demand for skill in areas and moments where the comparative advantage of agriculture production was more relevant. Moreover, we see little movements of less skilled workers to these areas in the short-term.

Columns 3-6 study proxies for human capital accumulation. Results in column 3 show that schools per capita have were not affected by railroads (which implies that schools grow at the same rate as the population). In turn, railroads had a negative impact in areas with more potential for agriculture. Note that results in columns 4-5 show that while the share of teachers in the population is not affected by railroads, the share of students decreases in areas with more agriculture potential when railroads arrive. The latter result may be a combination of changes in the supply and demand for education. We cannot disentangle between the two but the pattern is consistent with a model in which the increase in the demand for less-intensive occupations decreases human capital accumulation in equilibrium. Column 6 studies if railroads affected foreign migration. This is important as railroads were contemporaneous with an increase in foreign (skilled) immigration flows to some areas (González, 2019). However, the share of foreigners was not especially affected by railroads, but there is a negative interaction of railroads and the real exchange rate. This result is consistent with our previous evidence in that the flows of skilled immigrants grew at a lower rate when there were incentives to produce less skill-intensive goods (i.e. in agriculture).

Columns 7-8 study the effects of railroads on fertility outcomes. As previously discussed, the fact that railroads do not seem to have produced significant effects on human capital outcomes and seem to have produced a decrease in the demand for skills in some areas should imply that we do not observe a demographic transition in these areas (with decreases in fertility rates). Column 7 presents the effects on the ratio of births per woman. Results show that railroads did not affect this dimension. Finally, in order to measure the impact of railroads on the *quality* dimension of fertility, we collected data on the share of children born out-of-the-wedlock. This phenomenon is particularly interesting in the history of Chile and captures important features of the marriage market that affect the quality of life of families and newborns (Díaz et al., 2016). Results indicate zero effects of railroads on this dimension, suggesting that the shock associated to railroad expansion did not affect fertility patterns in terms of the type of family where the kids were born.

All in all, results in this section suggest that the expansion of railroads did not increase human capital accumulation and did not affect fertility patterns. This is consistent with the fact that railroads were specially important for the development of low-skill activities such as agriculture and, therefore, did not lead to a process of structural transformation.[31]

7 Conclusion

Societies are often confronted with understanding the economic and social consequences of the construction of transportation infrastructure. Our paper examines the case of the central and south areas of Chile during the First Globalization. We constructed a panel data set using historical sources to document a sixty-year process of railroad construction and designed an identification strategy using exogenous variation in an instrumental variable framework. We complement the literature with a new source of exogenous variation related to speed of construction given by the existing construction technology, which enable us to predict the timing of infrastructure availability. Deviations from this prediction are endogenous, but we overcome this endogeneity by instrumenting effective with predicted construction.

We find that railroads caused significant increases in urban, rural, and total population without affecting urbanization rates. We argue that there is a complementarity between railroads and competitive advantages related to agriculture motivated by a simple rationale based on trade models. It is worth noting that Chile is a small open economy and the regions we study have comparative advantage in agriculture. We test these ideas in two dimensions: whether areas with more potential for agriculture are more affected by the railroad shock and whether the effect of railroad construction is stronger when the prices of tradable goods are higher (proxied by the real exchange rate). We find support for both hypotheses. To further understand the mechanisms behind these effects, we extended our analyses and estimate the effects on economic specialization and sectoral composition. We further show that the above-mentioned effects are accompanied by increases in specialization, mostly driven by movements of employment to agriculture. We also argue that these patterns reflect important flows of internal migration from other regions. Thus, our results show, in

[31] We also studied whether railroads affected the size of the transport and communication sector, and state presence. Any of them can be interpreted as an alternative explanation for the potential effects of railroads on population and other variables. We do not find effects of railroads on these dimensions (Table A.10).

contrast to most of previous work, that railroad construction does not necessarily imply an increase in urbanization rates but this depends on the pattern of comparative advantage and how this complements with the decrease in transportation costs.

Finally, we find that railroad construction did not increase human capital accumulation and did not affect fertility outcomes. This is consistent with the fact that railroads decreased the demand for skills in the areas and moments where there were more incentives to move to agriculture production. Putting it differently, as railroads did not produce a process of structural transformation, then we do not observe that they moved the economy to a demographic transition with lower fertility rates and more human capital accumulation. We believe that the findings in this study contribute to our understanding about the role of transportation on economic and social development. We also hope that they shed some light on the mechanisms.

References

Adamopoulos, T. (2011). Transportation costs, agricultural productivity, and cross-country income differences. *International Economic Review*, 52(2):489–521.

Alesina, A., Giuliano, P., and Nunn, N. (2013). On the origins of gender roles: Women and the plough. *Quarterly Journal of Economics*, 128(2):469–530.

Alliende, M. P. (2006). *Historia del Ferrocarril en Chile*. Goethe Institut-Pehuén Editores.

Aschauer, D. A. (1989). Is public expenditure productive? *Journal of Monetary Economics*, 23(2):177–200.

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

Badia-Miro, M. and Diaz, J. (2017). The impact of nitrates on the chilean economy, 1880–1930. *Palgrave Studies in Economic History: The First Export Era Revisited Reassessing its Contribution to Latin American Economies*.

Badia-Miro, M. and Ducoing, C. (2015). The long run development of chile and the natural resources curse. linkages, policy and growth, 1850-1950. In Marc Badia-Miro, V. P. and

Willebald, H., editors, *Natural Resources and Economic Growth. Learning from History.* Routledge.

Banerjee, A., Duflo, E., and Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in china. *Journal of Development Economics.*

Brakman, S., Frankopan, P., Garretsen, H., and Marrewijk, C. V. (2019). The new silk roads: an introduction to china's belt and road initiative. *Cambridge Journal of Regions, Economy and Society.*

Cariola, C. and Sunkel, O. (1990). *Un Siglo de Historia Económica de Chile: 1830-1930.* Santiago de Chile, Universitaria.

Caselli, F. and Coleman, W. J. (2001). The U.S. structural transformation and regional convergence: A reinterpretation. *Journal of Political Economy*, 109(3):584–616.

Coatsworth, J. H. (1979). Indispensable railroads in a backward economy: The case of Mexico. *Journal of Economic History*, 39(4):939–960.

Conley, T. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45.

Desmet, K. and Rossi-Hansberg, E. (2014). Spatial development. *American Economic Review*, 104(4):1211–43.

Díaz, J., Gallego, F., and Lafortune, J. (2016). Nacimientos fuera del matrimonio en la historia de Chile: algunos hechos estilizados. *Estudios Publicos.*

Diaz, J., Luders, R., and Wagner, G. (2016). *Chile 1810-2010. La Republica en Cifras. Historical Statistics.* Ediciones Universidad Catolica de Chile.

Dinkelman, T. (2011). The effects of rural electrification on employment: New evidence from South Africa. *American Economic Review*, 101(7):3078–3108.

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934.

Donaldson, D. and Hornbeck, R. (2016). Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2):799–858.

Duflo, E. and Pande, R. (2007). Dams. *Quarterly Journal of Economics.*

Duncan, O. D. (1961). *A Socioeconomic Index for All Occupations.* New York: Free Press of Glencoe.

Espinoza, E. (1897). *Jeografía Descriptiva de la República de Chile.* Imprenta y Encuadernación Barcelona.

Faber, B. (2012). *Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System.* PhD thesis, London School of Economics.

Fajgelbaum, P. and Redding, S. (2018). Trade, structural transformation and development: Evidence from Argentina 1869-1914. *NBER Working Paper.*

Fishlow, A. (1965). *American Railroads and the Transformation of the Antebellum Economy.* Harvard University Press.

