A Global Perspective on Railway Efficiency and its Policy and Geographic Determinants, 1880-1912

Dan Bogart[1]

Department of Economics UC Irvine

3151 Social Science Plaza A

Irvine CA 92697-5100

dbogart@uci.edu

Abstract

Railway inefficiency had important consequences for economies in the early twentieth century. This paper estimates the inefficiency of railway sectors using stochastic frontier models and a newly constructed cross-country data set on railway outputs and costs. The results show that Belgium, France, Germany, and India had the most efficient railway sectors by the 1900s while Canada and the Netherlands had the least efficient. The paper then examines whether geography, nationalizations, state vs. private construction, and other factors can explain inefficiency across and within countries. The results show most strongly that inefficiency increased with greater nationalizations. There is mixed evidence that inefficiency was higher in countries with higher elevation and inefficiency decreased with greater state railway construction. More generally the results point to the importance of geography, ownership changes, and the differing incentives of private and state railway officials in determining efficiency.

Keywords: Railways; State Ownership; Cost Inefficiency; Stochastic Frontier Estimation;

Globalization

JEL Classifications: N40, N70, L92, P52

---

[1] I would like to thank seminar participants at Kyoto University, UCLA, IMT Lucca, Warwick University, and LSE for comments on earlier versions of this paper. I would also like to thank Jan-Luiten Van Zanden, David Good, Michael Mesch, Max-Stephan Schulze, Roman Studer, Petra Gerlach, Steve Nafziger, Giovanni Federico, and Ola Grytten for providing invaluable help in locating data. Finally I would like to thank Quan Trinh, Sarah Chiu, and Cecilia Lanata Briones for valuable research assistance.

1. Introduction

Railways were one of the largest sectors at the turn of the twentieth century. In 1910 railway revenues in the US equaled around 3.5% of GDP, they equaled around 6% in Britain, and around 6.5% in Germany.[2] Few sectors outside of agriculture came close to achieving such a large share of total income. Railway tracks, locomotives, and stations also constituted a relatively large share of the total capital stock. In the 1880s railways accounted for 15% of total wealth in the US, 9% of total wealth in Britain, and 8% in Germany (Mulhall 1982, p. 589).

The large size of the railway sector implies that any inefficiency in its operation posed a significant economic problem. In general, inefficiency arises when firms or organizations do not minimize costs with respect to a given output level, input price vector, and available technology. In the railway context technical inefficiency would be associated with the failure to adopt more energy-efficient locomotives, poor network design, or the lack coordination across railway lines. Allocative inefficiency would be associated with the misallocation of inputs, such as the failure to reduce the use of labor when wages increased relative to fuel or capital prices. Either type of inefficiency could have a substantial impact because it meant that large amounts of labor, capital and natural resources would be drawn away from other sectors to railways.

This paper addresses two questions. How did railway inefficiency differ across countries from the 1880s to 1912? What explains the differences in railway inefficiency across countries and over time? In terms of the latter question the analysis focuses on the role of geography and ownership policies. Elevation and its variation is one potential determinant of cross-country differences in efficiency. Locomotives and rails were originally designed in countries with

---

[2] The GDP figures are taken from Mitchell (1992). The sources for revenue will be discussed in the data section.

relatively low elevation like Britain. As these technologies improved in the 1890s and 1900s not all could be easily applied to high elevations and therefore some opportunities to lower costs may have been lost. The amount of coastal terrain or the number of miles of navigable rivers could also explain cross-country differences in inefficiency. A large coastal area or more navigable rivers increased the potential for competition between water and rail transport, which helped to drive down prices and encourage the adoption of new techniques.

The degree of state versus private ownership also stands out as a potential determining factor of efficiency. Private ownership and operation of railways predominated in most countries in 1870. Afterwards there was an evolution towards greater state ownership and operation through nationalizations of privately-owned railways and construction of state owned railways. Private ownership is generally believed to be more efficient than state ownership because it encourages competition and provides stronger incentives for investment and innovation. However, it is not obvious that private ownership contributed to efficiency in the context of railways. In this period private companies were often guaranteed interest or dividends and thus had less incentive to cut costs. Moreover it may have been difficult for private companies to achieve cost savings through the coordination of services. The process of nationalization may have also influenced efficiency. The transfer price was often linked to the profits of railway companies over the previous 5 or 7 years. As a result, companies had the incentive to cut maintenance expenditures in order to temporarily boost profits, but once ownership was transferred the added costs of forgone maintenance set in. More generally, private companies may have also delayed investment or misallocated inputs in anticipation they would be taken over by the state in the future.

Inefficiency is estimated in this paper using stochastic frontier models and new cross-country data on railway performance. The cross-country data are drawn from a number of sources like *the Statistical Abstract of Foreign Countries*, published by the British Board of Trade. It includes total expenses, railway miles, passenger-miles, ton-miles, train-miles, construction costs, fuel prices, and wages in 18 of the largest economies. The stochastic frontier methodology involves the estimation of a cost function with the addition of a random term measuring inefficiency.[3] Following other studies, the specification of the cost function includes variables for scale, density, input prices, and country fixed effects.

The results show that the average level of inefficiency across countries ranged between 0.05 and 0.10, implying the average country could have reduced its costs by 5 to 10% if it had eliminated all inefficiency. More significantly the estimates suggest that the trends in inefficiency differed substantially across countries. Belgium, France, Germany, and India had the most efficient railway sectors by the 1900s, while Canada and the Netherlands had the least efficient. In the 1890s the ranking of inefficiencies were quire different as Belgium, France, Germany, and India were near the bottom and Canada was near the top.

The determinants of inefficiency are analyzed using regressions with variables for the degree of nationalizations, state railway construction, the distribution of elevation, percentage of land area within 100 km of the coast, the number of miles of navigable rivers and canals per square mile of land area, and a host of other variables. The results show that inefficiency was higher in countries with greater elevation or with a positive skew in the distribution of their elevation. The results also show that the inefficiency increased more over time in countries with higher elevations. These results provide some suggestive evidence that greater elevation

---

[3] For an introduction to stochastic frontier models see Kumbhaker and Lovell (2000).

increased inefficiency partly because it limited the application of best-practice technologies. Interestingly, the estimates suggest that inefficiency was not greatly affected by the proportion of land area near the coast or the number of miles of navigable river.

Another main finding shows that inefficiency increased with nationalizations and that inefficiency decreased with greater state railway construction. By themselves these results do not imply causation because nationalizations and ownership choices are endogenous; however they are consistent with the interpretation is that private companies stopped maintaining their track or that managers exerted less effort once they anticipated nationalizations. They are also consistent with the view that interest or dividend guarantees dulled the incentives for private railways to cut costs and that the incentive problems facing state railways were less severe than one might think.

More generally the findings suggest two conflicting effects from greater state ownership. Increased nationalizations lowered efficiency, but greater construction of state-owned railways increased efficiency. The two opposing effects seem to have offset each other in countries like Germany and Belgium, which had significant nationalizations and greater state construction, but not in countries like Switzerland, which had nationalizations but little state construction.

