Locomotives of Local Growth: The Short- and Long-Term Impact of Railroads in Sweden*

Thor Berger† Kerstin Enflo

Preliminary Draft

Abstract

This paper uses city-level data to examine the impact of a first wave of railroad construction in Sweden, 1855-1870, from the 19th century until today. We estimate that railroads accounted for 50% of urban growth, 1855-1870. In cities with access to the railroad network, property values were higher, manufacturing employment increased, establishments were larger, and more information was distributed through local post offices. Today, cities with early access to the network are substantially larger compared to initially similar cities. We hypothesize that railroads set in motion a path dependent process that shapes the economic geography of Sweden today.

JEL: N73, N93, R12, R40.

Keywords: Railroads, Industrialization, Urban Growth, Path Dependence.

*Berger: PhD Candidate, Department of Economic History, School of Economics and Management, Lund University. Alfa 1, Scheelevägen 15 B, 22363 Lund. (E-mail: Thor.Berger@ekh.lu.se) Enflo: Department of Economic History, School of Economics and Management, Lund University. Alfa 1, Scheelevägen 15 B, 22363 Lund. (E-mail: Kerstin.Enflo@ekh.lu.se) We are grateful for comments and suggestions made by Dan Bogart, Joan Rosés, Nikolaus Wolf, and seminar participants at Copenhagen Business School, Humboldt University, Lund University, and the University of Southern Denmark on earlier drafts of this paper. An earlier version of this paper was published as an EHES working paper (#42).We gratefully acknowledge funding from the Swedish Research Council (grant no. 2008-2023) and the Crafoord Foundation (grant no. 20130812). The usual disclaimer applies.

†Corresponding author.

1 Introduction

More than 250,000 km of railroads were constructed in 19th-century Europe, making it the most ambitious pan-European infrastructure project to date (Mitchell 1975). Despite widespread perception that railroads caused some regions to thrive and others to decline, we know little about how local economies were affected (Pollard 1981).[1] Railroad construction also entailed substantial sunk investments, making these historical railroad lines remarkably persistent: about 70% of the European railroad network in service today was in place already by 1900 (Martí-Henneberg 2013).[2] This constitutes a largely unexplored link between historical investments in transport infrastructure and long-term patterns of local economic development.

This paper uses city-level data to analyze the impact of a *first wave* of railroad construction in Sweden, 1855-1870, from the 19th century until today. We begin by asking if railroads had a causal short-term effect on urban economic activity. This relates to a prominent historical debate about whether railroads that were built "ahead of demand" were capable of igniting a process of economic development (e.g, Fishlow 1965). In the second part of the paper, we ask if this first wave of expansion affected patterns of urban growth, over the last 200 years. These questions lend themselves to evaluating the impact of the railroad using a reduced-form approach, where we simply compare relative outcomes for cities with access to the railroad network to those without.

However, it is not straightforward to identify the impact of infrastructure, because investments typically are allocated to already growing areas. In light of this, the extension of the 19th-century Swedish railroad network provides a compelling setting for three key reasons. First, construction of the network remained largely under the auspices of the state (Heckscher 1954). The evolution of the netowork did therefore not merely reflect local differences in transport demand. Second, it largely followed a predetermined plan. Third, the main trunk lines were explicitly routed with a motive to promote development in disadvantaged regions (Rydfors 1906; Sjöberg 1956). This meant that many (important) cities and regions, with pressing transport needs, were avoided. By 1870, less than a third of all cities had gained access to the 1,727 km network (see Figure 1). This provides a setting that allows us to examine the impact of infrastructure on local economic development.

Our empirical approach centers around comparing the population of cities - a broad

[1]Economic historians have typically evaluated the contribution of railroads to the aggregate economy by estimating the 'social savings' - the difference between the cost of transporting a fixed amount of goods by rail compared to alternative transport modes. See Fogel (1964, 1979), Fishlow (1965), O'Brien (1977, 1983), and the more recent contributions of Leunig (2006), Herranz-Loncán (2006), Donaldson and Hornbeck (2012), and Bogart and Chaudhary (2013). However, irrespective of the contribution of railroads to the wider economy, even small differences in transport costs can have large effects on local economies (Krugman 1991a,b).

[2]Construction entailed investments in, for example, embankements, drainage ditches, and cuttings. This leads Atack et al. (2008, p.14) to argue that "once a railroad was built in a specific location, it stayed where it was because the bulk of the railroad's investment was not just fixed but also sunk (literally)."

proxy for economic activity - that gained access to the railroad network to cities that did not, using a difference-in-differences strategy.[3] We show that cities that gained access to the railroad network between 1855 and 1870 expanded by an additional 26% on average over the same period. A simple back-of-the-envelope calculation implies that in the absence of railroad construction the level of urbanization in 1870 would decrease by 15%, and the rate of aggregate urban growth between 1855 and 1870 would decrease by 50%. These effects are sizable, taking into account that only a tenth of the network at its peak size had been laid at this point (Nicander 1980).

To alleviate concerns about endogenous placement of lines we use three alternative identification strategies. First, we compare observationally similar cities using a matching strategy. This yields nearly identical estimates. It is therefore unlikely that are findings are driven by observable differences between cities with and without access to the network. Second, we draw upon the two existing plans of the network and low-cost routes between major cities in an instrumental variables strategy. This corroborates our findings, and suggests an even larger impact of the railroad. Third, we examine the effects for lines that were proposed but not ultimately built by 1870 and lines that were constructed between 1870 and 1880 - i.e., after the period that we examine - in three placebo specifications. Our estimates for these lines are close to zero and statistically insignificant. Conditional on these lines initially being assigned on similar grounds as those actually built, this suggests that unobservable differences are not driving our findings. Taken together, these results suggest that the expansion of railroads had a causal short-term effect on local economic activity, but doesn't identify the underlying mechanisms.

Cities grew because they attracted migrants. Railroads potentially induced migration by lowering the costs of relocating and by increasing urban employment opportunities, as emphasized in canonical cost-benefit models of migration (e.g., Sjaastad 1962; Lee 1966; Harris and Todaro 1970). Access to the railroad network enabled local firms to sell their goods in more distant markets and to obtain raw materials more cheaply.[4] This should encourage an increase in the scale of production, potentially leading to coordinated investments across firms causing a local "big push" (Rosenstein-Rodan 1943; Murphy et al. 1989). Industrial expansion should be reflected in higher labor demand, wages, and housing and land rents (Rosen 1979; Roback 1982; Glaeser and Gottlieb 2009). Railroads also lowered the cost of distributing information in the form of mail and newspapers. Ideas and new technologies

[3]See, for instance, Cantoni (2011), Dittmar (2011), and Nunn and Qian (2011) for recent contributions that also rely on city populations as a proxy for economic activity in a historical setting. In our data, virtually all outcomes - such as the share of population employed in industry, property values, and the activity of local post offices - are positively and significantly correlated with the size of cities.

[4]Chandler (1965) famously argued that the origins of large business units in the United States were to be found in the administration of the railroads. While this is an intriguing channel through which the expansion of railroads could have affected industrial organization, it is one we cannot address with our dataset. However, see Montgomery (1947, p.204) for a similar argument regarding the advent of railroads in Sweden.

would therefore spread faster, and overall trade costs would decrease (Anderson and van Wincoop 2004).

We explore these mechanisms by drawing upon cross-sectional data for 1870 on (i) manufacturing employment and the size of establishments (ii) housing and land prices (iii) distribution of information through local post offices. Using within-region variation in access to the network, we show that access to the railroad network was associated with an increase of manufacturing employment of 2.8 percentage points (more than 120% of the sample mean) on average. Manufacturing establishments were more likely to belong to incorporated firms as opposed to sole proprietors, were twice as large, and used more steam engines compared to establishments in cities without access to the network. Housing and land prices were also substantially higher, implying large productivity gains associated with access to the network. Using data from local post offices, we document that inhabitants in cities with access to the network consumed more mail, newspapers, and sent more parcels, generating higher incomes for local post offices. Taken together, these results suggest that economic expansion was underpinned by productivity gains due to economies of scale and higher rates of information diffusion.

In order to examine the long-term impact of this first wave of railroad expansion, we compare the population of cities with and without access to the railroad network by the end of the first wave of expansion, over the last 200 years. Early access to the railroad network translated into a persistent difference in city size, despite the fact that access to the network had been extended to virtually all cities well before the 20th century. Today, cities that gained access to the railroad network during the first wave of expansion are on average 62% larger and to be found 11 steps higher in the urban hierarchy, compared to initially similar cities.

What explains this long-term persistence? We hypothesize that the routing of the first railroad lines solved a coordination problem of future infrastructure investments. Once these first lines were in place, investments in roads and railroads were mainly directed at constructing branches to cities that already formed part of the network (Westlund 1998, 1992). This is why the "first lines mattered". Drawing upon data from maps of the mid-20th century railroad and highway networks we document that on average 80% more railroad lines and 50% more highways emanate from cities with early access. In long-differenced regressions of the change in population 1855-2010, differences in the modern railroad network account for a substantial fraction of the differences in population growth. We empirically evaluate alternative explanations based on sunk investments in housing and communications infrastructure as well as external economies and find that they explain less of the long-run persistence that we find.

These findings contribute to three strands in the literature. First, our results contribute to a growing body of evidence that documents the causal impact of 19th-century railroads on urbanization (Haines and Margo 2006; Atack et al. 2010), city growth (Hornung 2012), the reorganization of production from artisan shops to factories (Atack et al. 2008), market integration (Keller and Shiue 2008), and agricultural development (Atack and Margo 2011; Donaldson 2012; Donaldson and Hornbeck 2012). We document similar short-term effects, but our paper differs from this literature in that we link these effects to local development trajectories spanning more than 150 years.

Second, our findings contribute to the literature on the impact of modern transport improvements on regional and urban growth (Baum-Snow 2007; Banerjee et al. 2012; Baum-Snow et al. 2012; Duranton and Turner 2012; Storeygard 2013). Our paper contributes to this literature by documenting the impact of infrastructure on urban development in a poor, rural, and predominantly agricultural setting. In addition, we provide evidence on plausible mechanisms that underlie the "first stage relationship" between historical and contemporary infrastructure.[5]

Third, our finding that railroads had persistent effects on the distribution of economic activity contributes to an emerging literature on long-term urban development, path dependence, and the persistence of spatial equilibria (Davis and Weinstein 2002; Bosker et al. 2007; Redding and Sturm 2008; Davis and Weinstein 2008; Miguel and Roland 2011; Redding et al. 2011; Bleakley and Lin 2012). Whereas this literature mainly has focused on the stability of spatial equilibria, we document how historical investments in infrastructure shapes contemporary spatial patterns of economic activity.[6] In that sense, our paper is closely related to Jedwab and Moradi (2011) that examine the long-term impact of colonial railroads in Ghana, and find that areas that gained access to a railroad in the early 20th century are still more developed today.

The remainder of this paper is structured as follows. In the next section we present the historical background and describe our data. In section three we discuss our empirical strategy and analyze the short-term impact of railroads. Section four examines the long-term impact of early access to the railroad network on population, and evaluates channels of persistence. In section five we provide some concluding remarks.