Fogel, R. (1962). A quantitative approach to the study of railroads in american economic growth: A report of some preliminary findings. *Journal of Economic History*, 22(2):163–197.

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

Galor, O. and Mountford, A. (2008). Trading population for productivity: Theory and evidence. *The Review of Economic Studies*, 75(4):1143–1179.

Galor, O. and Ozak, O. (2016). The agricultural origins of time preference. *American Economic Review*, 106(10):3064–3103.

Garcia-Jimeno, C. and Robinson, J. (2011). The myth of the frontier. *Understanding Long-Run Economic Growth: Geography, Institutions, and the Knowledge Economy.*

González, F. (2019). Immigration and human capital: Consequences of a nineteenth century settlement policy. *Cliometrica.*

Gross, P. (1998). Desarrollo urbano y ferrocarril del sur: 1860–1960. impacto en ciudades y pueblos de la red. *Instituto de Estudios Urbanos, Pontificia Universidad Católica de Chile.*

Herranz-Loncán, A. (2014). Transport technology and economic expansion: The growth contribution of railways in Latin America before 1914. *Journal of Iberian and Latin American Economic History*, 32(1):13–45.

Herrendorf, B., Rogerson, R., and Valentinyi, Á. (2008). Growth and structural transformation. *NBER Working Paper*.

Hummels, D. (2007). Transportation costs and international trade in the second era of globalization. *Journal of Economic Perspectives*, 21(3):131–154.

Humud, C. (1974). Política económica Chilena desde 1830 a 1930. *Estudios de Economía*.

Hurtado, C. (1966). *Concentración de poblacion y desarrollo economico: el caso chileno*. Universidad de Chile.

Johnson, A. L. (1978). *Internal Migration in Chile to 1920: Its Relationship to the Labor Market, Agricultural Growth, and Urbanization.* PhD Thesis. University of California, Davis.

Katz, O. (2018). Railroads, economic development, and the demographic transition in the United States.

Kujovich, M. Y. (1970). The refrigerator car and the growth of the American dressed beef industry. *Business History Review*, 44(4):460–482.

Mamalakis, M. (1976). *The Growth and Structure of the Chilean Economy: From independence to Allende*. Yale University Press.

McGreevey, W. P. (1989). *Historia Económica de Colombia, 1845-1930*. Bogotá, Tercer Mundo Editores.

Meller, P. (1998). *Un Siglo de Economía Política Chilena (1890-1990)*. Editorial Andrés Bello.

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

Miller, R. (1976). *Social and Economic Change in Modern Peru*, chapter Railways and Economic Development in Central Peru, 1890-1930. Center for Latin American Studies, The University of Liverpool.

North, D. (1958). Ocean freight rates and economic development 1750-1913. *Journal of Economic History*, 18(4):537–555.

Nunn, N. and Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *The Review of Economics and Statistics*.

Oppenheimer, R. (1976). *Chilean Transportation Development: The Railroad and Socio-Economic Change in the Central Valley, 1840-1885*. PhD thesis, University of California, Los Angeles.

Perez, S. (2018). Railroads and the rural to urban transition: Evidence from 19th-century Argentina.

Ramírez, M. T. (2001). Los ferrocarriles y su impacto sobre la economía colombiana. *Journal of Iberian and Latin American Economic History*, 19(1):81–122.

Rossi-Hansberg, E. and Wright, M. L. J. (2007). Establishment size dynamics in the aggregate economy. *American Economic Review*, 97(5):1639–1666.

Rostow, W. W. (1967). *Las Etapas del Crecimiento Económico*. F.C.E.

Summerhill, W. R. (2005). Big social savings in a small laggard economy: Railroad-led growth in Brazil. *Journal of Economic History*, 65(1):72–102.

Thompson, I. and Angerstein, D. (1997). *Historia del Ferrocarril en Chile*. Dirección de Bibliotecas, Archivos y Museos.

Uribe-Castro, M. (2019). Caffeinated development: Exports, human capital, and structural transformation in Colombia.

Vandenbroucke, G. (2008). The U.S. western expansion. *International Economic Review*, 49(1):81–110.

Vasallo, E. and Matus, C. (1943). *Historia de los Ferrocarriles en Chile*. Editorial Rumbo.

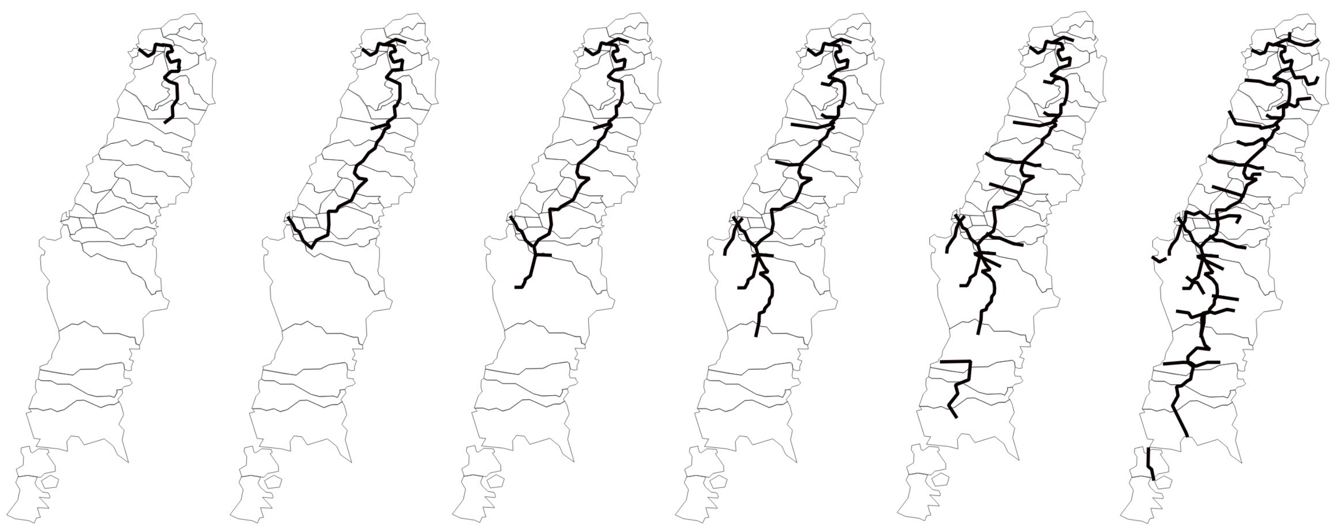
Veliz, C. (1961). *Historia de la marina mercante de Chile.* Editorial Universitaria.

Verniory, G. (2001). *Diez años en Araucanía, 1889-1899.* Pehuén Editores.

World Bank (2007). Evaluation of world bank support to transportation infrastructure. *Washington DC, World Bank Publications.*

Zegarra, L. F. (2011). Railroads in Peru: How important were they. *Revista Desarrollo y Sociedad*, 68.