This paper contributes to the historical literature on railway efficiency and productivity. It builds on the studies by Crafts, Mills, and Mulatu (2007), Crafts, Leunig, and Mulatu (2008), Leunig (2006), Arnold and McCartney (2005), and Herranz-Loncan (2006), which estimate freight charges, fares, travel speeds, rates of return, productivity, and cost inefficiency for British and Spanish railways before 1912. It also builds on the work of Foreman-Peck (1987), Millward

(2004), and Bogart (2009) who analyze construction costs and network development in a cross-country setting.[4]

The paper also offers a new perspective on the role of geography and institutions in the global economy from 1870 to 1913. These two factors have been highlighted as 'fundamental' determinants of economic development. The results here suggest that differences in elevation and state ownership policies affected railway efficiency and hence influenced development before 1913. The effects of state ownership are particularly interesting because they reflect policy choices made by political actors, in some cases to advance their own interests and in other cases in response to the wishes of their constituents.

The paper is organized as follows. Section 2 introduces the data. Section 3 discusses the methodology for estimating inefficiency. Section 4 reports the inefficiency estimates by country and decade. Section 5 analyses the determinants of inefficiency. Section 6 concludes.

2. Data

2.1 Railway Performance

Contemporaries collected a substantial amount of data on the financial performance of railways. Almost every aspect was measured including construction costs, operating expenses, revenues, and outputs. The British Board of Trade published summary data on railways in Britain, foreign countries, and British colonies annually starting in the 1870s. The main Board of

---

[4] The paper also relates to the contemporary literature on efficiency estimation. A number of studies have analyzed efficiency at the network or country-level and examined its relationship with ownership or regulatory policies. Some examples include Parisio (1999), Christopoulos, Loizides, and Tsionas (2001), and Cantos and Maudos (2001). Relatively few studies have incorporated country or railway line fixed effects in their models. One exception is the work by Farsi, Fillipini, and Greene (2005) who use stochastic frontier models to estimate cost inefficiency for Swiss railways in the 1980s and 1990s. This paper follows a similar approach using what is known in the literature as the 'true' fixed effects model.

Trade publications are *the Statistical Abstract for the United Kingdom*, *the Statistical Abstract for the Principal and Other Foreign Countries,* and *the Statistical Abstract for the Several Colonial and other Possessions of the United Kingdom*. These volumes marked one of the first attempts to compile and present comparable cross-national data. The information is clearly presented and is easily entered into spreadsheets for statistical analysis.[5]

The main variables drawn from the Statistical Abstracts are total railway miles, train miles, passengers transported, ton shipped, train miles, total expenses, and construction costs. Railway miles (or total route miles) measure the total length of the network. Passengers transported and tons shipped represent the volume of output, but they do not incorporate the distance of the trip or the haul. In some cases the Statistical Abstracts reported the average haul and the average trip, which along with the data on passengers and tons shipped can be used to calculate passenger miles and ton miles.[6] In other cases, International Historical Statistics and secondary sources provide data on ton miles and passenger miles. The appendix provides full details on the sources used in each country.

Train miles measure the number of miles travelled by locomotives in a given year. Train miles are often divided by total route miles to measure the density of traffic. In transport economics, density is often distinguished from total output like ton miles. Consider, for example, a train that moves 100 miles once a day. It can have the same ton-mileage as a train that moves 10 miles 10 times a day, but the cost level may be very different.

---

[5] The statistical Abstracts are now available through web-based versions of the Parliamentary papers. See www. parlipapers.chadwyck.com for details.

[6] Passenger miles are calculated by multiplying the number of passengers transported by the average trip in miles; ton miles are calculated by multiplying tons shipped by the average haul in miles.

Total expenses measure the operational costs of railways. The data in the Statistical Abstracts do not distinguish between different types of expenses, so it is not possible to calculate cost shares. Other data sources show that expenses include the wage bill for train staff and station staff, spending on fuel, spending on maintenance to the track, plant, and equipment, and purchases of new capital goods like locomotives.[7] The Statistical Abstracts do not report total expenses for Britain, Australia, India, and Canada; instead they report working expenses, which are total expenses less spending on new capital goods. Fortunately the ratio of working to total expenses is fairly constant in most countries and averages 0.94. In these cases total expenses were estimated using data from countries that report both working and total expenses.[8] The appendix provides more details.

Construction spending is usually provided separately and is not included in total expenses. Construction costs were expenses associated with building the track and purchasing land for the permanent way. Construction costs therefore increased as route mileage grew, but they did not necessarily increase at a constant rate because the price of land, materials, and labor evolved over time.

The financial data on railways needs to be converted into real terms and into a common currency in order to be comparable across countries and over time. The approach used here is to first deflate expenses and construction costs in the domestic currency using a domestic consumer price index with base year 1905 and then convert the figures into 1905 British pounds or

---

[7] See for example, *Estadistica de Los Ferrocarriles en Explotacion* for a detailed report on the total expenses of railways in Argentina.

[8] Total expenses were estimated using working expenses and the predicted ratio of working to total expenses. The predicted ratio in year t was estimated to equal 0.9444897 - 0.0763076*(Mileage Growth in year t).

shillings using the official exchange rate for 1905.[9] Within a particular country the series will represent real changes in the expenses of railways. Across countries the series will reflect differences in expenses valued at the exchange rate. There are other ways to deflate, say using a GDP deflator, or other exchange rates, say the market exchange rate of the Pound or the US Dollar. The differences between the deflated and converted series are likely to be small. Moreover, the efficiency analysis includes country fixed effects so the choice of the exchange rate will not matter for the estimation.

2.2 Input Prices

Different sources are needed to identify the prices of inputs—fuel, labor, and capital. Coal prices are used to measure the price of fuel because it was the primary source for most railways. Average coal prices are available for many countries in the *Returns Relating to the Production and Consumption of Coal* published by the British Board of trade. The *Returns* specify the 'pit-head' price of coal in shillings per metric ton for Britain, Russia, Sweden, Germany, Belgium, France, Spain, Austria, Hungary, Japan, the US, India, Canada, and Australia. For most years and countries the exchange rate was fixed so the series are effectively expressed as a domestic coal price series converted to the official exchange rate in 1905. The coal price series are deflated using the domestic consumer price index for each country to express coal prices in real terms. The *Returns* did not cover all countries for which there is railway data. Information on Dutch, Norwegian, Italian, and Argentine coal prices were taken from other sources which are described in the appendix. In each case the price was deflated by

---

[9] Consumer Prices indices are generally drawn from Global Financial Database. The exchange rates are available in the Statistical Abstracts. See the appendix for details on each country.

the consumer price index in a country and then converted to British pounds at the 1905 official exchange rate.

The earnings of railway employees in the US, UK, Germany, and several other European countries are available in a report by the Bureau of Railway Economics (1912). The report shows that the average weekly earnings in US dollars for railway employees were 13.02 for the US, 6.15 for the UK, 6.25 for Germany, and 5.55 for Italy from 1905 to 1909. The report does not have enough data to form a wage series, but fortunately the earnings of railways employees are similar to general wages for skilled and unskilled workers. Williamson's (1995) data shows that from 1905 to 1909 the average weekly earnings for skilled and unskilled builders in US dollars were 23.8 and 14.3 for the US, 9.1 and 6.2 for the UK, 6.9 and 5.2 for Germany, 3.4 and 2.1 for Italy (p. 184). These weekly earnings are broadly similar to the figures for railway employees and suggest that railways hired some combination of skilled and unskilled labor. Building on these data the price of labor is measured using the weekly wages of skilled and unskilled labor in each country. The appendix provides details on the nominal wage series in each country. The wages are all deflated using a domestic consumer price index with base year 1905 and then converted to British shillings using the 1905 official exchange rate.