[5] See, for instance, Duranton and Turner (2012) that exploit 19th-century railroad lines as the basis for an IV strategy to examine the impact of contemporary road infrastructure on urban growth in U.S. metropolitan areas. We find that the long-term impact of historical investments in railroads seems to run through later incarnations of the network, rather than through some other channel, which lends support to studies that rely on this exclusion restriction for identification.

[6] Sunk investments in infrastructure may be one potential explanation for the fact that urban economies are extremely resilient even in the face of extreme shocks (e.g., Davis and Weinstein 2002).

2 Historical Background and Data

This section provides a brief overview of developments in 19th century Sweden and the historical background of railroad construction. We then describe our city-level dataset and compare pre-railroad characteristics for cities with and without access to the railroad network.

2.1 Swedish Developments in the 19th Century

Sweden underwent a dramatic economic, political, and social transition over the latter half of the 19th century (Gårdlund 1942; Montgomery 1947; Heckscher 1954). A host of institutional reforms were enacted around the mid-19th century: the guilds were abolished (1846), passport requirements were revoked (1860), and through a string of legislation (beginning in the 1850s) free trade was gradually introduced (Schön 2010).

Between 1856, when the first railroad line opened, and the outbreak of World War I, per capita incomes grew 65% faster than in Britain and 20% faster than in the United States.[7] Rapid convergence was also manifest in terms of real wages, increasing from about half those paid to British workers to parity (Williamson 1995; Prado 2010). Despite a low degree of urbanization, the number of urban dwellers increased from less than 400,000 to 1.5 million and the share of the population employed in manufacturing tripled, over the same period (Statistiska Centralbyrån 1969; Krantz and Schön 2007).

Several explanations has been offered for this remarkable catch-up, emphasizing a disproportionate pre-industrial accumulation of human capital (Sandberg 1979) and a dynamic domestic market (Schön 1979). Another influential explanation rests on Heckscher-Ohlin logic emphasizing the role of the expanding 19th-century commodity trade, as well as capital inflows and mass emigration (O'Rourke and Williamson 1995a,b). Eli Heckscher, however, also underlined the importance of transport improvements, arguing that "[t]here is little doubt that the revolution in transport was far more important than foreign trade policies" (Heckscher 1954, p.240). Arguably, the economic transition during the latter half of the 19th century would have been inconceivable in the absence of substantial improvement of the internal infrastructure.

2.2 Transport Before and After the Railroad

Prior to the railroad network was constructed, transportation primarily took place by pack animals and horse-drawn carts on small unpaved roads, by sleigh over 'winter roads', and along navigable waterways, the coast, and canals (Heckscher 1954; Gårdlund 1942; Gadd 2000). Transport costs were high and distinctly seasonal, since canals, waterways, and har-

[7] Average annual GDP per capita growth between 1856-1914 was 1.6% in Sweden, 1.0% in Britain, and 1.4% in the United States. Our calculations based on data provided in Bolt and van Zanden (2013).

bors froze in the winter months.[8] In addition, goods were typically transported using several modes and therefore frequently had to be reloaded. Overland transport in excess of 200 km was not viable (Heckscher 1907), and important high weight-to-value goods, such as iron ore, could not profitably be hauled more than 30 km (Sjöberg 1956).

Railroads radically altered the means of transportation, offering transport at higher speed, lower cost, during all seasons, at unitary tariff rates (Montgomery 1947). Freight rates were cut by three-fourths, passenger costs decreased by half, and travel speeds increased tenfold (Sjöberg 1956).[9] Already by the end of the 1860s, the railroad had overtaken water transport as the primary means for internal transportation (Westlund 1992).

Whereas transportation had constituted a constraint on industrialization and city growth prior to the railroad era, the emerging network allowed cheap transportation of basic necessities to urban dwellers (Thorburn 2000). By 1870, grain and fuel (coal, wood, and charcoal) constituted more than one-fifth of the tonnage transported via rail, effectively reducing the barriers to urban expansion.[10]

2.3 Planning and Construction of the Railroad Network

Prospects of a railroad network was debated in the Riksdag of the Estates as early as the 1820s.[11] However, it would take the better part of another three decades before the first lines went into operation. Whether railroads should be primarily planned, constructed, and managed by private companies or the state became a politically contentious issue. Two proposals for a national railroad network emerged during the 1840s and 1850s: one adhering to a market-based approach and the other based on a *de facto* state monopoly.

2.3.1 Adolf von Rosen's 1845 Proposal

The first proposal for a railroad network, presented in 1845, was based on privately funded lines, that were to be managed by private companies. It was conceived by Count Adolf von Rosen, a major in the Naval Mechanical Corps. He presented an extensive plan of an entire network (see Figure 1), meant to address the disruption and inefficiencies arising from local political lobbying that had plagued piecemeal railroad construction elsewhere in Europe (Sjöberg 1956).

[8]Water transport remained available, with regional variations, for about eight months of the year. However, in landlocked areas, transport costs were generally lower in the colder months as 'winter roads' provided a cheaper alternative to road transport (Heckscher 1907, 1954).

[9]Rydfors (1906, p.86) reports that freight rates by road of high-weight and low-weight goods were 6-10 and 13-17 *öre* respectively; corresponding rates by rail were 3 and 10 *öre*.

[10]Calculated from the official railroad statistics (see Appendix A).

[11]Prior to its abolishment in 1866, the Riksdag of the Estates - henceforth referred to as the *Riksdag* - was a national diet where the four estates (the nobility, clergy, burghers, and peasants) were represented. This political structure unexpectedly led to protracted debates between the estates over the perceived need and desirability of railroad construction. See Rydfors (1906) for a general discussion.

Several of the proposed routes were surveyed by von Rosen in cooperation with British engineers, and the Riksdag ordered topographical surveys of additional proposed lines (Sjöberg 1956). These surveys collected detailed geographical information, and therefore lowered the cost of future railroad construction along these routes (Rydfors 1906). Figure 1 provides suggestive evidence of this, showing that several of the lines constructed by 1870 followed the initial routes proposed by von Rosen. In section 3.1 we motivate the use of von Rosen's proposal as the basis for an instrumental variable and placebo strategy.

In the end, von Rosen's market-based approach to railroad construction resulted in a spectacular failure due to an underdeveloped domestic capital market, a lack of demand for transport services, and inflationary pressures following the Crimean War (Nicander 1980). When the syndicate of British investors that were to finance the main lines withdrew from their commitments (following the speculation and inevitable collapse during the British *Railway Mania* of the 1840s) von Rosen became confined to raising domestic capital. Despite state concessions and interest guarantees amounting to 4% of construction costs, von Rosen repeatedly failed to raise sufficient capital. Scepticism mounted among politicians (see quote below) against leaving construction of the railroad network in the hands of foreign investors and private enterprise (Rydfors 1906).

2.3.2 Nils Ericson's 1856 Proposal

> *"I therefore believe, that if one wants to extend a helping hand to our industry ... the State cannot support the improvement of the country in a more efficient, appropriate, impartial and magnificent way, than by a firm action to bring about railroads."* -Johan August Gripenstedt, Minister of Finance[12]

In the Riksdag of 1853/54 it was decided that all major trunk lines of the network were to be planned, financed, and constructed by the state. In 1855, Nils Ericson, a colonel in the Navy Mechanical Corps, was commissioned by the Riksdag to lead the construction and was bestowed with "authoritarian powers" to route the main lines at will (Rydfors 1906).

Ericson's plan for the network, presented in 1856, centered around five main trunk lines, to be constructed by the state, on which private branch lines would then expand.[13] There were two main motives behind his plan: to connect the capital Stockholm with the other two major cities (Gothenburg and Malmö) and to stimulate development in disadvantaged regions (Sjöberg 1956).[14] In addition, due to military concerns, the trunk lines were to

[12]From a speech to the Riksdag, cited and translated by Kaijser (1999, p.223).

[13]Private initiatives had to undertake a survey of the proposed route by an experienced railroad technician, obtain a state concession, and undergo a review by the technical authorities. If a proposal was approved, a joint stock company had to be formed. Financial support from the State could be granted conditional on the company finding buyers of at least half of the offered stock. Construction, traffic, and maintenance were, however, to remain under direct state supervision. See Nicander (1980, p.15).

[14]This dimension of regional policy is emphasized in all historiographical work on railroad construction

be routed through the interior, avoiding cities located close to the coastline and previously important transport routes (Schön 2010). Ericson's plan was viciously criticized and ridiculed for its "horror of waterways and cities" (Heckscher 1954, p.241). Figure 1 lends support to his contemporary critics, documenting how Ericson's proposed railroad lines avoided the important mining region *Bergslagen*, west of Stockholm, as well as historically important naval cities in the southeast.

The Riksdag initially approved construction of the Southern and Western trunk line, and in November 1862 the 455 km Western trunk line, running from Stockholm to Gothenburg, was inaugurated. Three years later the Southern trunk line opened, connecting the three major cities by rail. As evident in Figure 1, several additional branch lines were constructed to link up cities to the emerging network. Construction costs were, however, to a large part determined by the distance to the main trunk lines. Placement of these main trunk lines - that were to follow the shortest routes between their terminal points - therefore indirectly influenced the roll-out of the entire network (Rydfors 1906).

In the Riksdag of 1857, Ericson's proposal was rejected due to conflicts between the estates and increasing financial strains. In the wake of this decision, local political groups gained the clout to block and influence the construction of remaining lines. Protracted debates in the Riksdag concerning the direction of each remaining line took place throughout the 1860s, and local politicians seized on the capital to ensure that lines were routed through their districts (Westlund 1998). Ensuing political infighting meant that only part of Ericson's plan had been realized by 1870. In section 3.1 we describe how this provides a set of lines to use as the basis for a placebo strategy, and Ericson's proposal as an instrumental variable for the network actually constructed by 1870.

By 1870, the first wave of railroad expansion had reached its end. A network spanning 1,727 km - two-thirds of which were directly owned by the state - had connected less than a third of all cities. Importantly, even though Ericson's plan was eventually rejected by the Riksdag, and despite his formal resignation in 1859, Figure 1 documents that he nevertheless was able to enforce the realization of his envisioned network with hardly any changes (Rydfors 1906; Heckscher 1954).

[Figure 1 about here.]

2.4 Data on Cities and Railroads

We have constructed a new panel dataset of all cities in Sweden, observed at decadal intervals over the period 1800-2010. Our sample is restricted to cities that held town charters in 1840, prior to when railroad construction began, to ensure that cities do not endogenously enter our

in Sweden. See, for instance, Westlund (1998, p.74) who argues that railroads "were that epoch's great instrument for regional policy for spreading industrialization and economic development to new regions".

sample as an effect of the railroads. Because we exclude all cities that gained town charters after the railroads were constructed we will, however, understate the long-term impact of the railroads, as there are many smaller urban agglomerations that we *ex post* know formed cities due to their location on railroad junctions.[15] In the rest of the paper we exclude the three main terminal cities (Stockholm, Gothenburg, and Malmö), where the impact of railroads arguably differed compared to the average city, and the two insular cities, that by our definition could not gain access to the network. These restrictions reduce our baseline sample used throughout the rest of the paper to 81 cities. Detailed information on sources and construction of our dataset is provided in Appendix A.