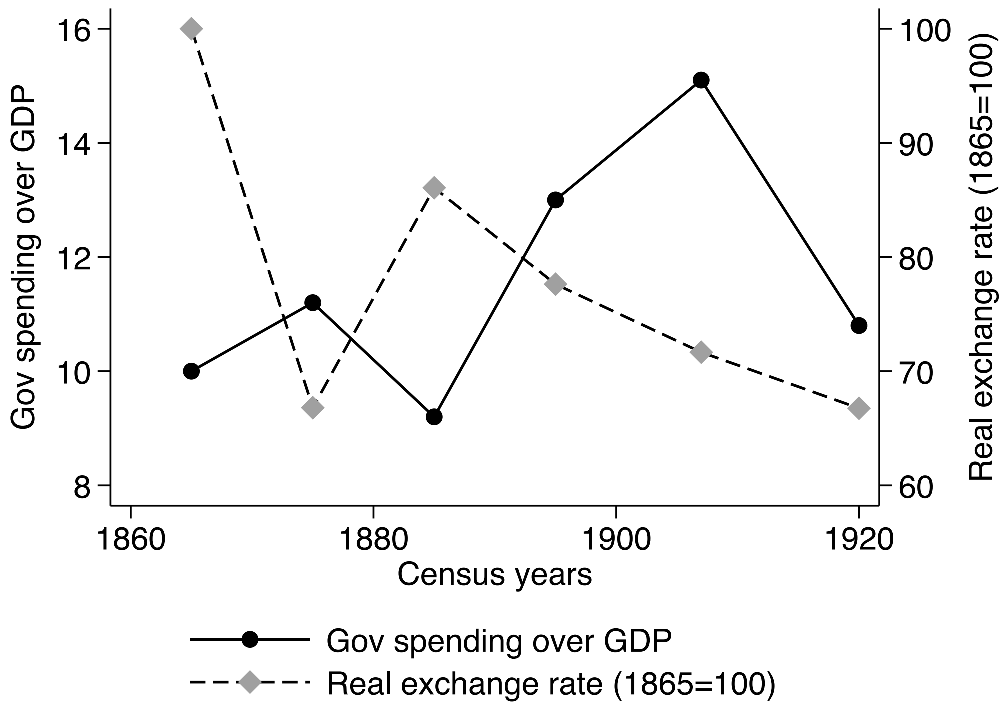
Figure 1: Railroad construction, 1865 – 1920



(a) 1865 (b) 1875 (c) 1885 (d) 1897 (e) 1907 (f) 1920

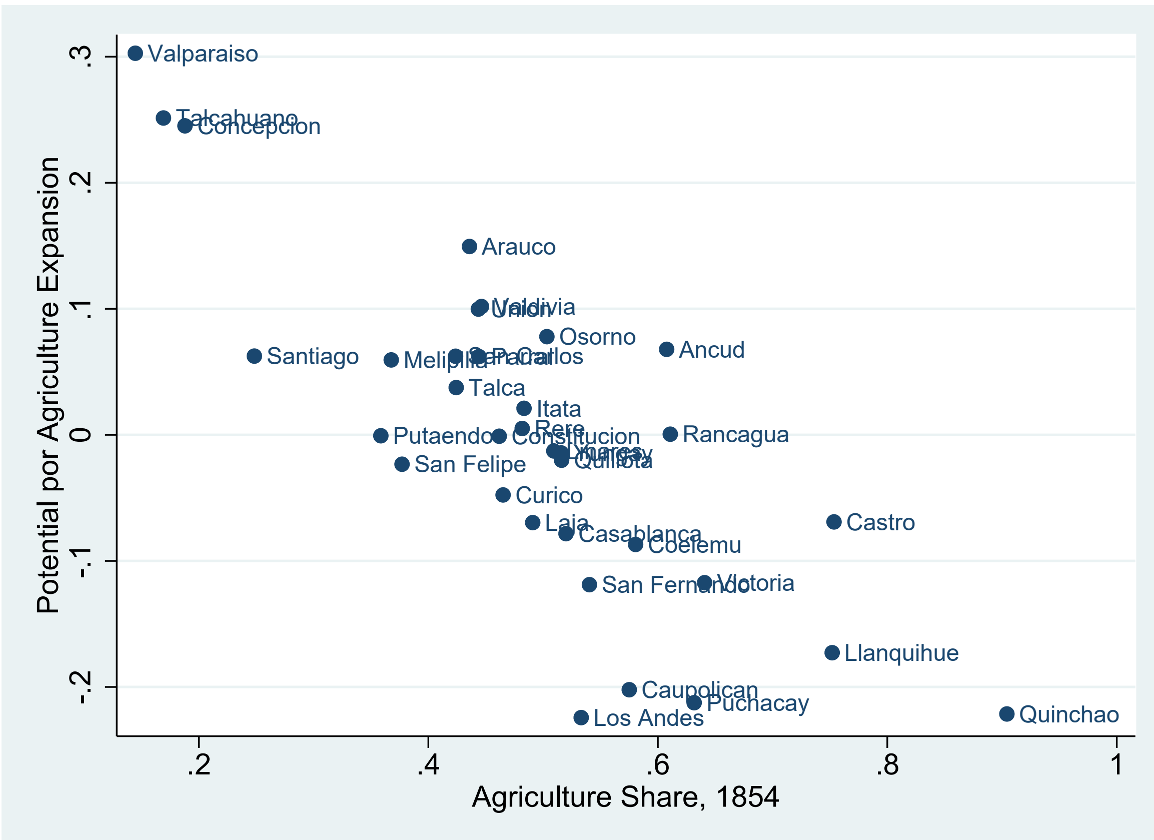
Notes: Own construction using information and maps from Espinoza (1897) for the the period 1854–1897 and Vasallo and Matus (1943) for the period 1897–1920. We complemented and checked this information with Thompson and Angerstein (1997), Alliende (2006), and several web sources.

Figure 2: Real exchange rate and government size



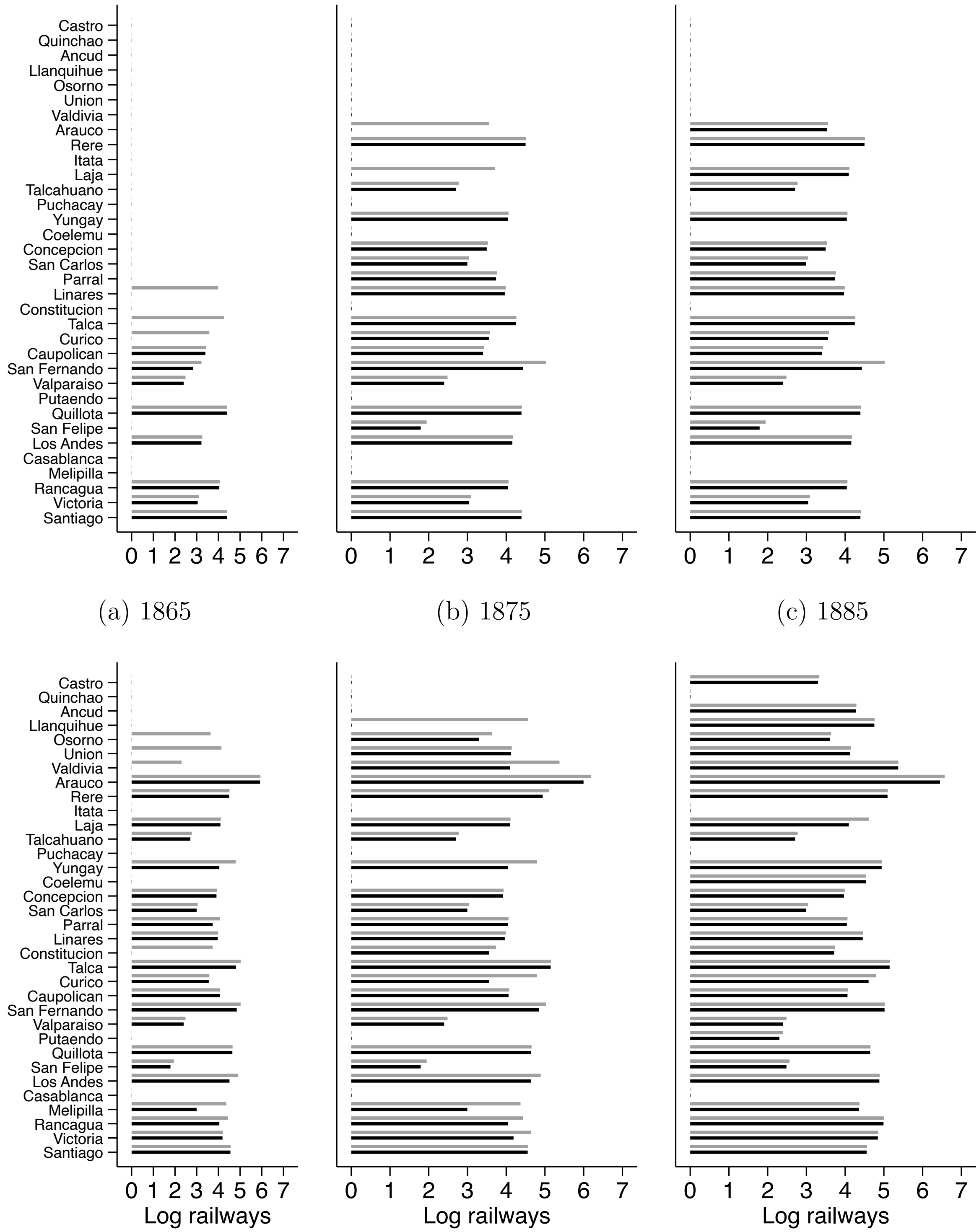
Notes: Own construction using data from Diaz et al. (2016).