The prices for new capital goods, like locomotives, and materials for repairs like timber, iron, steel, bricks, etc. are more problematic because these series are not available for most countries.[10] Fortunately, materials for repairs to the track and stations were the same or similar as those used for construction of the track and station. Therefore changes in construction costs per mile provide a measure of changes in the average price of capital goods used for repairs. It

---

[10] Collins and Williamson (2001) provide data on capital prices for a handful of countries but they are too few to cover the countries here and may not be applicable to the railroad sector.

should be noted that construction costs per mile are a noisy measure of capital prices because they also reflect the prices of other inputs used to build railways like unskilled labor and land.

Construction costs are published in the Statistical Abstracts for most countries. Secondary sources were used for Spain, Italy, and Belgium where construction costs are not reported (see the appendix for details). The average construction cost per mile was deflated and converted using the same procedures for wages and coal prices.

2.3 Nationalizations, State Construction, and other Variables

The degree of nationalizations and state ownership in the railway sector can be estimated using data in the Statistical Abstracts and related sources.[11] The basic variables are the fraction of miles nationalized in a country and the fraction of miles constructed by the state as opposed to private companies. Miles are defined to be constructed by the state if it initially owned and operated the line or it was the initial residual clamant. Table 1 summarizes the fraction of miles nationalized by 1880 and 1910 and the fraction of miles constructed by the state by 1880 and 1910 for the sample of countries analyzed here. In 1880 nationalizations were rare outside of a few countries like Belgium, Germany, India, and Italy. The state constructed a relatively large proportion of the miles in some countries, including Norway, Holland, Belgium, Argentina, Germany, and Australia. In others like Britain, Spain, and the US construction of state-owned railways was absent or negligible.

By 1910 nationalizations had become more common. State takeovers affected a relatively large proportion of the miles in Russia, Belgium, Italy, and Germany. The most extreme cases were Japan, Switzerland, and Austria where over 50% of all miles had been

---

[11] See Bogart (2009) for details on the measurement of nationalizations and state ownership.

nationalized by 1910. The degree of state construction also differed across countries by 1910. In Russia, Germany, Italy, and India the state constructed a greater proportion of the mileage, but in Norway, Japan, Argentina, Canada, and Australia the state constructed a smaller proportion.

The geography of a country can be measured in many ways. Here the emphasis is on the distribution of elevation and the percent of land area near a seacoast. The Center for International Earth Science Information Network (CIESIN) provides data on the percentage of land area in 12 different elevation classes (0-5 meters, 5-10 meters, etc). [12] From this information the average elevation level, the standard deviation of elevation, and the skew of elevation can be approximated.[13] CIESIN also provides data on the percentage of land area within 100 km of the coast. The country boundaries in the CIESIN data are modern and so there is some measurement error for countries like Russia and India, but this should not dramatically affect the results.

The number of miles of navigable rivers and canals is taken from Mulhall (1982, p. 102). The data are drawn from measurements of waterway networks in the 1880s. A more detailed analysis of waterway networks by Vernon-Harcourt (1896) suggests that the length of waterway networks was fairly stable over time even though some countries continued to build canals in the 1880s and 1890s.

Table 2 shows the distribution of elevation, the percent of land within 100km of the coast, and the number of miles of navigable rivers and canals per square mile across countries. Switzerland had the highest average elevation at 883 km. The Netherlands had the lowest

---

[12] See Center for International Earth Science Information Network, Columbia University, 2007. National Aggregates of Geospatial Data: Population, Landscape and Climate Estimates, v.2 (PLACE II), Palisades, NY: CIESIN, Columbia University. Available at: http://sedac.ciesin.columbia.edu/place/.

[13] The land area of all countries was normalized to be 100 square meters. The number of square meters within an elevation class was assumed to be equal to the percentage of land in each elevation class. Land was assumed to be equal to the lower bound in its elevation class to simplify the calculations. Then average, standard deviation, and skew were calculated over the distribution.

average elevation at 6.65 km. Argentina had the highest standard deviation of elevation and the Netherlands had the lowest. The skew measures the amount of land area in the upper tail of the elevation distribution. The Netherlands had the greatest skew while Austria had the lowest skew. Island countries like Britain and Japan had the greatest land area within 100km of the coast. A land-locked country like Switzerland had none of its land area within 100km of the coast. Finally the Netherlands had the greatest number of navigable rivers and canals per square mile. Canada had the fewest.

Other variables are also incorporated as controls. The average age of the network is calculated using the number of miles added in each year as an indicator for the 'birth' of track miles. Miles are assumed to never 'die' in the calculation. GDP per capita and GDP per capita growth are available from Maddison (2003) and are expressed in 1990 PPP adjusted US dollars.[14] The use of PPP dollars is not crucial in this context because the GDP figures are correlated with the inefficiency estimates, which are not based on a value measure.

3. Methodology for Estimating Inefficiency

Panel data allows for a number of approaches to estimating efficiency. One of the most common is the 'true fixed effects' model. In the railway context the baseline model is usually specified as a Cobb-Douglas cost function taking the following form:

$$\ln c_{it} = \sum_{k=1}^{K} \beta^k \ln q_{ii}^k + \sum_{j=1}^{J} \gamma^j \ln p_{ii}^j + \mu_1 t + \mu_2 t^2 + \alpha_i + u_{it} + v_{it} \quad (1)$$

---

[14] GDP data for Russia are taken from Gregory (1982).

where $\ln c_{it}$ is the natural log of real expenses for country i in year t, $\ln q_{it}^{k}$ is the log of ton-miles, passenger-miles, and rail miles for country i in year t, $\ln p_{it}^{j}$ is the log of the real price of unskilled labor, skilled labor, fuel, and construction costs per mile for country i in year t, t and t^2 are the year and year squared respectively, α_i is a fixed effect for country i, u_{it} is a half-normal random variable with mean 0 and variance σ_u^2, and v_{it} is a normal random variable with mean 0 and variance σ_v^2. The term u_{it} measures cost inefficiency so that a higher value corresponds with higher costs. Cost inefficiency is estimated for each country and year using the composite error term $\varepsilon_{it} = u_{it} + v_{it}$ and the parameter estimates for σ_u^2 and σ_v^2. The conditional mean for cost inefficiency is calculated using the following formula from Jondrow et. al. (1982):

$$E[u_{it} \mid \varepsilon_{it}] = \frac{\sigma\lambda}{1+\lambda^2}\left[\frac{\phi(z)}{1-\Phi(z)} - z\right] \qquad (2)$$

where ϕ is the p.d.f of the standard normal, Φ is the c.d.f. of the standard normal, $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$, $\lambda = \frac{\sigma_u}{\sigma_v}$, $z = \frac{\varepsilon_{it}\lambda}{\sigma}$, and $\varepsilon_{it} = u_{it} + v_{it}$.