Using geospatial software we have reconstructed the 19th century railroad network from georeferenced maps of railroad lines in Sweden today. To capture the impact of the first wave of railroad expansion, 1855-1870, we include all lines that were constructed by the end of 1870. In addition, we also digitize lines that were part of von Rosen's 1845 and Ericson's 1856 proposal. Our measure of access to the network is a binary indicator taking the value one for all cities that had direct access to the network through a rail line by 1870. To control for alternative modes of transport we also code binary indicators for all cities located at the coast and cities located by one of the four great lakes respectively.[16] This latter measure indirectly captures access to the major canals that primarily were constructed to provide direct connections between these lakes and the coast. Figure 1 shows the extent of the railroad network as of 1870, the proposed lines in the two alternative plans of the network, and the location of all cities in our (unrestricted) sample.

We collect data on population for each decade between 1800 and 2010, and for the year 1855, from historical population censuses. Many cities were little more than villages or small towns: an average city had 4,400 inhabitants in 1855, increasing to 6,200 in 1870, and eventually to 53,800 in 2010.

We have collected a richer set of outcomes in 1870 from a variety of historical sources, that allows us to explore potential mechanisms. From official statistical sources, we have digitized data on housing and land prices, manufacturing and artisanal employment, average size and ownership form of manufacturing establishments, and the number of steam engines used in production. For the local post office in each of the cities in our sample, we have also collected data on revenues and the distribution of mail, newspapers, and parcels.

In order to evaluate alternative explanations for the long-term impact of railroads, we have also collected an eclectic set of additional data, described when introduced later in the paper.

[15] See Heckscher (1907) who emphasizes the role of the railroad in creating new urban agglomerations. It is also worth noting that all 34 urban agglomerations that were awarded town charters between 1910 and 1950 had access to a railroad line, and that several of these towns owed their existence exclusively to the railroad (Westlund 1998, p.84).

[16] These being Vänern, Vättern, Hjälmaren, and Mälaren, as shown in Figure 1.

2.4.1 Descriptive Statistics and Pre-Railroad Differences

One concern is that cities that gained access to the network differed in important ways from cities that did not. If this is the case, any comparison of cities with and without access the network may reflect these differences, rather than the effect accruing from the railroad itself.

Table 1, Panel A, reports mean pre-railroad characteristics for cities with (column 1) and without (column 2) access to the railroad network by 1870, and the difference-in-means and corresponding Huber-White standard errors (column 3). Cities that gained access to the railroad network were on average larger than those that did not, were less likely to be located at the coast, and consequently had a smaller share of the population employed in the shipping sector. In terms of employment in the artisanal, trade, manufacturing, and service sector they were, however, broadly similar. Importantly, cities that gained access to the railroad network did *not* grow significantly faster in the period directly preceding railroad construction (1840-1855) suggesting that they shared a common growth trend. However, the observed differences in terms of geographical location, sectoral composition, or initial city size may reflect subtle differences between cities that did and did not gain access to the network during the first wave of expansion.

To mimic a more experimental setting we follow the evaluation literature and balance our sample on observables using propensity scores (Rosenbaum and Rubin 1983). Propensity scores are obtained from a probit regression of a binary indicator for having access to the network in 1870 on 12 pre-railroad characteristics.[17] Treatment and control groups are identified by excluding cities with very high or low propensity scores, resulting in a sample consisting of 42 out of the 81 cities included in our baseline sample.[18] Although this drastically reduces the size of our sample, we view it as a simple way to gauge the magnitude and direction of any bias arising from observable differences between cities with and without access to the network.

Table 1, Panel B, reports mean characteristics for our balanced sample. There are no remaining statistical differences between cities with and without access to the network (see column 6). Although cities that gained access to the railroad network by 1870 were marginally larger in 1855, cities that did not were slightly more industrial and had better access to urban markets. This restricted sample is therefore plausibly balanced on the characteristics that

[17]Propensity scores are estimated based on all variables in Table 1 and a first-order polynomial in longitude and latitude of the centroid of each city. Market potentials (MP) are constructed in the spirit of Harris (1954). For each city i we calculate:

$$MP_{it} = \sum_{j \neq i} P_{jt} D_{ij}^{-1}$$

where P is the population of city j in year t, and D is the geodesic distance between city i and j. This correspond to a distance-weighted measure of each city's access to domestic urban markets.

[18]We exclude all cities with propensity scores outside the interval $[0.15, 0.85]$.

we observe prior to when railroad construction began.

[Table 1 about here.]

3 The Short-Term Impact of Railroads (1840-1870)

This section examines the impact of railroad expansion on the population of cities over the period 1840-1870, documenting that cities that gained access to the railroad network grew substantially larger as a consequence. We then perform a simple back-of-the-envelope calculation of the aggregate contribution of railroads to urbanization and urban growth. Finally, we examine plausible channels through which access to the network operated, by a closer examination of manufacturing industries, property values, and local post offices in 1870.

3.1 Empirical Strategies

In order to test if access to the railroad network led to a surge in population, we compare cities with and without access to the network using a difference-in-differences approach. We regress the population P of city $i = 1, ..., 81$ in year $t = 1840, 1855, 1870$ on the indicator *Rail* that takes the value one in $t = 1870$ for all cities with access to the network by 1870, and zero for all other cities and periods, using the estimating equation:

$$\ln P_{ijt} = \alpha_i + \theta_{jt} + \lambda_t + \delta Rail_{it} + \varepsilon_{ijt} \quad (1)$$

We include a city fixed effect (α_i) that capture time invariant factors, potentially correlated with gaining access to the network, a period fixed effect (λ_t) that capture the fact nearly all cities expanded over this period, and a region-by-period fixed effect (θ_{jt}) that takes into account shocks common to all cities in region $j = 1, ...8$.[19]

Measuring the impact of railroads over a 15-year period allows firms and migrants to respond to the changes brought about by the railroad, and also reduces concerns about railroad construction resulting in a temporary local economic boom (e.g., due to employment of local navvies) that could affect our estimates. Identification in this setting demands that in the absence of railroad construction, cities that did and did not gain access to the railroad network would have grown at similar rates.[20] This cannot be tested directly, but we have shown above (see Table 1) that cities that gained access to the network did not grow faster prior to its construction, suggesting that this assumption is not violated.

[19]We include an indicator for each of the eight National Areas (*Riksområden*) interacted with period dummies. National Areas are aggregated from the 24 counties, as defined by historical administrative boundaries.
[20]In robustness checks presented in Appendix B we allow for differential trends across cities, which yields qualitatively similar results.

Another concern is heterogeneity in the effect of access to the network. Cities with initially higher levels of manufacturing employment, for example, may benefit more from getting a railroad than cities without any industry which would then be reflected in our estimate of δ. Here we rely on estimating equation (1) in our sample that is balanced on pre-railroad characteristics that eliminates this source of bias. Standard errors are clustered at the city-level in all specifications, allowing for arbitrary patterns of heteroscedasticity and serial dependence (Bertrand et al. 2004).

3.1.1 IV Strategy

We complement our difference-in-differences specification with an instrumental variable (IV) strategy to alleviate concerns about endogenous placement of lines. Our instruments draw upon the two existing plans of the network, as described in section 2.3.1 and 2.3.2, and approximate low-cost routes between major cities.

Von Rosen's 1845 and Ericson's 1856 plan of the network are valid instruments as they were not designed to connect cities with better preconditions for growth, were conceived under minimal political influence, and were dated *prior* to when railroad construction began. A separate regression of the annual percentage population growth between 1840 and 1855 on an indicator taking the value one for cities present in von Rosen's and Ericson's plans yields a coefficient of -0.08 (s.e. = 0.21) and 0.16 (s.e. = 0.24) respectively. Cities that were included in these plans therefore did not grow faster (the estimated difference is close to zero and statistically insignificant) consistent with the qualitative evidence discussed above. But there do exist a positive and statistically significant first stage relationship between these plans and actual railroad lines in place by 1870.[21] Based on these plans we construct an instrument that corresponds to a binary indicator taking the value one for cities included in each plan respectively.

We also create an instrument based on low-cost routes between Stockholm and the other central terminal points (Gothenburg and Malmö, the northern regions, and the Norwegian border), that we approximate by connecting them by "straight lines" (see Figure 1).[22] This instrument is based on the intuition that when building a railroad line to connect, for example, Stockholm and Gothenburg, cities located along the shortest (and therefore approximately the cheapest) route between these cities will exogenously gain access to a railroad. We then create a 10 km buffer zone around each of these lines, motivated by the fact that small

[21] A regression of an indicator of having access to the railroad network by 1870 on an indicator for being included in the von Rosen and Ericson plans yield a coefficient of 0.48 (s.e. = 0.10) and 0.53 (s.e. = 0.11) respectively.

[22] This strategy follows Banerjee et al. (2012) that examine the impact of transport infrastructure on economic growth in contemporary China. They exploit the fact that early railroad lines in China tended to be constructed along a straight line between the Treaty Ports, established following the Treaty of Nanking in 1842, and historically important cities such as Beijing, Taiyuan, and Chengdu.

deviations are less costly. Our instrument is a binary variable, taking the value one for all cities located in the buffer zone of these straight lines.[23]

3.1.2 Placebo Lines

We use lines that were planned but not constructed by 1870 and lines that were built after 1870 as the basis for a placebo strategy.[24] From von Rosen's 1845 and Ericson's 1856 plans we include all lines that were proposed, but not built by 1870. Several of these lines were surveyed and most were constructed after 1870, suggesting that they initially were assigned on similar grounds as lines that were actually built. In addition, we use all lines that were actually constructed between 1870 and 1880. If there are unobserved factors that correlate with gaining access to the railroad network, or issues of reverse causality, these three sets of placebo lines are likely to reflect the magnitude and direction of this bias. Conversely, if our estimates are picking up the causal effect of gaining access to the railroad network we would expect the estimated effects for these lines to be close to zero.

3.2 Results

3.2.1 Population

Table 2 presents our estimates of equation (1), documenting that cities that gained access to the railroad network prior to 1870 grew significantly larger between 1855 and 1870. Our baseline estimate in column 1 suggest that access to the network led to a population increase of 26% on average.[25] This effect is statistically significant at the 1% level. Taking into account region-specific shocks, such as differential regional migration patterns, does not affect our estimates in a meaningful way (column 2). Similarly, balancing the sample on pre-railroad characteristics produces an nearly identical estimate to that in column 1, suggesting that our findings are not driven by observable differences between cities with and without access to the railroad network (column 3).

Columns 4-6 report IV estimates, using the two plans of the network and our low-cost route instrument to predict actual railroad lines in place by 1870, that corroborate our baseline estimate and suggest an even larger impact of the railroad. Reassuringly, point estimates are close to zero and statistically insignificant for the two sets of railroad lines that were included in von Rosen's 1845 and Ericson's 1856 proposal, but not constructed by

[23]Recall that we always exclude the endpoints (Stockholm, Gothenburg, and Malmö) from the sample.

[24]We adopt this strategy from Donaldson (2012) that examine the impact of railroads in colonial India and exploit the four-stage planning hierarchy of Indian railroads and three major proposals, as the basis for a placebo strategy.

[25]Throughout the paper we calculate percentage effects as $\left(e^{\delta}-1\right)\cdot 100$. Using the consistent (and almost unbiased) estimator suggested by Kennedy (1981) for semi-logarithmic equations with independent binary indicators $(e^{\delta-1/2V[\delta]}-1)$ yields similar results.