Figure 3: Agricultural potential and initial share of labor in agriculture



Notes: We define "Potential Agricultural Expansion" (y-axis) corresponds to the negative of the residual of a regression of the share of agriculture in 1854 (before railroads) on wheat suitability, distance to santiago, and socioeconomic variables (see Table A.3).

Figure 4: Actual and expected railroad construction

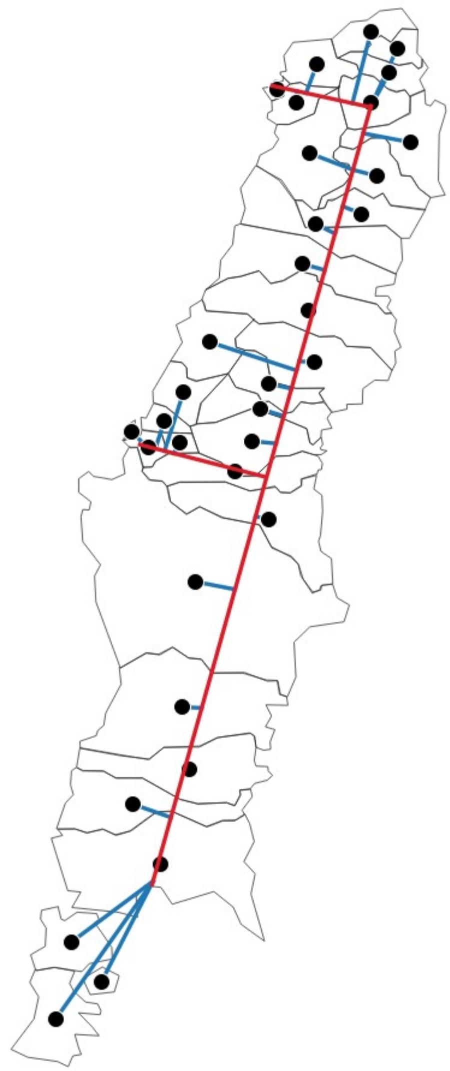


(a) 1865 (b) 1875 (c) 1885

(d) 1897 (e) 1907 (f) 1920

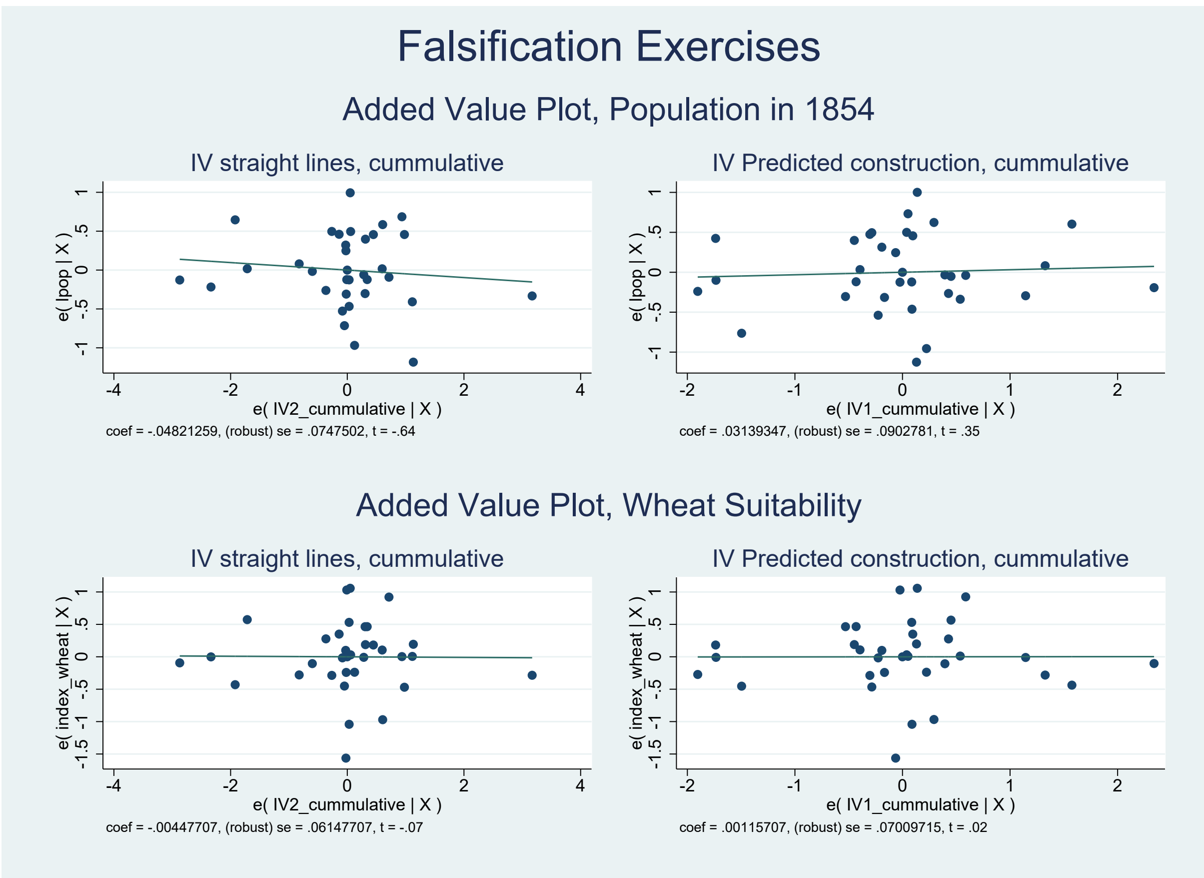
Notes: Departments are ordered in the vertical axis according to distance to Santiago. Gray lines depict expected construction of railroads and black lines actual railroads constructed.