The true fixed effects model is preferred in this setting because it has several advantages over alternatives like the fixed effects model and the pooled normal half-normal model (Greene 2005). First, unlike the fixed effects model, it allows cost inefficiency to be time-varying. This is an important property in this context because there is no reason to believe that the efficiency of railways remained constant over a thirty-year period. Second, unlike the pooled normal half-normal model, the true fixed effects model includes a country fixed effect which captures all time-invariant unobservable factors. Time-invariant factors should be separated from inefficiency because they will cause some countries to operate at higher or lower costs because of factors beyond their control. The true-fixed effects model provides a methodology precisely for separating time-invariant unobservable factors from inefficiency.[15]

The Cobb-Douglas specification with a half normal random variable for the inefficiency term u_{it} provides a 'baseline' model. Several other specifications are analyzed to check the robustness of the estimates. One adds the length of the average trip and the average haul. Another adds the log of train miles divided by route miles. Both of these specifications incorporate different controls for density. A third extension imposes the assumption that the cost function is homogenous of degree one in prices. In the Cobb-Douglas specification this implies that the coefficients on the price variables sum to one. Substituting the expression $\gamma^{unskilled} = 1-(\gamma^{skilled} + \gamma^{fuel} + \gamma^{capital})$ into the cost function and simplifying gives the following estimating equation.

$$\ln\left(\frac{c_{it}}{p_{ii}^{unskilled}}\right) = \sum_{k=1}^{K} \beta^k \ln q_{ii}^k + \sum_{j=1, j\neq unskilled}^{J} \gamma^j \ln\left(\frac{p_{ii}^j}{p_{ii}^{unskilled}}\right) + \alpha_i + u_{it} + v_{it} \qquad (3)$$

The left hand side variable is the log of expenses divided by wages for unskilled labor and on the right-hand side everything else is the same except now it includes the log of real coal prices, real skilled wages, and real construction costs per mile all divided by unskilled wages. A fourth extension replaces the Cobb-Douglas specification with a translog cost function. The Cobb-Doulgas assumes a constant elasticity of substitution across inputs. The translog relaxes this assumption by allowing the elasticity to vary with different combinations of outputs and inputs.

---

[15] Notice that geography is a time-invariant variable: elevation levels do not change over time. Therefore, the country fixed effects will partly subsume the effects of geography on railway costs. However, geography arguably affects the adoption of technologies and so its influence may be time-varying. The issue of geography will be revisited later in the analysis of the determinants of inefficienc

A final set of extensions considers the gamma, exponential, and truncated normal distributions as an alternative distribution for the inefficiency term u_{it}.

4. Inefficiency Estimates

The estimates reveal significant differences in inefficiency across countries and over time. Table 3 reports the average inefficiency estimate by country in the 1880s, 1890s, and 1900s. It also lists the rankings from the most efficient to the least efficient in the 1900s. The results show that Belgium, India, France, and Germany had the most efficient railway sectors in the 1900s. All four experienced significant decreases in inefficiency from the 1880s to the 1900s. In Belgium for example, average inefficiency went from 0.105 in the 1880s to 0.094 in the 1890s to 0.051 in the 1900s. These figures imply that railway expenses were reduced by 5.4% in Belgium because of the lowering of inefficiencies between the 1880s and 1900s. The estimates suggest similar cost savings were realized in France, India, and Germany.

There were a large number of countries with ‘medium’ efficiency levels by the 1900s. The US and Spain had average inefficiency levels of 0.073 and 0.075; Australia and Norway had inefficiency levels of 0.08 and 0.088; Japan, Argentina, Britain and Russia had average inefficiencies of 0.093, 0.093, 0.095, and 0.096. The trends were varied in this group. Britain’s inefficiency decreased in the 1890s and then increased in the 1900s. In the US inefficiency increased in the 1890s and then decreased in the 1900s. In Argentina inefficiency increased throughout.

Canada and the Netherlands had the least efficient railway sectors in the 1900s. The estimates suggest that both countries could have reduced their costs by at least 13% if they eliminated all inefficiency. Canada arrived at this point because its inefficiency rose over time.

The Netherlands experienced an increase in inefficiency in the 1890s. Ironically it had the most efficient railway sector in the 1880s. Sweden, Italy, Austria, and Switzerland, and were slightly more efficient than Canada and the Netherlands. Inefficiency increased in all four countries from the 1890s to the 1900s.

Overall the estimates suggest that countries with the most efficient railway sectors in the 1880s or 1890s were not necessary the most efficient by the 1900s. There is a negative correlation between cost inefficiency in the 1880s and 1890s and a negative correlation between cost inefficiency in the 1890s and 1900s. The average (un-weighted) cost inefficiency decreased over time from 0.113 to 0.067. This was not the case though in some of the world's largest railway countries such as the US and Britain.

The estimated inefficiency differences across countries are very robust to different specifications, but the levels of inefficiency do differ. Table 4 reports the average inefficiency for alternative specifications and compares them with the estimates using the baseline model reported in column (1). Specifications (2) and (3) add variables for density like the average haul and the average trip or the log of train miles divided by rail miles. Specifications (4) to (6) assume the cost function is homogenous. Specifications (7) to (9) replace the half-normal normal distribution with the exponential, gamma, and truncated normal distributions. Specification (10) uses a translog cost function instead of the Cobb-Douglas.

Average inefficiency varies across the specifications. In the baseline model, the mean inefficiency is around 0.09, but it is 0.046 when the cost function is assumed to be homogenous of degree one in prices, and it is 0.044 in the translog specification. Although there are differences in the means, the inefficiency estimates are highly correlated across the specifications. The bottom panel of table 3 shows the correlation coefficient with the baseline model. Most correlation coefficients are above 0.96. The only exception is the tranlog specification where the correlation coefficient is 0.78. Overall the estimates yield consistent results on which countries had relatively high efficiency and which had relatively low efficiency, but they are not consistent on how severe inefficiency was for the average country.

The estimates of the cost function are also of interest. Table 5 reports the results for the first three specifications. There is evidence of economies of scale and economies of density. In the baseline model, the coefficients for ton miles and passenger miles sum to around 0.75, implying economies of scale. The coefficient for average trip length is negative implying that density is associated with lower costs. The coefficient for the log of train miles divided by rail miles is essentially zero in (3), but it is negative and significant in an unreported specification that assumes the cost function is homogenous. The coefficients on the prices for fuel, unskilled labor, skilled labor, and construction costs per mile are all similar and highly significant. Interestingly the coefficient on unskilled labor is negative. It is possible that the effects of unskilled wages are partly captured by construction costs per mile. Recall that construction costs per mile are essentially a weighted average of the price of capital goods and other inputs like unskilled labor, but the weighting between them is unknown. Lastly in all specifications, the time trend coefficients (i.e. the year and the year square) are not significant. This suggests that the effects of world-wide technological change are either small or they are being offset by other omitted factors which raised railway costs for all countries over time.