1870, and the set of lines constructed in the 1870s (columns 7-9). This provides compelling evidence in favor of our exclusion restriction, supporting a causal interpretation of the effect of access to the network on population growth.

A plausible objection is that our sample is relatively small and that our results may be sensitive to outliers or other specifics. In Appendix B, we show that our results are robust to excluding all large cities (>75th percentile), excluding all small cities (<25th percentile), excluding all terminal cities, allowing for differential effects for public and private lines, and including city- and region-specific linear trends.

Overall, these results suggest that access to the railroad network was associated with a substantial increase in population, but do not identify the underlying mechanisms. In the next section we examine the aggregate impact of the railroad and in section 3.2.3 we proceed to explore potential mechanisms.

[Table 2 about here.]

3.2.2 Evaluating the Aggregate Impact: "Removing" all Railroads in 1870

This section presents a simple back-of-the-envelope calculation to evaluate the aggregate contribution of railroads to changes in urbanization and urban growth between 1855 and 1870. To this end, we construct a counterfactual scenario in the spirit of Fogel (1964) where we "remove" all railroads that had been constructed by 1870.

Table 3 provides the intermediate steps of our calculations. Rows 1-3 present the total population in 1870 and the urban population in 1855 and 1870 respectively.[26] In 1855, the number of urban dwellers were 379,539, increasing to 539,649 by 1870. This corresponds to an aggregate urban growth of 42%, resulting in a level of urbanization of close to 13% in 1870 (rows 4 and 5).

To obtain the urban population consistent with no railroads being built, we subtract our baseline estimate of the effect of railroads on urban population (Table 2, column 1) from the actual log population of each city with access to the railroad network in 1870, and sum over all cities (row 6).[27] This implies that the urban population would be 459,640 had no railroads been constructed. In this counterfactual scenario, the level of urbanization in 1870 would decrease to 11% and the rate of urban growth between 1855 and 1870 would slow to 21% (rows 7 and 8). In other words, the level of urbanization and aggregate urban growth would decrease by 15% and 50% respectively.[28]

[26]Throughout this section we include *all* cities that existed in 1855 and 1870 respectively.

[27]Weighting the regression by population, so that it more properly reflects the average increase in urban population, produces a similar estimate (0.224 compared to 0.234 from the unweighted regression reported in Table 2, column 1). In practice, this has little impact on our estimates of the aggregate contribution of railroads.

[28]These calculations are based on the data in rows 4, 5, 7, and 8. The contribution to urban growth is $(12.9 - 11.0)/12.9 = 0.15$ and the contribution to aggregate urban growth is $(42.2 - 21.1)/42.2 = 0.50$.

These are economically meaningful effects, taking into account that only 1,727 of the 16,886 km network, at its peak size, had been constructed by 1870 (Nicander 1980). Although these results come with several caveats since we are ignoring general equilibrium effects and investments in other means of transportation had railroads not been constructed, they do suggest that even relatively small investments in transport infrastructure can have important effects on the aggregate patterns of urban growth.

[Table 3 about here.]

3.2.3 Cross-Sectional Evidence on Mechanisms

In the previous two sections we documented that cities that gained access to the railroad network grew substantially larger. This section aims to uncover potential mechanisms underlying this expansion by a closer examination of manufacturing industries, housing and land prices, and the activity of local post offices.

Because data is not available for the pre-railroad period, we are confined to estimate the impact of the railroad based on a cross-section of cities in 1870. Our estimating equation is:

$$Y_{ij}^{1870} = \gamma_j + \delta Rail_i + X_i\beta_i + \varepsilon_{ij} \tag{2}$$

where Y denotes an outcome (such as the income of a local post office or the average size of manufacturing establishments in city i in 1870) and $Rail$ is a binary indicator that equals one for all cities with access to the railroad network by 1870. We also include a set of region fixed effects (γ_j) and a vector of control variables (X_i), further specified below. Including a set of region fixed effects soaks up potentially important regional variation in alternative means of transportation, income, and natural endowments and ensures that identification of the effect of access to the network (δ) comes from within-region variation.

Manufacturing Industries Table 4 presents estimates of equation (2), documenting that access to the railroad network was associated with an overall expansion and modernization of industrial activity. The regressions control for (i) log 1870 population and market potential (ii) binary indicators for direct access to the sea or one of the great lakes (iii) the percentage share of population employed in manufacturing in 1855 (iv) a full set of region fixed effects.

As reported in column 1, the share of population employed in manufacturing was on average 2.8 percentage points higher in cities with access to the railroad network by 1870. This estimate reflect the increase in industrialization over this period, since we control for manufacturing employment in 1855. Manufacturing workers were not only more plentiful, but also displaced artisanal workers in relative terms (column 2). Although artisanal employment expanded in tandem with the diffusion of the factory system due to increasing demand for custom-made tools and machines (Schön 2010), this relative displacement suggests that railroads indirectly may have promoted a deskilling of the local labor force and the transition from artisanal production to the factory.

Table 4, Panel B, explores how manufacturing establishments differed across cities with and without access to the network. Establishments more commonly belonged to incorporated firms as opposed to sole proprietors (column 3) and were more than twice as large in cities with access to the network (columns 4 and 5).[29] More generally, this suggests that railroads contributed to the increase in the average size of manufacturing establishments during this period (Gårdlund 1942).[30] Railroads also lowered the cost of transporting imported coal, further fueling an increase in the size of establishments by promoting the use of steam engines. Consequently, establishments in cities with access to the railroad network used significantly more steam engines (column 6).

These results suggest that economies of scale were an important rationale for industrial expansion in cities with access to the railroad network, plausibly by widening the markets for local firms' output and lowering costs of obtaining raw materials and other inputs.[31]

[Table 4 about here.]

Housing & Land Prices Table 5 reports our estimates of equation (2) using average housing and land prices as outcomes for all 63 cities that reported both.[32] The regressions control for (i) log 1870 market potential (ii) binary indicators for direct access to the sea or one of the four great lakes (iii) a full set of region fixed effects.

In cities with access to the network, average housing prices were about 160% higher and average land prices were more than 90% higher (columns 1 and 2). Although we cannot rule out that these differences reflect some unobserved factor such as quality of housing, underlying differences in soil quality, or the presence of other amenities, the magnitude of our estimates imply that these unobservables would have to be substantial to explain away the observed differences.

[29] The fact that the average difference in the size of establishments is substantially larger when measured as average gross output (i.e., including intermediate goods) per establishment (column 5) than when measured as the number of workers per establishment (column 4) suggest that cities with access to the railroad network specialized in production of goods where intermediates, that likely had to be transported, constituted a large share of the gross value of output.

[30] Manufacturing establishments, however, remained characteristically small despite this increase in average size: an establishment in a city with access to the railroad network employed on average 28 workers in 1870.

[31] Modig (1971) documents that backward linkages from the railroad sector was of limited importance to the domestic industry: 86% of the rail and 95% of the coal used were imported, and the engineering industry delivered only about 10% of its output to the railroad sector.

[32] We rely on the taxed value of housing and land as a proxy for prices, since data on actual housing and land prices are not available. Six cities with access to the network and 12 cities without access did not report the taxed value of housing or land. See Appendix A for further description of this data.

These results therefore suggest that productivity gains associated with access to the railroad network were reflected in property values and that these gains likely were substantial already by 1870.

[Table 5 about here.]

Local Post Offices Table 6 presents estimates of equation (2) for seven different outcomes for local post offices, documenting that post offices in cities with access to the railroad network generated more revenue and distributed substantially more information. The regression control for (i) log 1870 population and market potential (ii) the number of postal roads that emanated from each city (iii) binary indicators for having direct access to the sea or one of the four great lakes (iv) a full set of region fixed effects. Controlling for postal roads and access to water transport effectively controls for the alternative means of postal transportation (stagecoaches and boats).

Column 1 documents that the total income of a post office was on average 24% higher in a city with access to the railroad network. Similarly, the sale of stamps - an important source of revenue - was almost 38% higher (column 2). Although the estimated difference in the profitability of post offices in column 3 is positive, although statistically insignificant, it is perhaps more telling that all of the post offices that made a loss (12% of all post offices) were located in cities without access to the railroad network.

Columns 4-7 examine the distribution of different types of information. Inhabitants in cities with access to the network sent around 20% more mail and parcels (columns 4 and 5). Circulation of newspapers was also higher: inhabitants of cities with access to the railroad network consumed more than twice as many foreign newspapers and about 10% more domestic newspapers, although this latter difference is estimated with a large standard error (columns 6 and 7).

Railroads therefore plausibly increased the rate of information diffusion. Although speculative, this should have provided firms with timely updates on market movements as well as facilitated matching on an increasingly national labor market (Lundh et al. 2005). Foreign newspapers and periodicals were further important as they spread technological information that "practically removed the veil of secrecy in which new techniques and processes used to be wrapped" (Heckscher 1954, p.212). Although elusive to quantify, this plausibly contributed to economic expansion in this period and beyond.

[Table 6 about here.]

4 The Long-Term Impact of Railroads (1855-2010)

This section provides evidence that cities that gained access to the network in the first wave of railroad expansion, between 1855 and 1870, are significantly larger today compared to cities that did not. We evaluate different explanations and hypothesize that long-term persistence is driven by successive investments in infrastructure over the 20th century.

4.1 Empirical Strategy

Our empirical approach centers around comparing the population of cities with and without access to the railroad network by 1870, on a decade-by-decade basis over the last 200 years. The estimating equation takes the following form:

$$\ln P_{it} = \alpha_i + \lambda_t + \delta_t Rail_i + \varepsilon_{it} \tag{3}$$

where P is the population in city $i = 1, ..., 81$ in year $t = 1800, ..., 2010$, and $Rail$ is an indicator that equals one for all cities with access to the railroad network by 1870. We include a full set of city (α_i) and decade fixed effects (λ_t). We are interested in the coefficient δ that is allowed to vary by decade. This coefficient returns the average difference in log population between cities with and without access to the network at the end of the first wave of railroad expansion, relative to the year 1855 that we omit. Our identifying assumption implies that there should be no difference prior to the railroad network was constructed ($\delta_{t=1800} \approx \delta_{t=1810} \approx ... \approx \delta_{t=1855} \approx 0$), whereas we expect to find a positive effect after construction had taken place ($\delta_{t>1855} > 0$). Standard errors are clustered at the city-level to allow for arbitrary patterns of heteroscedasticity and serial dependence.

4.2 Results

4.2.1 Population

Figure 2 graphs our results, where solid lines correspond to the differential effect for cities with and without access to the railroad network by 1870 (δ_t) from equation (3) and dashed lines correspond to a 95% confidence interval. Panel A reports estimates from our baseline sample and Panel B from our sample balanced on pre-railroad characteristics.