Figure 5: Straight lines



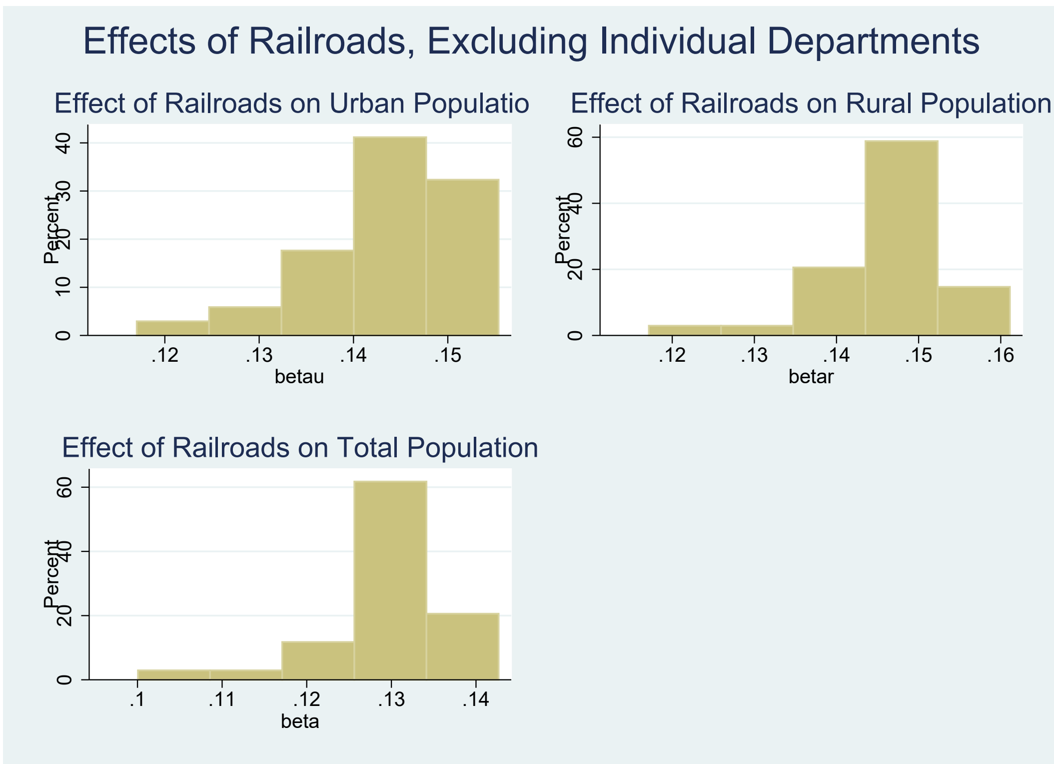
Notes: This map shows the shortest distance (blue lines) from the 34 departments in our analysis to the straight lines (red lines) that connected the three largest cities in Chile: Valparaiso, Santiago, and Concepcion. Black circles represent the centroid of departments.

Figure 6: Added value plots, falsification exercises



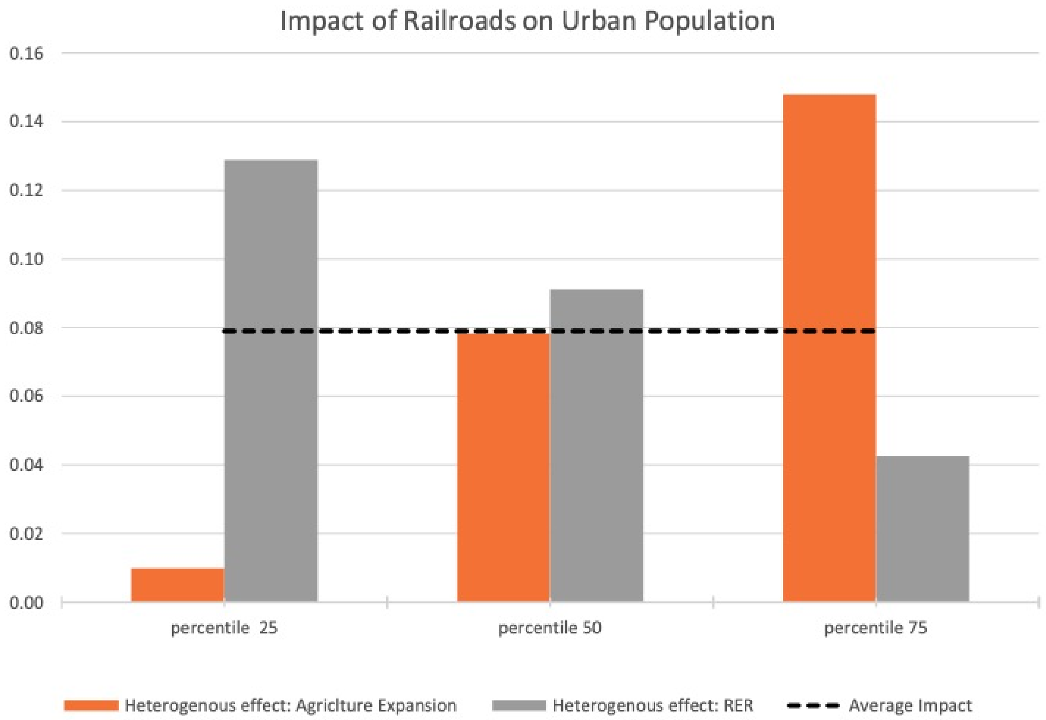
Notes: Added value plots correspond to regressions presented in Table A.4.

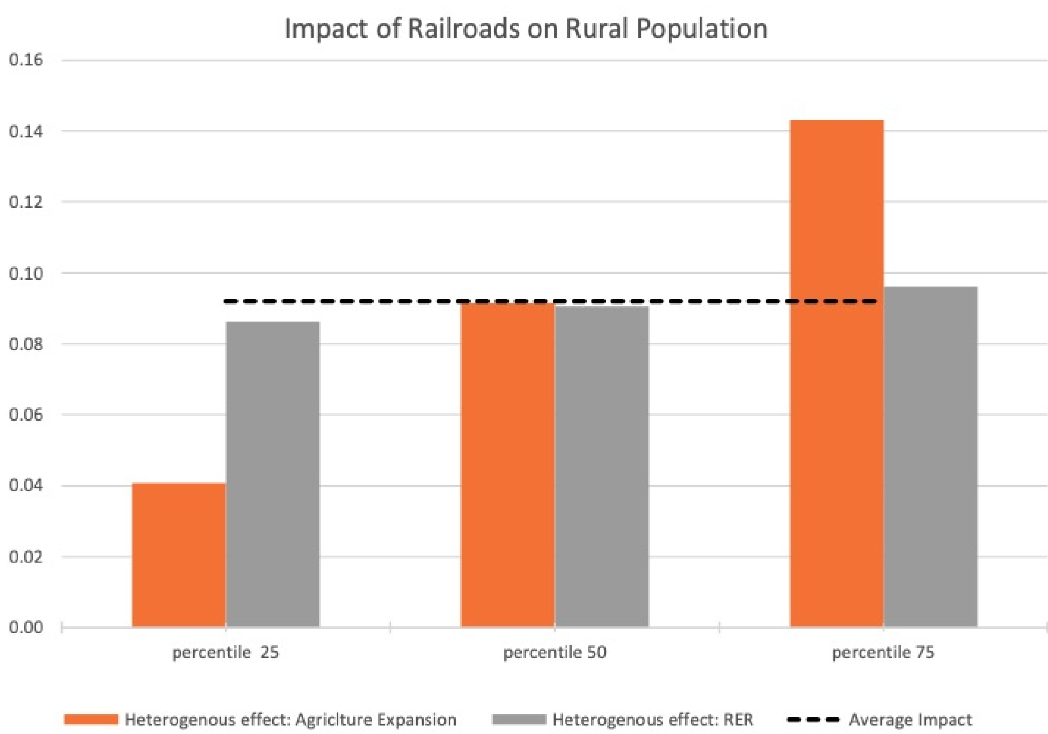
Figure 7: Effects of Railroads, Excluding Individual Departments



Notes: Coefficients on log of railroads from our main instrumental variables regression excluding one of the departments in each regression.