5. Determinants of Inefficiency

What caused inefficiency to be high in some countries and low in others? What caused inefficiency to increase within some countries and decrease within others? This section addresses these questions by focusing on the role of ownership polices and geographic factors. One hypothesis is that the degree of private or state construction or the extent of nationalizations influenced efficiency. Geographic features, like higher elevation levels or elevation extremes, are hypothesized to influence the degree to which new technologies could be applied. The amount of land area near the coast and the number of miles of navigable rivers and canals are also potentially important factors because they allowed for greater competition between water and rail transport. Other possible determinants of inefficiency are considered as well. Railways with older networks may have been less efficient because of difficulties adopting the latest technologies; the rate of economic growth or the level of per capita income could also influence inefficiency through a variety of channels.

Two empirical approaches are used to investigate the determinants of inefficiency. The first regresses inefficiency levels in country i and year t on variables measuring ownership, geography, economic development, and the average age of the track mileage. The second approach regresses inefficiency levels in country i and year t on the same with country fixed effects added. This latter regression captures the effect of changes in ownership policies, development, and average track age within countries and controls for time-invariant unobserved heterogeneity.[16]

Table 6 shows the results for the specifications without country fixed effects. Columns (1)-(4) use the inefficiency estimates from the baseline model. The results in column (1) show a

---

[16] A third approach is to use tobit regressions with and without country random effects. The conclusions are very similar and so these results are not reported.

strong, positive relationship between nationalizations and inefficiency. The variable for the fraction of miles nationalized is positive and statistically significant at the 1% level. The estimate implies that a nationalization of 50% of the network, as occurred in Japan or Switzerland, could increase inefficiency by 0.025 which is about half of a standard deviation for inefficiency. There is no statistically significant relationship between greater state railway construction and inefficiency. Since the alternative to greater state railway construction was private railway construction, it appears that countries with more privately-built railways were not any more efficient than countries with more state-built railways.

The results in column (2) address the role of geographic factors. A higher average elevation and a higher positive skew in the distribution of elevation were both associated with greater inefficiency. An increase in the average elevation by 200km would increase inefficiency by 0.016 which is a little more than a quarter of a standard deviation in efficiency. An increase in the skew of the elevation by 1 would increase inefficiency by 0.022 which is a little less than one-half of a standard deviation in inefficiency. The standard deviation of elevation had little influence after controlling for the average elevation and the skewness of elevation.

The percentage of land area within 100 km of the coast and the number of miles of navigable rivers and canals per square mile are the other geographic variables of interest. The estimates show that neither had a large influence on inefficiency. This is somewhat surprising because countries with greater coastline or greater rivers should presumably have greater competition between water and rail transport. The data suggest, however, that the potential for water-rail competition was not enough to increase efficiency.

The results in column (3) address the effects of elevation over time by adding an interaction between the time trend and the average elevation. The coefficient on elevation is now negative and its interaction with the year is positive. This implies that in countries with higher elevation the railway sector tended to become less efficient over time. A similar interaction term was included for the skew of elevation but there was no statistically significant effect. Higher average elevation is the only geographic variables where the effect varies significantly over time.

There are several explanations for the connection between elevation and inefficiency. Operating railways in mountainous terrain was more expensive because greater fuel was required and maintenance was more difficult. Countries with greater elevation levels or elevation extremes also had difficulties adapting technologies. Many railway technologies, like locomotives, were initially designed in Britain and Germany which have relatively low elevation. Technologies were then transferred to other countries. Not all were adopted though. Prefabricated bridges provide one example. French engineers developed a standard construction kit for building bridges in France. French engineers tried to use the same prefabricated bridges in the high Andes of Peru but were unsuccessful (Faith 1990, p. 131). In this case, high extreme elevation limited or at least delayed the adoption of technologies.

Column (4) in table 6 includes some additional explanatory variables like GDP per capita growth, GDP per capita, and the average age of the network. The results show that GDP per capita growth and GDP per capita are not correlated with inefficiency in the cross-country analysis. A higher average age for track mileage is negatively and significantly related to inefficiency. This finding is somewhat surprising because one might expect that a greater age would be positively related to inefficiency because of difficulties adapting the latest technologies. Instead it appears that 'experience' paid dividends and networks of a younger age tended to be less efficient. The connection between network age and efficiency should not be over-stressed though. The regressions with country fixed effects analyzed below show that increases in the average age of the network are not strongly associated with lower inefficiency.

Most of the results are very similar if inefficiency estimates from the alternative specifications are used. The results do differ, however, when the inefficiency estimates from the translog specification are used. Column (5) reports the results using inefficiency from the translog model. The only variable that has any statistically significant effect on inefficiency is the fraction of miles nationalized. The signs of most other variables are the same, except the coefficients are smaller and the standard errors are larger. This suggests that the relationship between elevation and inefficiency is not as statistically robust as the earlier results would indicate. This is perhaps not surprising because there is less variation in geography with only 18 countries in the sample. There is also less variation in the inefficiency estimates under the translog model. Still these results cast some doubt on the view that elevation was a key determinant of inefficiency.

The preceding analysis focused on the cross-sectional variation, but it is also revealing to examine how inefficiency changed within countries as nationalizations or state railway construction increased. Table 7 addresses this issue by reporting results from specifications that add country fixed effects. The variables for geography are necessarily dropped because they do not vary within countries. The results in column (1) show that a positive and significant relationship between inefficiency and the fraction of miles nationalized. This implies that inefficiency increased as nationalizations increased within a country. The results also show a negative and significant relationship between inefficiency and the fraction of miles built by the state. This implies inefficiency decreased as state railway construction increased within a country.

Column (2) adds variables for GDP per capita, GDP per capita growth, and the average age of the network. The results still show the same relationship between nationalizations, state railway construction, and inefficiency. The average age is not significantly related to inefficiency. Comparing the results in tables 6 and 7 it appears that countries with greater network age tended to be more efficient, but efficiency did not increase as a countries' network aged. Therefore, greater age does not appear to be a robust determinant of efficiency. The results also show that increases in GDP per capita were associated with higher inefficiency. This finding discounts the possibility that greater efficiency in the railway sector simply reflects greater labor productivity in the economy. It is not immediately obvious though why greater income per capita contributed to higher inefficiency. Here the omitted variables problem is perhaps most concerning. Higher GDP per capita might be associated with greater unionization or higher legal costs both of which would have an influence on efficiency. The relationship between income per capita and inefficiency needs further study before firm conclusions can be drawn.

Column (3) reports results for the same variables using the inefficiency estimates from the translog model. The coefficient for nationalizations is positive and significant like before. The coefficient on state railway construction is still negative, but in this case it is not significant. Thus is not clear that state railway construction was associated with lower inefficiency. The results are still noteworthy though because they show that private railway construction was not associated with lower inefficiency.

Together the results provide some suggestive evidence regarding the effects of ownership policies on railway inefficiency before 1913. First, state ownership per se did not contribute to inefficiency. If that were the case the coefficients on the fraction of miles nationalized *and* the fraction of miles built by the state should have been positive, but the latter is generally negative. Instead the results suggest that large changes in ownership were associated with increases in inefficiency. This makes sense if private companies stopped maintaining their track or exerted less effort in monitoring employees once they anticipated nationalizations. It might also be the case that the *transfer* of managerial control from the private sector to the state proved to be very problematic.