There were no significant difference in terms of population between cities with and without access to the railroad network prior its construction, consistent with our identifying assumption. After railroad construction began, in 1855, we observe a positive difference in the population of cities with and without access to the network that turns statistically significant by 1870, consistent with the results in section 3. Cities with early access to the network continued to grow faster over the first half of the 20th century. After a period of relative decline between 1950 and 1970 - a period characterized by the breakthrough of highway construction and motoring - these cities are on average 51% (0.42 log points) larger today, measured relative to 1855 and compared to cities that did not gain access to the railroad network in the first wave of expansion. This difference is statistically significant at the 5% level. When we compare initially similar cities (Panel B), it suggests an even larger long-term effect of 62% (0.48 log points). Because this reduces our sample size considerably, the standard errors are larger. The difference is, however, still statistically significant at the 10% level.

[Figure 2 about here.]

4.2.2 Long-Term Impact on the Urban Hierarchy

Another way to convey our results is to estimate the impact on the urban hierarchy, simply defined as the ranking of cities by their size. We sort all cities by their size S_i in year t, such that $S_i^1 > S_i^2 > ... > S_i^{81}$, and assign each city a rank, increasing from largest to smallest. Regressing the rank in 2010 on a binary indicator for having access to the railroad network by 1870, controlling for each city's rank in 1855, yields a slope coefficient of -11.3 (s.e. = 4.0).[33] In other words, cities that became connected to the network in the first wave of railroad expansion are to be found on average 11 steps higher in the urban hierarchy today.[34] This suggests that the first wave of railroad expansion substantially reshaped the urban hierarchy.

4.2.3 Channels of Persistence

This subsection empirically evaluates potential explanations for the fact that cities that gained access to the network during the first wave of railroad expansion were substantially larger by 2010, compared to cities that did not. Our main explanation is that access to the railroad network in the first wave of railroad expansion solved a coordination problem of future infrastructure investment. Once these initial lines were in place, additional lines constructed after 1870 were routed through this initial set of cities, entrenching their roles as nodes in the network.[35] While it may have made sense to connect a large number of different cities prior to the network was constructed - as explicitly discussed at the time, and

[33] A similar regression in our balanced sample yields a coefficient of 11.2 (s.e. = 5.2).

[34] There are several illuminating trajectories of individual cities. Karlskrona, the second largest city in our sample in 1855, that did not gain access to the railroad network during the first wave of expansion, fell to place 23 by 2010. Skövde, the 58th largest city in 1855, and located on the Western trunk line, had by 2010 reached place 24. Södertälje, similarly located on the Western trunk line, rose from place 55 in 1855, to being the 13th largest city in our sample by 2010. See Heckscher (1907, pp.129-130) for a discussion on historical changes in the urban hierarchy.

[35] Westlund (1992, p.67) argues that there were little change in cities' relative nodality in the 20th century once the road and railroad networks had been established. The important 'revolution' was the early period of railroad expansion.

manifested in the different proposals - the benefits of building a line to a city already part of the network was higher than building one to a city that was not, *once* the network had been constructed. These first lines therefore gave rise to path dependence in future infrastructure investments.[36]

Following Bleakley and Lin (2012) we contrast this explanation with mechanisms working through sunk investments and external economies. If large sunk investments were made in cities that gained access to the railroad network early, we would expect to find persistence over the medium term. For example, investments in housing are slowly depreciating and during this depreciation period people and firms might choose to locate in a city with ample housing supply, rather than incur the cost of construction at another location. The fact that cities with early access to a railroad line declined relatively after 1950 (see Figure 2) is consistent with some form of slowly depreciating asset, or with the relative decline of railroads as a mode of transport. External economies may similarly be important if the growth of manufacturing in these cities gave rise to external economies derived from input-output linkages, thick labor markets, and knowledge spillovers as emphasized by Marshall (1890), or cross-sectoral spillovers as emphasized by Jacobs (1969). If external economies were important, firms may choose to stay in these cities even though a concentration of firms would bid up factor prices.[37]

In order to evaluate the plausibility of these explanations we run long-differenced regressions on the form:

$$\triangle \ln P_{ij}^{t} = \gamma_j + \delta Rail_i + \theta Z_i^t + \varepsilon_{ij} \tag{4}$$

where $\triangle \ln P_{ij}^{t} \equiv \ln P_{ij}^{2010} - \ln P_{ij}^{1855}$, and P is the population of city i in region j in year 2010 and 1855 respectively. *Rail* is a binary indicator that equals one for all cities with access to the railroad network by 1870. We include a set of region fixed effects (γ_j) and condition on Z_i^t, corresponding to an intermediating variable in some year t. Here we are interested in how conditioning on Z_i^t affects the magnitude and statistical significance of our estimated effect of early access to the railroad network (δ).

Column 1 of Table 7 provides the baseline impact ($\delta = 0.58$ log points) of early access to the railroad network, obtained from equation (4) without any intermediating variable. In each remaining column we then add one potential intermediating channel (Z_i^t). In the following three subsections we discuss how the inclusion of these intermediating variables affect the estimated long-term impact of early access to the railroad network.

[36]This mechanism is implicit in recent work on the contemporary impact of infrastructure that rely on the first stage relationship between historical and contemporary levels of infrastructure. See, for instance, Duranton and Turner (2012) and Banerjee et al. (2012).

[37]See Duranton and Puga (2004) and Rosenthal and Strange (2004) for an overview.

Sunk Investments We proxy for sunk investments by the stock of old housing units, the presence of a grammar school, and the number of telephones per capita. From the housing census of 1939 we have obtained the number of housing units constructed prior to 1880 - roughly corresponding to the period of early railroad expansion - that were still in use in the late 1930s for each city in our sample. For each city where a grammar school was present in 1880 we code a binary indicator that equals one if a school was present. To proxy for sunk investments in communications infrastructure, we calculate the number of telephones per inhabitant in 1900.[38]

In columns 2-4 of Table 7 we condition on each of these measures. When conditioning on sunk investments, the effect of early access decreases by at most 14% (column 3), and the early access indicator remains statistically significant at the 5% level in all three cases. Although these sunk investments were positively correlated with long-run population growth, it is therefore unlikely that they account for any significant fraction of the effect that is attributed to early access to the railroad network.

External Economies We proxy for external economies by a measure of sectoral specialization, manufacturing employment, and employment in the transport sector. To measure the diversity of sectoral employment in each city we calculate a Herfindahl–Hirschman index (HHI) of sectoral specialization in 1930.[39] We use the manufacturing share in total employment in 1930 as a rough proxy for the scope for external economies. Lastly, we include the share of the population employed in the transport sector. This serves as a check on the argument that early access to the railroad network simply may be running through employment opportunities in railroad-related sectors.

In columns 5-7 of Table 7 we condition on each of these measures. Sectoral specialization is positively correlated with long-term population growth, but the impact of early access remains largely unaffected. Manufacturing employment is also significantly correlated with long-run population growth. When we condition on this variable, the effect of early access to the railroad network decreases by roughly a third, but retains its statistical significance at the 5% level. It is therefore unlikely that the effect of early access is singularly running through manufacturing employment, or being explained by external economies more generally. Similarly, controlling for the share of the population employed in the transport sector has little impact on the early access indicator. Persistence does therefore not reflect differences in employment opportunities in this sector.

[38]See Appendix A for a more detailed description of our data.

[39]We calculate the Herfindahl–Hirschman index as $\mathrm{HHI_i} = \sum \mathrm{s_{si}^2}$ where s is the share of total employment in sector s in city i, across five sectors (agriculture, industry, trade, transport, and services). If all employees work in one sector - i.e., if a city is completely specialized - the index takes the value one.

Modern Infrastructure To measure modern infrastructure networks we rely on maps of the mid-20th century railroad and road networks.[40] Based on these maps we calculate the number of "rays" that emanate from each city, akin to the method used by Baum-Snow (2007). We think of this as a measure of the *cumulative* investments in infrastructure and a measure of the contemporary centrality of a city in each respective network. If early access to the railroad network coordinated future infrastructure investments to these cities we would expect them to be more central in the latter incarnations of these networks. Indeed, our data shows that cities with early access to the railroad network had on average 80% more railroad rays and 50% more highway rays emanating from them in the mid-20th century.

In column 8 of Table 7 we condition on the number of railroad rays in the mid-20th century. The estimated effect of early access to the network decreases by more than 60% and it is no longer statistically significant at conventional levels. A large share of the impact of early access to the railroad network is therefore attributable to differences in centrality in the modern railroad network. When conditioning on the number of rays in the mid-20th century highway network the effect of early access to the railroad network decreases by about 30%, although it retains its statistical significance (column 9). A large share of the effect of early access to the network therefore primarily runs through the later incarnations of the railroad network.

[Table 7 about here.]

5 Conclusions

We have shown that during a first wave of railroad construction, between 1855 and 1870, cities that gained access to the network experienced an economic expansion: their population increased and they became more industrialized. Cities with early access to the railroad network continued to grow faster for a better part of the 20th century. Today they are considerably larger compared to initially similar cities that only gained access to the network later. Our main explanation for this long-term persistence is that successive infrastructure investments over the 20th century was directed toward cities with early access to the railroad network.

Our results strongly suggest that railroads were a causal factor in promoting economic development in 19th-century Sweden, and that railroads that were built "ahead of demand" were capable of igniting a process of sustained economic development. More generally, we argue that historical investments in infrastructure ignited a path dependent process, that shapes the economic geography of Sweden today. This constitutes an intuitive yet unex-

[40] Specifically, we use maps of the road network as of 1957 and the railroad network as of 1968. These networks are very similar to those in existence today. See Appendix A for a further description of the data.

plored mechanism that likely is at work in many countries. Understanding how historical investments in infrastructure shapes local development trajectories and disparities today constitutes an area that merits future work.

References

Anderson, J. E. and van Wincoop, E. 2004. Trade costs. *Journal of Economic Literature 42,* 3 (September), 691–751.

Atack, J., Bateman, F., Haines, M., and Margo, R. 2010. Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850–1860. *Social Science History 34,* 2, 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, A., Duflo, E., and Qian, N. 2012. On the road: Access to transportation infrastructure and economic growth in China. *NBER Working Paper 17897*.

Baum-Snow, N. 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, H. and Lin, J. 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.

Bolt, J. and van Zanden, J. L. 2013. The First Update of the Maddison Project; Re-Estimating Growth Before 1820. *Maddison-Project Working Paper WP-4, University of Groningen, January 5.*

Bosker, M., Brakman, S., Garretsen, H., and Schramm, M. 2007. Looking for Multiple Equilibria When Geography Matters: German City Growth and the WWII Shock. *Journal of Urban Economics 61,* 1, 152–169.

Cantoni, D. 2011. The economic effects of the protestant reformation: testing the weber hypothesis in the german lands. *Mimeo.*.

Chandler, A. 1965. The Railroads: Pioneers in Modern Corporate Management. *The Business History Review*, 16–40.

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

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.

Dittmar, J. E. 2011. Information technology and economic change: the impact of the printing press. *The Quarterly Journal of Economics 126,* 3, 1133–1172.

Donaldson, D. 2012. Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *Mimeo*.

Donaldson, D. and Hornbeck, R. 2012. Railroads and American Economic Growth: New Data and Theory. *Mimeo*.

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 Ante-bellum Economy.* Vol. 127. Cambridge: Harvard University Press.

Fogel, R. 1964. *Railroads and American Economic Growth.* Baltimore: John Hopkins Press.

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

Gadd, C.-J. 2000. *Den Agrara Revolutionen (Det Svenska Jordbrukets Historia Del 3).* Natur och kultur/LTs förlag,.