Figure 8: Heterogenous Effects of Railroad Expansion:





Notes: Own construction using results in Table 8.

Table 1: Economic and Social Variables, 1865-1920

Panel A: Macroeconomic Variables

	GDP growth (per-capita)	GDP per capita to US	GDP per capita to LAC	Inflation	Government Size (over GDP)	Terms of Trade (1865=100)	Real Exchange Rate (1865=100)	Price of Wheat (1865=100)	Exports (over GDP)
1865	1.8	0.49	1.33	0.0	10.0	100.0	100.0	100.0	12.9
1875	2.1	0.59	1.46	2.0	11.2	105.9	66.8	101.6	12.8
1885	1.5	0.52	1.41	6.0	9.2	96.0	86.1	86.7	16.5
1895	2.3	0.54	1.46	4.4	13.0	91.2	77.6	72.0	21.9
1907	1.2	0.48	1.48	3.9	15.1	229.7	71.7	81.9	18.6
1920	0.9	0.53	1.61	6.0	10.8	157.7	66.8	84.0	22.2

Panel B: Social Variables

	Years of Schooling	Gini Index	Population Growth	Birth Rate (per 10,000)	Death rate (per 10,000)	Share of Foreign born
1865	3.4	0.57	1.7	44	25	1.2
1875	3.8	0.58	1.4	45	26	1.2
1885	4.2	0.54	1.8	41	28	4.2
1895	4.4	0.51	1.2	38	33	2.9
1907	5.0	0.52	1.2	41	31	4.1
1920	6.1	0.56	1.3	43	31	3.1

Panel C: Economic Structure Variables

	GDP Share				Employment Share				Exports Share			
	Agr.	Mining	Manuf.	Services	Agr.	Mining	Manuf.	Services	Agr.	Mining	Manuf.	Rest
1865	30	18	28	24	43	3	25	29	31	64	0	5
1875	25	19	27	29	45	4	24	28	38	61	1	0
1885	18	25	28	29	41	4	23	33	13	83	4	0
1895	13	29	23	35	41	3	20	35	11	86	3	0
1907	11	31	20	38	38	3	15	45	7	91	2	0
1920	12	38	21	29	36	4	16	43	11	85	4	0

Sources: Own construction using data from Diaz et al. (2016).

Table 2: Descriptive statistics

	Average for variable:										
Year:	Railways (in km.)	Population	Urbanization Rate	Herfindahl Index	Agriculture Share	Manufacturing Share	Services Share	Fertility per woman	Literacy Rate	Duncan Index	Schools per 1,000 people
1865	9.50	45,564.21	0.23	0.38	0.49	0.29	0.19	0.09	0.13	18.05	0.06
1875	25.00	52,944.12	0.27	0.41	0.53	0.28	0.19	0.11	0.21	18.63	0.07
1885	27.76	62,657.59	0.35	.	.	.	.	0.09	0.26	.	.
1895	45.65	67,284.09	0.35	0.39	0.51	0.25	0.18	0.11	0.26	17.76	0.18
1907	55.68	79,093.97	0.32	0.35	0.49	0.26	0.22	0.12	0.35	19.89	0.26
1920	90.76	92,973.09	0.35	0.35	0.50	0.21	0.26	0.12	0.45	19.41	0.46

Sources: Own construction using data from historical censuses. Averages for all variables.

Table 3: Railroads and population

Dependent variable:	Log urban population		Log rural population		Urbanization		Log population	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log railways	0.225***	0.099***	0.117***	0.104***	0.017***	0.001	0.141***	0.088***
	(0.025)	(0.031)	(0.022)	(0.026)	(0.004)	(0.009)	(0.021)	(0.025)
Department fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	202	202	202	202	202	202	204	204
R^2	0.875	0.915	0.901	0.904	0.908	0.935	0.894	0.906

Notes: OLS regressions. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Railroads and population, robustness checks

Dependent variable:	Log urban population		Log rural population	
	(1)	(2)	(3)	(4)
Log railways	0.099***	0.109***	0.083***	0.107***
	(0.035)	(0.031)	(0.027)	(0.026)
Department fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Sample	Excl. Big cities	Excl. 1885	Excl. Big cities	Excl. 1885
Observations	184	168	184	168
R^2	0.898	0.932	0.883	0.901

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: First-stage

Dependent variable:	Log railways					Dummy for Railroads
	(1)	(2)	(3)	(4)	(5)	(6)
Log predicted railways	0.731***		0.420***	0.412***	0.434***	0.124***
	(0.073)		(0.095)	(0.090)	(0.098)	(0.027)
Log straight lines (dummy)		0.850***	0.632***	0.639***	0.609***	0.121***
		(0.039)	(0.089)	(0.087)	(0.094)	(0.027)
Department fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Complete	Complete	Complete	Excl. Connecting cities	Excl. Big cities	Complete
Observations	204	204	204	204	174	186
R-squared	0.879	0.915	0.946	0.894	0.953	0.944

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6: Population and railroads, instrumental variables

Dependent variable:	Log urban population			Log rural population			Urbanization			Log population		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log railways	0.135***	0.159***	0.144***	0.138***	0.160***	0.146***	0.001	0.003	0.002	0.123***	0.140***	0.129***
	(0.038)	(0.044)	(0.036)	(0.035)	(0.039)	(0.033)	(0.005)	(0.006)	(0.005)	(0.033)	(0.036)	(0.032)
	[0.045]	[0.056]	[0.044]	[0.043]	[0.042]	[0.039]	[0.004]	[0.005]	[0.003]	[0.041]	[0.045]	[0.038]
Department fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	202	202	202	202	202	202	202	202	202	204	204	204
Inst. Var.	SL	PR	Both	SL	PR	Both	SL	PR	Both	SL	PR	Both
Under-id test (p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Over-id test (p-value)	-	-	0.533	-	-	0.476	-	-	0.701	-	-	0.527

Notes: PR=Predicted Railroads, SL= Straight lines. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors robust to spatial correlation in square brackets.

Table 7: Main estimates using different panel data methods

Dependent variable:	Log urban population			Log rural population			Urbanization			Log population		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log railways	0.190***	0.055***	0.239***	0.154***	0.048**	0.256***	0.004	0.001	-0.002	0.153***	0.045***	0.194***
	(0.033)	(0.022)	(0.081)	(0.025)	(0.021)	(0.074)	(0.005)	(0.004)	(0.011)	(0.022)	(0.017)	(0.074)
Method	Random Effects	First Differences	Long Differences	Random Effects	First Differences	Long Differences	Random Effects	First Differences	Long Differences	Random Effects	First Differences	Long Differences
Observations	202	168	32	202	168	32	202	168	32	204	170	34

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: The role of population and economic variables

Dependent variable	Log of Urban Pop.	Log of Rural Pop.	Urbanization Rate	Herfindhal Index	Share of Agriculture	Share of Primary Non-Agriculture	Share of Industry	Share of Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log railways	0.079**	0.092***	-0.000	-0.007	0.001	0.000	0.004	-0.003
	(0.037)	(0.027)	(0.005)	(0.005)	(0.006)	(0.001)	(0.004)	(0.003)
Log railways × Pot. Agr. Expansion	0.993***	0.736***	0.044	0.106***	0.120***	-0.011	-0.084**	-0.028
	(0.312)	(0.254)	(0.052)	(0.037)	(0.041)	(0.013)	(0.035)	(0.023)
Log railways × Log RER	-0.339**	0.039	-0.060***	0.038*	0.064**	0.004	-0.023	-0.027*
	(0.138)	(0.078)	(0.021)	(0.020)	(0.026)	(0.003)	(0.015)	(0.016)
Department fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	202	202	202	170	170	170	170	170