The results also suggest that greater state railway construction opposed to private railway construction did not inhibit inefficiency. Recall that most private railway companies received some type of guaranteed return on their dividends or bonds. Guarantees were deemed necessary because private companies were building feeder lines in areas with marginal economic potential. However, they arguably dulled the incentives for private railways to cut costs. Some contemporaries vigorously opposed giving guarantees for private railways and preferred state railway construction. Prince Thewawong of Thailand gave the following argument on why state railway construction should be preferred over a particular proposal for a private railway concession with a 5% guarantee:

> "This concession is not suitable for us. It looks as if we will freely give them profit and in return they will deliver us several types of troubles. It would be better if we were able to construct the line ourselves. The use of a loan for construction is still better [than a concession] if we have no funds."[17]

---

[17] Quoted in Ichiro (2005, pp 95-97).

The estimates suggest that state railway construction may have indeed been more effective. By the 1890s and 1900s state bureaucracies were becoming increasingly sophisticated and they seem to have been able to resolve the incentive problems highlighted by private ownership advocates at the turn of the century and today.

6. Conclusion

Railway inefficiency imposed significant costs on economies in the early twentieth century because the railway sector was relatively large. This paper estimates inefficiency using stochastic frontier models and a newly constructed cross-country data set on railway outputs and costs. The results show there were significant differences in inefficiency across countries and over time. Specifically, Belgium, France, Germany, and India had the most efficient railway sectors by the 1900s, but none of these four countries was in the top ranks of efficiency in the 1880s or the 1890s.

The determinants of inefficiency are also analyzed using variables for geography, nationalizations, state vs. private construction, and other factors. The results show most strongly that inefficiency increased with greater nationalizations. There is mixed evidence that inefficiency was higher in countries with high average elevation and that inefficiency decreased with greater state railway construction. More broadly the results speak to importance of two of the 'fundamental' causes of economic development: geography and institutions. Both impacted inefficiency in one of the world's largest and most important sectors at the turn of the twentieth century.

Data Appendix by Country

Russia. The complete data cover the years from 1891 to 1907. The Statistical Abstracts provide data on railway miles, passengers carried, average journey, tons shipped, average haul, total expenses, and train miles from 1899 to 1910. Expenses are converted into British pounds using the official exchange rate of 0.105 shillings for one Rouble in 1905. The 1899 figures for the average haul and the average journey were assumed to apply for the years 1891 to 1898. Ton miles and passenger miles are calculated by multiplying tons by the average haul and passengers by the average journey. Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1891 to 1907. The prices are reported as average values in shillings. The official exchange rate was essentially constant (0.10 from 1891 to 1896 and 0.105 from 1897 to 1907) so the reported coal price series is already expressed in shillings at the 1905 official exchange rate. Weekly wages for unskilled and skilled workers are based on the real annual earnings of building and railroad staff reported in Allen (2003, p. 38). The series were converted into weekly nominal earnings using Gregory's (1982) retail price index and then converted into British shillings using the 1905 exchange rate. Construction costs per mile were estimated using Mulhall's (1892) figure construction costs in 1888 and Gregory (1982) annual series on railway investment from 1888 to 1907. Gregory's annual investment series is converted to 1905 shillings using the exchange rate. Mulhall's figure for construction is already expressed in British pounds. All financial variables are deflated using the retail price index from Gregory (1982) which was converted to the base year 1905.

Norway: The complete data cover the years from 1880 to 1912. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, total expenses, train miles, and construction costs. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.056 pounds per Kroner in 1905. Ton miles and passenger miles are taken from Mitchell (1992). Coal prices per metric ton were provided by Ola H Grytten. They were converted into British pounds using the 1905 official exchange rate. Weekly wages for unskilled and skilled workers are based on day wages for workers in road construction and railway staff in Grytten (2007 pp. 281-282, 315-316). The wages are converted into British shillings using the 1905 exchange rate. All financial variables are all deflated using the Norwegian consumer price index from Grytten (2004) with base year 1905.

Sweden: The complete data cover the years from 1896 to 1912. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, total expenses, train miles, and construction costs. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.056 British pounds per Kronor. Ton miles and passenger miles are taken from Mitchell (1992). Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1896 to 1912. The prices are reported as average values in shillings. The official exchange rate was constant from 1896 to 1912 so the reported coal price series is already expressed in shillings at the 1905 official exchange rate. Weekly wages for unskilled and skilled workers are taken from the yearly wages of workers in sawmills and engineering listed in Bjorklund and Stenlund (1995, p.253). Wages are converted into shillings using the 1905 exchange rate. All financial variables are deflated using the Swedish consumer price index from Mitchell (1992) with base year 1905.

Netherlands: The complete data cover the years from 1884 to 1902, 1904, 1906, and 1910. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, total expenses, train miles, and construction costs (state railways only). Expenses and construction costs are converted into British pounds using the official exchange rate of 0.083 British pounds per Guilden. Ton miles and passenger miles are taken from Mitchell (1992). Coal prices are provided by Arthur van Riel through the Global Price and Income History Group (http://iisg.nl/hpw/data.php#netherlands). The official exchange rate was constant from 1884 to 1910 so the reported coal price series is already expressed in shillings at the 1905 official exchange rate. The wages for unskilled and skilled workers are based on day wages for laborers and craftsman in Amsterdam (Allen 2001). The wages of builders are converted into an index with base year 1905 and multiplied by the weekly earnings of unskilled workers in 1905 shillings, which is estimated using Williamson's figures for the Netherlands and Britain (1995). The wages for skilled workers are equal to the weekly wages of unskilled in shillings multiplied by the ratio of craftsman to builders wages. All financial variables are deflated using the Dutch consumer price index from Mitchell (1992) with base year 1905.

Belgium: The complete data cover the years from 1883 to 1910 except 1903 and 1908. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, total expenses, and construction costs. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.04 British pounds per Belgian Franc in 1905. The Statistical Abstracts provide the average journey from 1901 to 1912 and the average haul in 1908, 1911, and 1912. The average journey in 1901 was assumed to apply to the period from 1883 to 1900. The average haul in 1908 was assumed to apply for the entire period from 1883 to 1910. Ton miles and passenger miles are then calculated by multiplying tons by the average haul and passengers by the average journey. The Statistical abstracts provide information on train miles for all railways from 1901 to 1910 and for state railways from 1883 to 1902. Total train miles from 1883 to 1900 are assumed to grow at the same annual rate as state railway train miles. Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc.

of Coal from 1884 to 1912. The prices are reported as the average value 'when on the truck for transportation' in shillings. The official exchange rate was constant from 1883 to 1912 so the reported coal price series is already expressed in shillings at the 1905 official exchange rate. Weekly earnings for unskilled and skilled workers is based on the nominal annual earnings for manufacturing workers and civil servants in Scholliers (1995, p. 204). The weekly earnings are converted to shillings using the 1905 exchange rate. All financial variables are deflated using the Belgian consumer price index from Mitchell (1992) with base year 1905.