Gårdlund, T. 1942. *Industrialismens Samhälle.* Tidens Förl.

Glaeser, E. L. and Gottlieb, J. D. 2009. The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in United States. *Journal of Economic Literature 47,* 4, 983–1028.

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

Harris, C. D. 1954. The Market as a Factor in the Localization of Industry in the United States. *Annals Of The Association Of American Geographers 44,* 4, 315–348.

Harris, J. R. and Todaro, M. P. 1970. Migration, unemployment and development: a two-sector analysis. *The American Economic Review 60,* 1, 126–142.

Heckscher, E. 1907. *Till Belysning af Järnvägarnas Betydelse för Sveriges Ekonomiska Utveckling*. Stockholm: Centraltryckeriet.

Heckscher, E. 1954. *An Economic History of Sweden*. Cambridge: Harvard University Press.

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

Hornung, E. 2012. Railroads and micro-regional growth in prussia. Mimeo.

Jacobs, J. 1969. *The Economy of Cities*. New York: Random House.

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

Kaijser, A. 1999. The helping hand. in search of a swedish institutional regime for infrastructural systems. *Andersson-Skog, L. & Kranz, O.(red.), Institutions in the transport and communications industries (Canton, Ma)*, 223–44.

Keller, W. and Shiue, C. H. 2008. Institutions, Technology, and Trade. *NBER Working Paper 13913*.

Kennedy, P. E. 1981. Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations. *American Economic Review 71,* 4.

Krantz, O. and Schön, L. 2007. Swedish historical national accounts 1800-2000.

Krugman, P. 1991a. Increasing Returns and Economic Geography. *Journal of Political Economy 99*, 483–499.

Krugman, P. R. 1991b. *Geography and Trade*. MIT press.

Kungl. Järnvägsstyrelsen. 1956. *Sveriges järnvägar hundra år.* Kungl. Järnvägsstyrelsen, Stockholm.

Lee, E. S. 1966. A theory of migration. *Demography 3,* 1, 47–57.

Leunig, T. 2006. Time is money: a re-assessment of the passenger social savings from victorian british railways. *Journal of Economic History 66,* 3, 635.

Lundh, C., Schön, L., and Svensson, L. 2005. Regional wages in industry and labour market integration in sweden, 1861–1913. *Scandinavian Economic History Review 53,* 3, 71–84.

Marshall, A. 1890. *Principles of Economics.* London: MacMillan for the Royal Economic Society, 1961.

Martí-Henneberg, J. 2013. European Integration and National Models for Railway Networks (1840–2010). *Journal of Transport Geography 26,* 0, 126 – 138.

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

Mitchell, B. R. 1975. *European historical statistics, 1750-1970.* Macmillan London.

Modig, H. 1971. *Järnvägarnas efterfrågan och den svenska industrin 1860-1914.* Vol. 8. Läromedelsförlagen.

Montgomery, A. 1947. *Industrialismens Genombrott i Sverige.* Almquist & Wiksell.

Murphy, K. M., Shleifer, A., and Vishny, R. W. 1989. Industrialization and the Big Push. *Journal of Political Economy 97,* 5 (October), 1003–26.

Nicander, E. 1980. *Järnvägsinvesteringar i Sverige 1849-1914.* Ekonomisk-Historiska Föreningen, Lund: Tryckbaren.

Nilsson, L. 1992. *Historisk tätortsstatistik. D. 1, Folkmängden i administrativa tätorter 1800-1970.* Stads-och kommunhistoriska institutet.

Nunn, N. and Qian, N. 2011. The potato's contribution to population and urbanization: Evidence from a historical experiment. *The Quarterly Journal of Economics 126,* 2, 593–650.

O'Brien, P. 1977. *The New Economic History of the Railways.* Croom Helm.

O'Brien, P. 1983. *Railways and the Economic Development of Western Europe, 1830-1914.* Macmillan.

O'Rourke, K. and Williamson, J. 1995a. Education, Globalization and Catch-Up: Scandinavia in the Swedish Mirror. *Scandinavian Economic History Review 43,* 3, 287–309.

O'Rourke, K. and Williamson, J. 1995b. Open Economy Forces and Late Nineteenth Century Swedish Catch-Up. A Quantitative Accounting. *Scandinavian Economic History Review 43,* 2, 171–203.

Pollard, S. 1981. *Peaceful Conquest: The Industrialization of Europe, 1760-1970.* Oxford University Press Oxford.

Prado, S. 2010. Fallacious convergence? williamson's real wage comparisons under scrutiny. *Cliometrica 4,* 2, 171–205.

Redding, S. J. and Sturm, D. M. 2008. The Costs of Remoteness: Evidence from German Division and Reunification. *American Economic Review 98,* 5 (September), 1766–97.

Redding, S. J., Sturm, D. M., and Wolf, N. 2011. History and Industry Location: Evidence from German Airports. *Review of Economics and Statistics 93,* 3, 814–831.

Roback, J. 1982. Wages, rents, and the quality of life. *The Journal of Political Economy*, 1257–1278.

Rosen, S. 1979. Wage-based indexes of urban quality of life. *Current issues in urban economics 3.*

Rosenbaum, P. R. and Rubin, D. B. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika 70,* 1, 41–55.

Rosenstein-Rodan, P. N. 1943. Problems of industrialisation of eastern and south-eastern europe. *The Economic Journal 53,* 210/211, pp. 202–211.

Rosenthal, S. S. and Strange, W. C. 2004. Evidence on the Nature and Sources of Agglomeration Economies. *Handbook of regional and urban economics 4*, 2119–2171.

Rydfors, A. 1906. *Statens järnvägar 1856-1906: historisk-teknisk-ekonomisk beskrifning i anledning af statens järnvägars femtioåriga tillvaro, utgifven på Kungl. Maj: ts nådiga befallning af Järnvägsstyrelsen.* Centraltryckeriet, Chapter Politisk Historik, 22–213.

Sandberg, L. G. 1979. The Case of the Impoverished Sophisticate: Human Capital and Swedish Economic Growth Before World War I. *Journal of Economic History 39,* 1, 225–241.

Schön, L. 1979. *Från Hantverk till Fabriksindustri: Svensk Textiltillverkning 1820-1870*. Vol. 9. Arkiv för studier i arbetarrörelsens historia.

Schön, L. 2010. *Sweden's Road to Modernity: An Economic History*. SNS förlag Stockholm.

Sjaastad, L. A. 1962. The costs and returns of human migration. *The journal of political economy 70,* 5, 80–93.

Sjöberg, A. 1956. *Sveriges järnvägar hundra år*. Kungl. Järnvägsstyrelsen, Stockholm, Chapter Järnvägarna i svenskt samhällsliv, Några huvuddrag i deras utveckling, 1–159.

Stads och Kommunhistoriska Institutet. 2012. Electronic Source. Available at: http://www2.historia.su.se/urbanhistory/cybcity/, Accessed: 2012-10-31.

Statistiska Centralbyrån. 1969. *Historisk Statistik for Sverige Del 1: Befolkning 1720-1967*.

Storeygard, A. 2013. Farther on down the road. *Mimeo*.

Thorburn, T. 2000. *Economics of Transport: The Swedish Case, 1780-1980*. Vol. 12. Almqvist & Wiksell International.

Westlund, H. 1992. *Kommunikationer, tillgänglighet, omvandling. En studie av samspelet mellan kommunikationsnät och näringsstruktur i Sveriges mellanstora städer 1850-1970.*

Westlund, H. 1998. State and Market Forces in Swedish Infrastructure History. *Scandinavian journal of history 23,* 1-2, 65–88.

Williamson, J. G. 1995. The evolution of global labor markets since 1830: background evidence and hypotheses. *Explorations in Economic History 32,* 2, 141–196.

A Data Appendix

This appendix describes the construction of our dataset and provides detailed information on sources used. Our sample consists of all cities that held town charters prior to the railroad era. We merge two cities (Skanör and Falsterbo) that had been a under joint political rule since 1754 and formed a single municipality from 1863 and onwards. We also exclude the three major cities Gothenburg, Malmö, and Stockholm and the two insular cities Borgholm and Visby. This brings the sample size down to 81 cities, that constitute the baseline sample used in the paper.[41]

A.1 Population

For each city in our sample we collect population data at decadal intervals 1800-2010, as well as for the year 1855. We obtained our data for the period 1800-1950 from Nilsson (1992) and Stads och Kommunhistoriska Institutet (2012). For the period 1960-2010 our data was obtained from Statistics Sweden. For a small number of cities that did not hold town charters in the early 1800s, and therefore were not reported in the official statistics, we have assumed that their growth equalled the average growth of all other cities.[42]

A.2 Sectoral Employment

For the year 1855 our data on sectoral employment is based on census materials (*Tabellverkets Folkmängstabeller*), obtained from Stads och Kommunhistoriska Institutet (2012). As female employment is only sporadically reported our data only include male employees. For 1870, data on manufacturing and artisanal employment is reported by Statistics Sweden in the official industrial statistics (*Bidrag till Sveriges officiella statistik D: Fabriker och Manufakturer*). From this source we calculate the total number of manufacturing workers and artisans in each city. We also obtain data on the share of manufacturing establishments that belonged to incorporated firms, number of active establishments, gross output of the manufacturing industry, and the number of steam engines used in each city.

We also obtained the sectoral composition of employment in each city in 1930 from the population census (*Folkräkningen*), obtained from Stads och Kommunhistoriska Institutet

[41]Cities included in our baseline sample are: Alingsås, Arboga, Askersund, Borås, Eksjö, Enköping, Eskilstuna, Falkenberg, Falköping, Falun, Filipstad, Gränna, Gävle, Halmstad, Haparanda, Hedemora, Helsingborg, Hjo, Hudiksvall, Härnösand, Jönköping, Kalmar, Karlshamn, Karlskrona, Karlstad, Kristianstad, Kristinehamn, Kungsbacka, Kungälv, Köping, Laholm, Landskrona, Lidköping, Lindesberg, Linköping, Luleå, Lund, Mariefred, Mariestad, Marstrand, Nora, Norrköping, Norrtälje, Nyköping, Piteå, Sala, Sigtuna, Simrishamn, Skanör (Falsterbo), Skara, Skänninge, Skövde, Strängnäs, Strömstad, Sundsvall, Säter, Söderhamn, Söderköping, Södertälje, Sölvesborg, Torshälla, Trosa, Uddevalla, Ulricehamn, Umeå, Uppsala, Vadstena, Varberg, Vaxholm, Vimmerby, Vänersborg, Västervik, Västerås, Växjö, Ystad, Åmål, Ängelholm, Örebro, Öregrund, Östersund, and Östhammar.

[42]Results reported in the paper (Figure 2) are nearly identical when instead using an unbalanced panel.

(2012). Based on this source we calculate a Hirschmann-Herfindahl of sectoral specialization (see main text for calculation) and the share of manufacturing and the transport sector in city-level employment.