Notes: RER stands for "real exchange rate." Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: The role of human capital and fertility

Dependent variable	Duncan Index	Literacy Rate	Log of Schools to Population	Share of Teachers	Share of Students	Share of Foreigners	Births per Woman	Share of Out of Wedlock Births
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log railways	-0.056	0.000	0.006	-0.000	0.002	-0.001	0.002	0.008
	(0.142)	(0.004)	(0.008)	(0.000)	(0.002)	(0.001)	(0.002)	(0.006)
Log railways × Pot. Agr. Expansion	-1.842**	-0.040	-0.101*	-0.000	-0.027*	-0.002	0.034	0.016
	(0.908)	(0.032)	(0.059)	(0.000)	(0.015)	(0.006)	(0.047)	(0.055)
Log railways × Log RER	-1.238*	0.017	0.040	-0.000	-0.002	-0.004*	-0.013	0.002
	(0.749)	(0.013)	(0.032)	(0.000)	(0.005)	(0.002)	(0.013)	(0.018)
Department fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	170	204	170	170	137	204	168	168

Notes: RER stands for "real exchange rate." Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix

Railroads, specialization, and population growth in small open economies: Evidence from the First Globalization

Andrés Forero, Francisco A. Gallego, Felipe González, and Matías Tapia

List of Tables

Table A.1: Definition, departments

Consolidated Department	Area (km^2)	Pop in 1865	Pop in 1920	Other Areas included
Ancud	3,873	21,008	28,383	San Carlos, Chacao, Dalcahue
Arauco	36,411	63,963	465,246	Angol, Traiguén, Mariluan, Collipulli, Nacimiento, Mulchen, Lautaro, Temuco, Llaima, Lebu, Imperial, Cañete
Casablanca	1,683	13,678	12,010	–
Castro	18,041	26,614	63,633	Lemuy, Chonchi
Caupolicán	3,323	72,448	75,624	–
Coelemu	3,689	30,688	26,411	–
Concepción	617	29,478	82,662	Quirihue
Constitución	4,101	75,458	72,893	Chanco, Cauquenes
Curicó	7,714	90,589	108,148	Vichuquén, Santa Cruz
Itata	2,309	37,141	40,338	–
Laja	7,607	37,415	65,187	–
Linares	8,023	53,220	83,904	Loncomilla
Llanquihue	84,871	21,342	74,809	Calbuco, Carelmapu
Los Andes	1,891	29,991	29,022	–
Melipilla	4,152	28,986	70,211	San Antonio, Cachapoal
Osorno	6,805	16,259	62,397	–
Parral	2,187	22,164	35,380	–
Puchacay	754	21,771	17,070	–
Putaendo	2,025	22,598	15,687	–
Quillota	2,936	54,220	83,219	Limache
Quinchao	341	11,400	18,532	Quenac
Rancagua	9,137	60,285	96,098	Maipo
Rere	4,433	32,777	42,358	–
San Carlos	3,378	38,586	47,027	–
San Felipe	1,910	27,261	27,774	–
San Fernando	6,664	70,008	90,718	–
Santiago	3,689	168,553	553,498	Colina, Renca, Ñuñoa, Lampa
Talca	9,948	100,575	133,957	Lontue, Curepto
Talcahuano	167	4,933	30,098	–
Unión	5,571	11,310	42,829	Rio Bueno
Valdivia	15,880	12,119	132,312	Villarrica
Valparaíso	440	74,731	225,169	–
Victoria	5,501	41,479	61,649	–
Yungay	5,445	86,823	123,398	Bulnes, Chillán

Source: Own construction based on National Censuses.

Table A.2: Definition, occupations

Category	Occupational Score	Occupations
1. Hunting and Fishing	10	Divers, fishermen
2. Agriculture	8	Farmers, *amansadores*, beekeepers, laborers, horticulturists, gardeners, loggers, pruners, cheese makers, vintners.
3. Mining	12	*Canteros*, miners.
4. Industry	30	Sharpeners, watermen, masons, *almidoneros*, shipowners, gunsmiths, millers, embroiderers, bronzesmiths, woodpeckers, coal workers, caulkers, basket weavers, coppersmiths, *caleros*, brewers, *cedaceros*, chocolate makers, seamstresses, mattress makers, leaves cutter, firework makers, shipbuilders, tanners, rope makers, distillers, gilders, confectioners, carpenters, paperhangers, bookbinders, spur makers, *estereros*, broom makers, *estucadores*, piano makers, noodle makers, gunpowder manufacturers, textile manufacturers, manufacturers of rigging, florists, stokers, smelters, recorders, guitar makers, farriers, weavers, tinsmiths, bakers, soap makers, jewelry, laundresses, lamp makers, potters, mechanics, millers, tailors, confectioners, seamstresses, umbrella stand makers, *pelloneros*, painters, goldsmiths, plumbers, watchmakers, rein makers, chair makers, hat makers, saddlers, carvers, upholsterers, tile makers, dyers, printers, *tinajero*, coopers, braiders, sailboat makers, glaziers, shoemakers.
5. Transport	30	Carriers, *careneros*, teamsters, coachmen, boaters, sailors, marines, machinist, telegraphers.
6. Commerce	40	
7. Liberal	90	
8. Medical	92	
9. Arts	52	
10. Teaching	72	
11. Cult	52	
12. Public servants	50	
13. Public force	40	Militars.
14. Domestic service	20	
15. Other	12	

Source: Own construction based on national censuses. The occupational scores follow Duncan (1961)

Table A.3: Estimation, potential for agricultural expansion

	(1)	(2)	(3)
	Share of Agriculture in Labor in 1854		
Wheat Suitability	0.102***	0.098***	0.066**
	(0.027)	(0.023)	(0.027)
Distance to Santiago		0.222***	0.124*
		(0.071)	(0.068)
Share of Literate Pop.			0.300
			(0.386)
Urbanization			-0.414**
			(0.193)
Population Density			-0.000
			(0.000)
Observations	43	43	36
R-squared	0.219	0.394	0.631

Notes: Robust standard errors in parentheses. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Robustness, falsification exercises

	Log of Population Density in 1854	Wheat Suitability Index	
	(1)	(2)	(3)
Predicted railroads, cumulative	0.031	0.001	
	(0.090)	(0.070)	
Straight lines, cumulative	-0.048	-0.004	
	(0.075)	(0.061)	
Observations	34	34	
R-squared	0.594	0.257	

Notes: Robust standard errors in parentheses. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table A.5: Robustness, main estimates

Dependent variable:	Log urban population			Log rural population			Urbanization			Log population		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log railways	0.146***	0.161***	0.143***	0.126***	0.142***	0.146***	0.004	0.002	0.004	0.121***	0.147***	0.129***
	(0.041)	(0.039)	(0.035)	(0.036)	(0.033)	(0.033)	(0.005)	(0.005)	(0.005)	(0.034)	(0.033)	(0.031)
Department fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Exclusion	Big cities	Connecting cities	None	Big cities	Connecting cities	None	Big cities	Connecting cities	None	Big cities	Connecting cities	None
Observations	184	172	202	184	172	202	184	172	202	186	174	204
Inst. Var.	SL, PR	SL, PR	SLD, PR	SL, PR	SL, PR	SLD, PR	SL, PR	SL, PR	SLD, PR	SL, PR	SL, PR	SLD, PR
Under-id test (p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Over-id test (p-value)	0.610	.620	0.521	0.465	0.606	0.470	0.888	0.671	0.677	0.472	0.541	0.500