France: The complete data cover the years from 1883 to 1911. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, total expenses, train miles, and construction costs. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.04 British pounds per French Franc in 1905. Ton-miles and passenger-miles are taken from Mitchell (1982). Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1884 to 1912. The prices are reported as the average value at the pit's mouth in shillings. The official exchange rate was constant from 1883 to 1912 so the reported coal price series is already expressed in shillings at the 1905 official exchange rate. The wages for unskilled and skilled workers are based on the day wages for laborers and craftsman in Paris (Allen 2001). The wages of builders are converted into an index with base year 1905 and multiplied by the weekly earnings of unskilled workers in 1905 shillings from Williamson (1995). The The wages for skilled workers are equal to the weekly wages of unskilled in shillings multiplied by the ratio of craftsman to builders' wages. All financial variables are deflated using the French consumer price index from Mitchell (1992) with base year 1905.

Switzerland: The complete data cover the years from 1884 to 1912. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, total expenses, train miles, and construction costs. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.04 British pounds per Swiss Franc in 1905. Ton miles and passenger miles are then calculated by multiplying tons by the average haul and passengers by the average journey. The Statistical abstracts report average haul and the average journey for 1906 to 1912. The average haul and journey in 1906 was assumed to apply for the whole period from 1884 to 1912. Coal prices are published in Siegenthaler (1996) and were converted to shillings per ton using the 1905 exchange rate. The wages for unskilled and skilled workers are based on the day wages for laborers and craftsman (Studer forthcoming). The wages are converted in 1905 shillings per week using the exchange rate. All financial variables are deflated using the Swiss consumer price index from Mitchell (1992) with base year 1905.

Spain: The complete data cover the years from 1884 to 1909. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, and total expenses. The statistical abstracts provide data on total expenses from 1898 to 1909. For earlier years total expenses are estimated from working expenses using the formula discussed in the text. Expenses are converted into British pounds using the official exchange rate of 0.04 British pounds per Pasetas in 1905. Ton-miles and passenger-miles are taken from Mitchell (1982). Construction costs per mile were estimated using Mulhall’s (1892) figure construction costs in 1880 and Herranz-Locain’s (2005) annual series on railway investment from 1880 to 1909. Herranz’s annual investment series is converted to 1905 shillings using the exchange rate. Mulhall’s figure for construction is already expressed in British pounds. Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1884 to 1912. The prices are reported as the average value at the pit's mouth in shillings. The official exchange rate was constant from 1884 to 1909 so the reported coal price series is already expressed in shillings at the 1905 official exchange rate. The wages for unskilled and skilled workers are based on the day wages for laborers and craftsman in Madrid (Allen 2001). The wages of builders are converted into an index with base year 1905 and multiplied by the weekly earnings of unskilled workers in 1905 shillings from Williamson (1995). The wages for skilled workers are equal to the weekly wages of unskilled in shillings multiplied by the ratio of craftsman to builders' wages. All financial variables are deflated using the Spanish consumer price index from Flandreau and Zumer (2004) with base year 1905.

Italy: The complete data cover the years from 1890-91, 1898-1901, 1903, 1906, 1911. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, train miles, and total expenses. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.04 British pounds per Lire in 1905. Ton-miles and passenger-miles are taken from Mitchell (1982). Italian coal prices per ton are published in Cianci (1933). They are converted to British shillings using the 1905 exchange rate. The wages for unskilled and skilled workers are based on the day wages for construction and engineering (Scholliers and Zamagni 1995, p. 231). The wages are converted in 1905 shillings per week using the exchange rate. All financial variables are deflated using the Italian consumer price index from Mitchell (1992) with base year 1905.

Japan: The complete data cover the years from 1894 to 1911. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, total expenses, train miles, and construction costs. Ton miles and passenger miles are taken from Mitchell (1995). Expenses and construction costs are converted into British pounds using the official exchange rate of 0.102

British pounds per Yen in 1905. Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1894 to 1911. The prices are reported as the average market value in shillings. The exchange rate was essentially constant from 1894 to 1911 (0.105 in 1895 and 1896 and 0.108 in 1897) so the reported coal price series is essentially expressed in shillings at the 1905 official exchange rate. The wages for unskilled and skilled workers are based on day wages for laborers and craftsman (masons) in (Allen 2001), which are provided by David Jacks, Peter Lindert, and Salvadore Puente through the Global Price and Income History Group and are based on the Financial and Economic Annual of Japan. The wages are converted into shillings per week using the exchange rate in 1905. All financial variables are deflated using the Japanese consumer price index from Global financial data with base year 1905.

US: The complete data cover the years from 1883 to 1911. The Statistical Abstracts provide data on railway miles, passengers carried, passenger miles, tons shipped, ton miles, total expenses, train miles, and construction costs. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.208 British pounds per dollar in 1905. Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1883 to 1911. The prices are reported as the average spot value in shillings. The exchange rate was constant from 1883 to 1911 so the reported coal price series is expressed in shillings at the 1905 official exchange rate. The wages for unskilled and skilled workers are based on weekly earnings for low skilled labor and railway staff in (Margo 2006). All financial variables are deflated using the US consumer price index from Global financial data with base year 1905.

Argentina: The complete data cover the years from 1897 to 1912. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, total expenses, train miles, and construction costs. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.2 British pounds per peso in 1905. Ton miles, passenger miles, and average coal prices are taken from *Estadistica de Los Ferrocarriles en Explotacion*. Coal prices are converted in shillings using the 1905 exchange rate. The wages for unskilled and skilled workers are based on weekly earnings for railway construction workers and railway staff in *Estadistica de Los Ferrocarriles en Explotacion*. All financial variables are deflated using the Argentina consumer price index from Williamson (1995) with base year 1905.

Britain: The complete data cover the years from 1883 to 1912. The Statistical Abstracts for the United Kingdom provide data on railway miles, passengers carried, tons shipped, train miles, and construction costs. Total expenses are estimated from working expenses using the formula discussed in the text. Ton miles and passenger miles are taken from Crafts, Mills, and Mulatu (2007). Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1883 to 1912. The prices are reported as the average value at the pit's mouth in shillings. The wages for unskilled and skilled workers are based on the day wages for laborers and craftsman in London (Allen 2001). The wages of laborers are converted into an index with base year 1905 and multiplied by the weekly earnings of unskilled workers in 1905 shillings from Williamson (1995). The wages for skilled workers are equal to the weekly wages of unskilled in shillings multiplied by the ratio of craftsman to builders' wages. All financial variables are deflated using the UK consumer price index from Mitchell (1992) with base year 1905.

Germany: The complete data cover the years from 1883 to 1912. The Statistical Abstracts provide data on railway miles, passengers carried, passenger miles, tons shipped, ton miles, total expenses, train miles, and construction costs. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.08 British pounds per mark in 1905. Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1883 to 1912. The prices are reported as the average value at the pit's mouth in shillings. The official exchange rate was constant from 1883 to 1912 so the reported coal price series is expressed in shillings at the 1905 official exchange rate. The wages for unskilled and skilled workers are based on the day wages for laborers and craftsman in Leipzig (Allen 2001). The wages of laborers are converted into an index with base year 1905 and multiplied by the weekly earnings of unskilled workers in 1905 shillings from Williamson (1995). All financial variables are deflated using the German consumer price index from Mitchell (1992) with base year 1905.