A.3 Railroads, Highways, and Postal Roads

Historical maps of the railroad network that include all lines built in each year were obtained from Statistics Sweden (*Bidrag till Sveriges officiella statistik L: Statens järnvägstrafik* 1862-1910). This is combined with modern GIS maps of the Swedish railroad network from Digital Chart of the World (http://www.diva-gis.org). Using ArcGIS, these two sources were combined to recreate the national railroad network as of 1870. We exclude all minor railroad lines that did not link up to the network. All cities were linked to this spatial layer based on the longitude and latitude of the centroid of each city.[43] In addition, we digitized the two alternative plans of the railroad network based on maps provided by Kungl. Järnvägsstyrelsen (1956, Map 1).

Maps of the 1968 railroad network were obtained from a historical atlas (*Atlas över Sverige, Svenska Sällskapet för Antropologi och Geografi, Stockholm: Generalstabens Litografiska Anstalts Förlag, Railways, Map 9: "Railway network 1968"*). Based on this map we calculated the number of railroad lines that emanated from each city. Maps of the 1957 highway network was obtained from a road atlas (*S-N Bilkarta över Sverige 1957, Generalstabens Litografiska Anstalt: Stockholm 1960*). From this source we calculated the number of major roads (*Europavägar* and *Riksvägar*) that emanated from each city.

A map of 19th century postal roads was obtained from a historical atlas (*Generatlasen*, *Inns and Stage-Coach System About 1850*, *Figure 9: "Mail-coach routes and railways in Sweden in 1868"*). From this map we calculated the number of postal roads that emanated from each city.

A.4 Housing & Land Prices

Each city had to report the value of housing and land for taxation purposes, reported by the Governor of each county and summarized by Statistics Sweden at five-year intervals (*Bidrag till Sveriges officiella statistik H: Kungl. Maj:ts befallningshavandes femårsberättelser 1856-1905*). From these reports we collect data on the taxed value of land and housing in 1870. We calculate the value of housing as the total taxed value of housing divided by the number of plots in each city, and the taxed value of land as the total taxed value divided by the land area in square km. A total of 63 (out of our 81) cities reported both the taxed value of housing and land, as Governors of some counties only reported aggregates for all cities in

[43] Longitude and latitude was obtained from: http://www.findlatitudeandlongitude.com/batch-geocode/

their county.

A.5 Post Offices

Data on local post offices were obtained from Statistics Sweden (*Bidrag till Sveriges officiella statistik M: Postverket 1870*). Data on incomes and costs is taken from *Bilaga litt. I.* and is measured in contemporary currency units (*riksdaler*). We calculate the profit of each post office as total income less total costs. Data on domestic and foreign mail is obtained from *Bilaga litt. Da.*, and is measured as the number of mails distributed on an annual basis. Our data on the number of annually distributed domestic and foreign newspapers is taken from *Bilaga litt. Dc.* Data on the number of domestic and foreign parcels (including registered mail) is obtained from *Bilaga litt. Db.*, and data on the total value of sold stamps is obtained from *Bilaga litt. H.*.

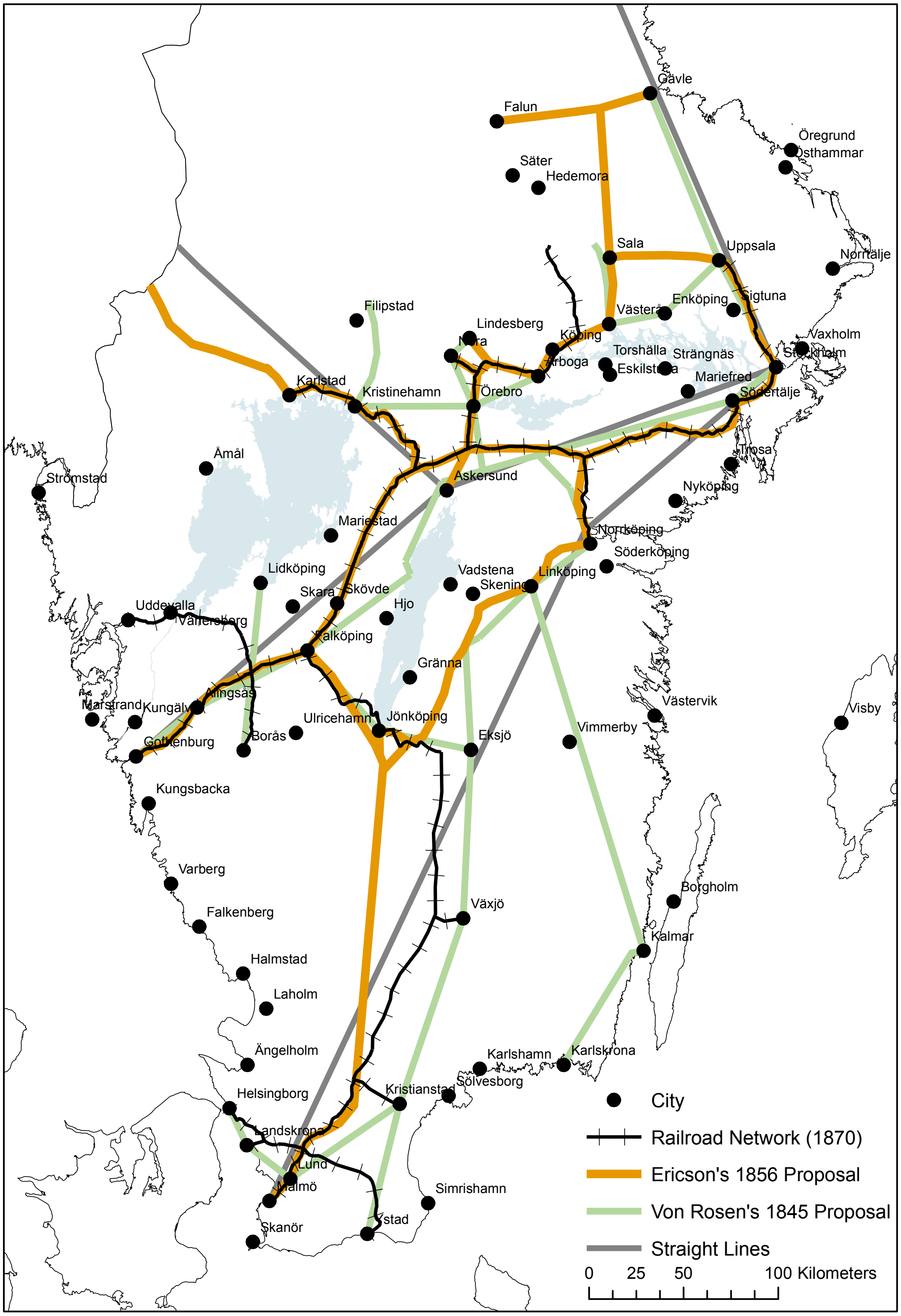
A.6 Other

From the housing census of 1939 (*Allmänna Bostadsräkningen, Tabellbilaga, Ortstabeller, 1939*) we have obtained the number of housing units constructed prior to 1880, and the total number of housing units in use in 1939. Based on the educational statistics (*Bidrag till Sveriges officiella statistik P: Undervisningsväsendet 1880-1881*) we have coded a binary indicator for whether or not a grammar school (*Allmänna Högre Läroverk*) existed in a city in 1880. From the official statistics on the telegraph network (*Bidrag till Sveriges officiella statistik. I: Telegrafväsendet 1900*) we have calculated the number of telephones per inhabitant in 1900.

B Robustness Checks

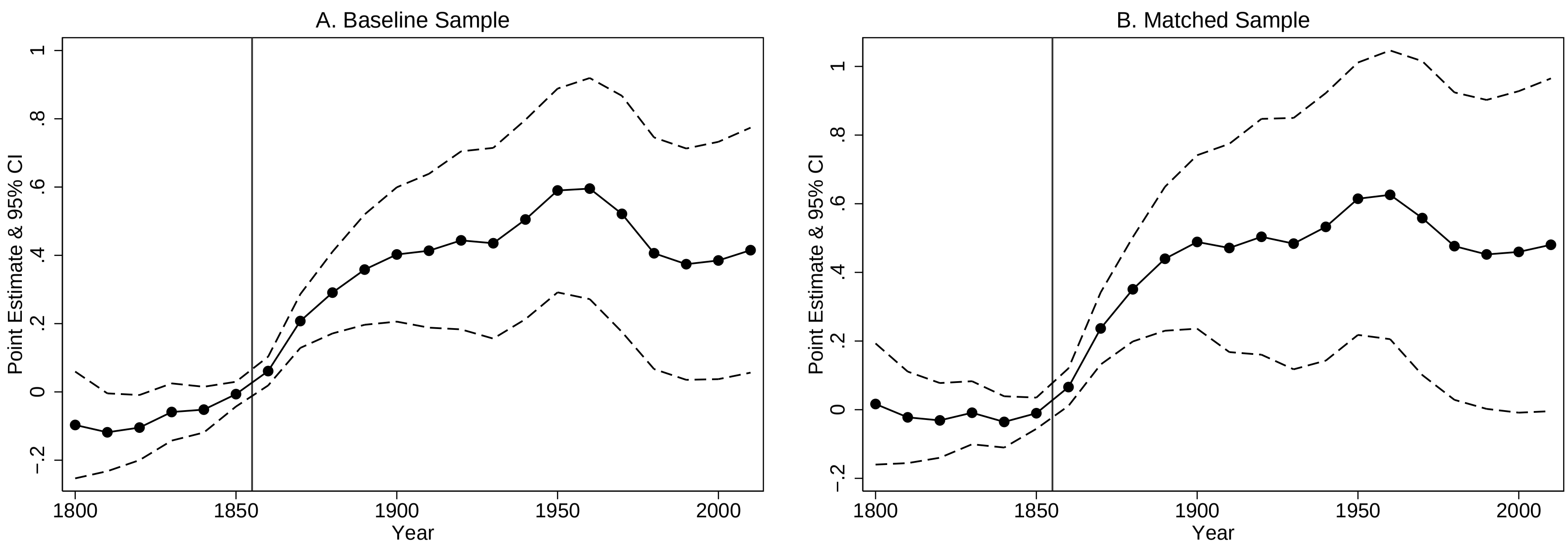
Table 8 presents robustness checks on our main results provided in Table 2, based on estimations of equation (1). Columns 1 and 2 excludes all large and small cities (above and below the 75th and 25th percentile in 1855) respectively. Column 3 excludes all terminal cities. Column 4 separates the effect for public and private lines. Column 5 includes a full set of city-specific linear trends. Column 6 includes a set of region-specific linear trends. All estimates retain their statistical significance and are of similar magnitude to those presented in Table 2.

[Table 8 about here.]



Notes: This map shows the actual railroad network as of 1870, lines proposed by Adolf von Rosen in 1845 and Nils Ericson in 1856, the four largest lakes, a set of straight lines that form the basis of our instrument, and all cities (N=86) that existed prior to when railroad construction began. Note that we exclude minor railroad lines and that cities in northern Sweden are not shown for ease of exposition. See the main text and Appendix A for a description of the underlying data.

Figure 1: The Swedish Railroad Network, 1870.



Notes: These figures plot the δ_t-coefficients from equation (3). Connected solid lines correspond to point estimates and dashed lines to a 95% confidence interval. Panel A reports estimates obtained from our baseline sample and Panel B reports estimates from our sample balanced on pre-railroad characteristics (see section 2.4.1). A solid vertical line denotes the year 1855 (the base year) when railroad construction was initiated.