Notes: PR=Predicted Railroads, SL= Straight lines with dummy, SLD= Straight lines with distance. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Robustness, extensive and intensive margin

Dependent variable:	Log urban population	Log rural population	Urbanization	Log population
	(1)	(2)	(3)	(4)
Log railways	0.173**	0.177**	0.001	0.136*
	(0.077)	(0.073)	(0.014)	(0.073)
Extensive Margin of railways	-0.128	-0.136	0.003	-0.029
	(0.293)	(0.257)	(0.063)	(0.255)
Department fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	202	202	202	204

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.7: Mechanisms, robustness to macroeconomic variables

Dependent variable	Log of Urban Pop.	Log of Rural Pop.	Urbanization Rate	Herfindhal Index	Share of Agriculture	Share of Primary Non-Agriculture	Share of Industry	Share of Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: GDP growth								
Log railways	0.065*	0.082***	-0.001	-0.007	-0.000	0.001	0.003	-0.003
	(0.035)	(0.026)	(0.005)	(0.005)	(0.006)	(0.001)	(0.004)	(0.003)
Log railways *	1.025***	0.758***	0.046	0.106***	0.120***	-0.011	-0.083**	-0.028
Pot. Agr. Expansion	(0.312)	(0.254)	(0.052)	(0.038)	(0.043)	(0.013)	(0.034)	(0.023)
Log railways *	-0.288**	0.066	-0.055***	0.033**	0.055***	0.004	-0.027*	-0.024
Log RER	(0.131)	(0.076)	(0.020)	(0.017)	(0.021)	(0.003)	(0.014)	(0.017)
Log railways *	-7.916***	-1.893	-1.119**	0.438	0.889	0.026	0.523	-0.338
GDP growth	(2.980)	(1.925)	(0.502)	(0.796)	(0.983)	(0.083)	(0.596)	(0.356)
Panel B: Mining GDP growth								
Log railways	0.081**	0.093***	-0.001	-0.007	0.001	0.001	0.004	-0.002
	(0.036)	(0.027)	(0.005)	(0.005)	(0.006)	(0.001)	(0.004)	(0.003)
Log railways *	0.972***	0.725***	0.046	0.110***	0.127***	-0.012	-0.082**	-0.026
Pot. Agr. Expansion	(0.312)	(0.251)	(0.051)	(0.038)	(0.043)	(0.013)	(0.035)	(0.023)
Log railways *	-0.316**	0.034	-0.052**	0.040	0.075*	0.004	-0.017	-0.012
Log RER	(0.126)	(0.075)	(0.021)	(0.033)	(0.041)	(0.004)	(0.022)	(0.016)
Log railways *	-0.031***	-0.007	-0.002	-0.000	0.002	0.000	0.001	0.004***
Mining GDP growth	(0.010)	(0.005)	(0.002)	(0.004)	(0.005)	(0.000)	(0.002)	(0.001)
Panel C: Government Expenditure (as percentage of GDP)								
Log railways	0.071*	0.082***	0.000	-0.008*	0.000	0.001	0.004	-0.002
	(0.040)	(0.027)	(0.005)	(0.005)	(0.006)	(0.001)	(0.004)	(0.003)
Log railways *	1.067***	0.815***	0.045	0.092***	0.103***	-0.010	-0.081**	-0.015
Pot. Agr. Expansion	(0.323)	(0.265)	(0.052)	(0.032)	(0.038)	(0.013)	(0.033)	(0.022)
Log railways *	-0.362**	0.031	-0.069***	0.050**	0.073***	0.004	-0.024	-0.040***
Log RER	(0.162)	(0.090)	(0.023)	(0.019)	(0.025)	(0.003)	(0.016)	(0.015)
Log railways *	-0.007	-0.004	-0.002	0.005**	0.004	-0.000	-0.000	-0.005***
Government Expenditure	(0.015)	(0.009)	(0.002)	(0.002)	(0.003)	(0.000)	(0.002)	(0.001)
Department fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	202	202	202	170	170	170	170	170

Notes: RER stands for "real exchange rate." Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.8: Mechanisms, robustness to the price of wheat

Dependent variable	Log of Urban Pop.	Log of Rural Pop.	Urbanization Rate	Herfindhal Index	Share of Agriculture	Share of Primary Non-Agriculture	Share of Industry	Share of Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log railways	0.065*	0.081***	-0.002	-0.006	0.001	0.000	0.004	-0.004
	(0.033)	(0.027)	(0.005)	(0.006)	(0.007)	(0.001)	(0.005)	(0.003)
Log railways * Pot. Agr. Expansion	-1.071***	-0.783***	-0.051	-0.104**	-0.118*	0.011	0.085**	0.026
	(0.292)	(0.234)	(0.047)	(0.049)	(0.061)	(0.008)	(0.040)	(0.029)
Log railways * Log Wheat Price	-0.549***	-0.024	-0.079***	0.049**	0.077***	0.003	-0.005	-0.038***
	(0.126)	(0.101)	(0.020)	(0.023)	(0.028)	(0.004)	(0.019)	(0.014)
Department fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	202	202	202	170	170	170	170	170

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.9: Mechanisms, additional exercises

Dependent variable	Log of Urban Pop.	Log of Rural Pop.	Urbanization Rate	Herfindhal Index	Share of Agriculture	Share of Primary Non-Agriculture	Share of Industry	Share of Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Distance to Santiago								
Log railways	0.065*	0.082***	-0.001	-0.007	-0.000	0.001	0.003	-0.003
	(0.035)	(0.026)	(0.005)	(0.005)	(0.006)	(0.001)	(0.004)	(0.003)
Log railways *	0.770***	0.576**	0.041	0.113***	0.137***	-0.012	-0.106***	-0.022
Pot. Agr. Expansion	(0.286)	(0.227)	(0.051)	(0.042)	(0.046)	(0.013)	(0.037)	(0.024)
Log railways *	-0.316**	0.028	-0.054**	0.037*	0.064**	0.004	-0.023	-0.027
Log RER	(0.138)	(0.076)	(0.021)	(0.020)	(0.025)	(0.003)	(0.015)	(0.016)
Log railways *	0.096***	0.084***	-0.001	-0.005	-0.010*	-0.000	0.011***	-0.001
Distance to Santiago	(0.027)	(0.026)	(0.005)	(0.005)	(0.006)	(0.001)	(0.004)	(0.003)
Panel B: Terrain Ruggedness								
Log railways	0.075*	0.084***	-0.000	-0.006	0.001	0.001	0.004	-0.004
	(0.038)	(0.028)	(0.005)	(0.005)	(0.006)	(0.001)	(0.004)	(0.003)
Log railways *	0.970***	0.708**	0.047	0.105***	0.120***	-0.012	-0.085**	-0.027
Pot. Agr. Expansion	(0.324)	(0.275)	(0.051)	(0.038)	(0.041)	(0.013)	(0.035)	(0.024)
Log railways *	-0.335**	0.039	-0.061***	0.040*	0.065**	0.005	-0.022	-0.029*
Log RER	(0.138)	(0.074)	(0.020)	(0.020)	(0.026)	(0.003)	(0.015)	(0.016)
Log railways *	-0.000	-0.001	0.000	0.000	0.000	0.000	-0.000	-0.000
Terrain Ruggedness	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Department fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	202	202	202	170	170	170	170	170

Notes: RER stands for "real exchange rate." Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.10: Mechanisms, alternative channels

Dependent variable	Share of State Empl.	Share of Transport	Share of Telegraph
	(1)	(2)	(3)
Log railways	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Department fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	170	170	170

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.