Figure 2: Long-Term Impact of Early Access to the Railroad Network on Population, 1800-2010.

	A. Baseline Sample			B. Matched Sample		
	Connected (1)	Unconnected (2)	Difference (1)-(2) (3)	Connected (4)	Unconnected (5)	Difference (4)-(5) (6)
Population (ln)	8.072	7.412	0.660***	7.932	7.682	0.251
	(0.778)	(0.772)	[0.193]	(0.679)	(0.717)	[0.217]
% Population Growth, 1840-55	1.670	1.321	0.348	1.662	1.484	0.178
	(0.803)	(1.104)	[0.223]	(0.812)	(0.695)	[0.238]
Market Potential (ln)	7.706	7.574	0.132	7.692	7.733	-0.041
	(0.226)	(0.508)	[0.082]	(0.242)	(0.222)	[0.073]
Access to Sea (=1)	0.273	0.492	-0.219*	0.278	0.250	0.028
	(0.456)	(0.504)	[0.117]	(0.461)	(0.442)	[0.141]
Access to Big Lakes (=1)	0.273	0.186	0.086	0.278	0.292	-0.014
	(0.456)	(0.393)	[0.109]	(0.461)	(0.464)	[0.144]
% Industrial Employment	2.303	1.438	0.865	1.733	1.933	-0.199
	(3.021)	(1.956)	[0.687]	(1.760)	(2.582)	[0.671]
% Artisan Employment	13.925	12.943	0.981	13.873	14.114	-0.241
	(2.373)	(3.826)	[0.707]	(1.928)	(4.485)	[1.024]
% Trade Employment	2.447	2.620	-0.173	2.625	2.295	0.331
	(1.226)	(2.989)	[0.469]	(1.293)	(0.949)	[0.360]
% Service Employment	3.360	3.104	0.256	3.611	3.117	0.494
	(3.081)	(3.323)	[0.782]	(3.348)	(2.595)	[0.949]
% Shipping Employment	0.562	2.304	-1.742***	0.540	0.467	0.073
	(0.671)	(3.474)	[0.476]	(0.642)	(0.580)	[0.192]
% Military Employment	1.461	1.198	0.263	1.729	1.256	0.473
	(2.940)	(3.293)	[0.755]	(3.202)	(2.648)	[0.927]
No. of Cities	22	59	81	18	24	42

Notes: Columns 1, 2, 4, and 5 report mean pre-railroad characteristics and standard deviations (in parentheses) for cities with and without access to the railroad network by 1870. Columns 3 and 6 report difference-in-means and corresponding Huber-White standard errors (in brackets). Panel A reports characteristics for our baseline sample and Panel B reports characteristics for our sample balanced on all pre-railroad characteristics in this table and a first-order polynomial in the longitude and latitude of each city. All pre-railroad characteristics are measured in the year 1855. Sectoral employment is calculated as a percentage of total population in 1855. See Appendix A for a description of the data. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 1: Pre-Railroad Differences Between Connected and Unconnected Cities, 1855.

	Baseline			IV (2SLS)			Placebo Lines		
	Baseline	Baseline	Matched	1845 Plan	1856 Plan	Straight Line	1845 Plan	1856 Plan	Built After 1870
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Access to Network (=1)	0.234***	0.278***	0.233***	0.242***	0.341***	0.321**	0.240***	0.238***	0.261***
	(0.048)	(0.047)	(0.057)	(0.082)	(0.078)	(0.144)	(0.053)	(0.050)	(0.058)
Placebo Line (=1)							0.025	0.035	0.040
							(0.049)	(0.046)	(0.057)
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region × Period FE	No	Yes	No	No	No	No	No	No	No
Observations	243	243	126	243	243	243	243	243	243
No. of Cities	81	81	42	81	81	81	81	81	81
R-squared	0.74	0.82	0.84	0.74	0.73	0.73	0.74	0.74	0.74

Notes: This table presents estimates of equation (1) where the dependent variable is log city population. In column 3 we use a sample balanced on pre-railroad characteristics (see section 2.4.1). In columns 4-6 the instruments correspond to being included in von Rosen's 1845 and Ericson's 1856 proposal of the network, or being located on a straight line between major cities (see section 3.1.1). Columns 7 and 8 includes lines that were included in the plans of 1845 and 1856, but that were not built by 1870. Column 9 includes lines that were built after 1870. Statistical significance based on standard errors clustered at the city-level is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 2: Short-Term Impact of Access to the Railroad Network on Population, 1840-1870.

A. Observed Outcomes		Year	Calculation	Outcome
(1)	Total population	1870	-	4,168,525
(2)	Urban population	1855	-	379,539
(3)	Urban population	1870	-	539,649
(4)	Urbanization (%)	1870	(3)/(1)	12.9
(5)	Urban Growth (%)	1855-1870	$(3)/(2) - 1$	42.2
B. Counterfactual Scenario: No Railroads				
(6)	Urban population	1870	$\sum_i e_i^{\ln P_i - \delta \times Rail_i}$	459,640
(7)	Urbanization (%)	1870	(6)/(1)	11.0
(8)	Urban growth (%)	1855-1870	$(6)/(2) - 1$	21.1

Notes: This table calculates the urbanization and urban growth in the counterfactual scenario that no railroad construction would have taken place. In row 6, P is the observed population in 1870, δ is the estimated short-term impact of access to the railroad network (Table 2, column 1), and *Rail* is an indicator that equals one for all cities with access to the railroad network by 1870. Total and urban population is obtained from from Statistics Sweden (Statistiska Centralbyrån 1969, Tables 3 and 4, pp.45-46).

Table 3: Aggregate Impact of Railroads on Urbanization and Urban Growth, 1855-1870.

	Panel A. All Cities		Panel B. Cities with at Least One Manufacturing Establishment			
	Workers/Pop. (1)	Workers/Artisans (2)	Incorporated (3)	Workers/Est. (4)	Output/Est. (5)	Steam Engines/Est. (6)
Access to Network (=1)	2.753***	0.820**	8.318***	0.798***	1.309***	0.127*
	(0.484)	(0.253)	(1.965)	(0.177)	(0.310)	(0.057)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	81	81	71	71	71	71
R-squared	0.36	0.36	0.38	0.40	0.37	0.17

Notes: This table reports estimates of equation (2). Dependent variables are defined as follows: the percentage share of the population employed in manufacturing (column 1), the ratio of manufacturing workers to artisanal workers (column 2), the percentage share of establishments that are owned by an incorporated firm (column 3), the log workers per manufacturing establishment (column 4), the log output per manufacturing establishment (column 5), and the number of steam engines used per manufacturing establishment (column 6). All specifications include controls for log 1870 population, log 1870 market potential, binary indicators that equal one if a city is located on the coast or has direct access to one of the four big lakes, and the percentage of the population employed in manufacturing in 1855. Statistical significance based on standard errors clustered at the region-level is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 4: Access to the Railroad Network and Manufacturing Industries, 1870.

	Average Housing Price	Average Land Price
	(1)	(2)
Access to Network (=1)	0.942***	0.645**
	(0.148)	(0.238)
Region FE	Yes	Yes
Controls	Yes	Yes
Observations	63	63
R-squared	0.33	0.17

Notes: This table reports estimates of equation (2). The dependent variables are the log average housing price per plot and the log average land price per square km respectively. See the main text and Appendix A for a description of the data. All specifications include controls for log 1870 market potential and binary indicators that equal one if a city is located on the coast or has direct access to one of the four big lakes. Statistical significance based on standard errors clustered at the region-level is denoted by: *** p<0.01, ** p<0.05, * p<0.10.

Table 5: Access to the Railroad Network and Housing and Land Prices, 1870.

	Panel A. Revenues & Profits			Panel B. Distribution of Information			
	Income	Stamps	Profits	Mail	Parcels	Domestic News	Foreign News
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Access to Network (=1)	0.213***	0.322***	2,260.349	0.186**	0.181***	0.095	0.736**
	(0.049)	(0.053)	(1,491.806)	(0.072)	(0.049)	(0.098)	(0.274)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	81	81	81	81	81	81	81
R-squared	0.83	0.71	0.83	0.80	0.79	0.71	0.57

Notes: This table reports estimates of equation (2). Dependent variables are defined as follows: the log income of a post office (column 1), the log value of stamps sold (column 2), the annual profit of a post office in contemporary currency units (column 3), and the the log number of distributed mail, parcels, domestic newspapers, and foreign newspapers (columns 4-7). All specifications include controls for log 1870 population, log 1870 market potential, binary indicators that equal one if a city is located on the coast or has direct access to one of the four big lakes, and the number of postal roads that emanate from each city. See the main text and Appendix A for a description of the data. Statistical significance based on standard errors clustered at the county-level is denoted by: *** p<0.01, ** p<0.05, * p<0.10.

Table 6: Access to the Railroad Network and Local Post Offices, 1870.

	Baseline	Sunk Investments (19th Century)			Sectoral Differences (1930s)			Modern Infrastructure (1960s)	
Additional Variable	None	Old Housing	Schools	Telephones	HHI	Manufacturing	Transport	Railroad Rays	Highway Rays
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Early Access (=1)	0.577***	0.528**	0.494**	0.539**	0.504**	0.379**	0.545**	0.219	0.399**
	(0.150)	(0.213)	(0.180)	(0.159)	(0.165)	(0.159)	(0.160)	(0.167)	(0.122)
Additional Variable	-	0.050	0.378*	3.175	2.137	0.025***	-0.016	0.237***	0.147***
	-	(0.096)	(0.196)	(2.478)	(1.183)	(0.005)	(0.013)	(0.042)	(0.033)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	81	81	81	81	81	81	81	81	81
R-squared	0.13	0.13	0.18	0.14	0.16	0.24	0.15	0.29	0.20

Notes: This table presents estimates of equation (4), where the dependent variable is the long differenced log population (2010-1855). “Early access” is a binary indicator that equals one for all cities with access to the railroad network by 1870. See the main text for description of the additional variables. Statistical significance based on standard errors clustered at the county-level is denoted by: *** p<0.01, ** p<0.05, * p<0.10.

Table 7: Explaining Long-Term Persistence, 1855-2010.

	No Large Cities	No Small Cities	No Terminal Cities	Public/Private	City Trends	Regional Trends
	(1)	(2)	(3)	(4)	(5)	(6)
Access to Network (=1)	0.224***	0.186***	0.289***		0.156***	0.278***
	(0.071)	(0.052)	(0.061)		(0.047)	(0.047)
Access via Public Line (=1)				0.319***		
				(0.060)		
Access via Private Line (=1)				0.149***		
				(0.050)		
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	183	183	213	243	243	243
No. of Cities	61	61	71	81	81	81
R-squared	0.68	0.77	0.72	0.75	0.97	0.82

Notes: This table presents robustness checks of the main results provided in Table 2, based on estimations of equation (1). The dependent variable is the log city population. Columns 1 and 2 excludes all large and small cities (above and below the 75th and 25th percentile in 1855) respectively. Column 3 excludes all terminal cities. Column 4 separates the effect for public and private lines. Column 5 includes a full set of city-specific linear trends. Column 6 includes a set of region-specific linear trends. Statistical significance based on standard errors clustered at the city-level is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 8: Robustness Checks, 1840-1870.