India: The complete data cover the years from 1891 to 1912. The Statistical Abstracts for British India provide data on railway miles, passengers carried, passenger miles, tons shipped, ton miles, train miles, and construction costs. Total expenses are estimated from working expenses using the formula discussed in the text. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.067 British pounds per rupee in 1905. Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1883 to 1912. The prices are reported as the average value at the pit's mouth in shillings using various exchange rates before 1898 and a constant exchange rate afterwards. Before 1898 Coal prices are converted back to rupees and then back to shillings using the official exchange rate in 1905. The wages for unskilled and skilled workers are based on the day wages for laborers and craftsman in five areas Karachi, Beglum, Ahmadgar, Bombay, and Ahmadbad. The wage data comes from the Statistical Abstract for British India. All financial variables are deflated using the Indian consumer price index from Global Financial data with base year 1905.

Canada: The complete data cover the years from 1886 to 1912. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, train miles, and construction costs. Total expenses are estimated from working expenses using the formula discussed in the text. Expenses and construction costs are converted into British pounds using the official exchange rate of 0.205 British pounds per Canadian dollar in 1905. Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1886 to 1912. The prices are reported as the average value in shillings. The official exchange rate was constant from 1886 to 1912 so the reported coal price series is expressed in shillings at the 1905 official exchange rate. Wage rates for unskilled and skilled workers are taken from Historical Statistics of Canada and are converted to British pounds. All financial variables are deflated using the consumer price index from Canadian consumer prices are from Minns and MacKinnon (2006) with base year 1905.

Austria: The complete data cover the years from 1894 to 1910. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, train miles, total expenses, and construction costs. Expenses and construction costs are reported in Gulden before 1900 and are converted into British pounds using the official exchange rate of 0.083 British pounds per Gulden. After 1905 expenses and construction costs are reported in Kronen and are converted into British pounds using the official exchange rate of 0.042 British pounds per Kronen. Ton miles and passenger miles are taken from Mitchell (1992). Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1883 to 1912. The prices are reported as the average value at the pit's mouth in shillings. The prices appear to be reported using the official exchange rate for Kronen. Weekly wages for unskilled workers are taken from Mesch (1984). They converted into British pounds using the exchange rate from 1905. The wages for skilled workers are equal to the weekly wages of unskilled in shillings multiplied by the ratio of craftsman to builders' wages in Leipzig drawn from Allen (2001). All financial variables are deflated using the Austrian consumer price index from Global Financial data with base year 1905.

Australia: The complete data cover the years from 1905 to 1912. The Statistical Abstracts provide data on railway miles, passengers carried, tons shipped, train miles, and construction costs. Total expenses are estimated from working expenses using the formula discussed in the text. Expenses and construction costs are expressed in British pounds. Coal prices per metric ton are reported in Returns relating to the Production, Consumption, etc. of Coal from 1883 to 1912. The prices are reported as the average value at the pit's mouth in shillings. Wages for unskilled and skilled workers are taken from the Statistical Abstracts for British Colonies. They are already expressed in British pounds. All financial variables are deflated using the Australian consumer price index from Global Financial data with base year 1905.

References

Allen, Robert. "The Great Divergence in European Wages and Prices from the Middle Ages to the First World War." *Explorations in Economic History* 38 (2001), pp. 411-447.

Arnold A. and S. McCartney. "Rates of return, concentration levels and strategic change in the British railway industry 1830-1913." *Journal of Transport History* 25 (2005), pp. 41 60.

Bjorklund, Jorgen and Hans Stenlund. "Real Wages in Sweden, 1870-1950: a Study of six Industrial Branches." In Scholliers, P. and Zamagni, V., *Labour's Reward: Real Wages and economic change in 19th and 20th century Europe*, New York: Edward Elgar, 1995.

Bogart, Dan. "Nationalizations and the Development of Transport Systems: Cross-Country Evidence from Railroad Networks: 1860-1912." *Journal of Economic History* 69 (March 2009): 202-237.

British Board of Trade, *The Statistical Abstract for the Principal and Other Foreign Countries*, Various Years.

British Board of Trade, *The Statistical Abstract for the Several Colonial and otherPossessions of the United Kingdom*, Various Years.

British Board of Trade, *The Statistical Abstract for the United Kingdom*, Various Years.

British Board of Trade, *Returns Relating to the Production and Consumption of Coal*, Various Years.

Broadberry, Stephen. *Market Services and the Productivity Race, 1850-2000*. Cambridge: Cambridge University Press, 2006.

Bureau of Railway Economics. *A Comparative Study of Railway Wages and the Cost of Living in The United States, the United Kingdom, and the Principal Countries of Continental Europe*. New York: 1912.

Cantos, Pedro and Joaquin Maudos, "Regulation and Efficiency: the Case of European Railways," *Transportation Research Part A* 35 (2001): 459-472.

Center for International Earth Science Information Network (CIESIN), Columbia University, 2007. National Aggregates of Geospatial Data: Population, Landscape and Climate Estimates, v.2 (PLACE II), Palisades, NY: CIESIN, Columbia University. Available at: http://sedac.ciesin.columbia.edu/place/.

Christopoulos, Dimitris, John Loizides, and Efthymios Tsionas. "Efficiency in European Railways: Not as inefficient as one might think," *Journal of Applied Economics* 4 (May 2001): 63-88.

Cianci, Ernesto. *Annali di Statistica series 6 Vol. 20* Rome: ISTAT, 1933.

Collins, William J. and Jeffrey G. Williamson. Capital-Goods Prices and Investment, 1870 1950. *Journal of Economic History 61* (March 2001): 59-94.

Crafts, Nicholas, Terence C. Mills, and Abay Mulatu. "Total factor productivity growth on Britain's railways, 1852–1912: A reappraisal of the evidence," *Explorations in Economic History* 4 (Oct 2007): 608-634.

Crafts, Nicholas, Timothy Leunig, and Abay Mulatu. “Were British Railway Companies Well Managed in the early twentieth Century?,” Forthcoming, *Economic History Review*.

Faith, Nicholas. *The World the Railways Made*. The Bodley Head, 1990.

Farsi, Mehdi, Massimo Filippini, and William Greene. “Efficiency Measurement in Network Industries: Application to Swiss Railway Companies,” *Journal ofRegulatory Economics* 20 (2005): 69-90.

Flandreau, Mark and Frédréric Zumer. The Making of Global Finance, 1880-1913, OECD 2004

Foreman-Peck, James. “Natural Monopoly and Railway Policy in the Nineteenth Century.” *Oxford Economic Papers* 39 (1987): 699-718.

Greene, William. “Reconsidering Heterogeneity in Panel Data Estimation of the Stochastic Frontier Model,” *Journal of Econometrics* 126 (2005): 269-303.

Gregory, Paul. *Russian National Income, 1885-1913*. Cambridge: Cambridge University Press, 1982.

Grytten, Ola H. “Nominal wages in Norway 1726–1940 by occupation,” in Eitrheim, Øyvind, Jan T. Klovland and Jan F. Qvigstad (eds.). *Historical Monetary Statistics for Norway Part II.* Norges Bank Occasional Papers no. 38, 2007.

Grytten, O.H. "A consumer price index for Norway 1516-2003", 47-98, *Historical Monetary Statistics for Norway 1819-2003*, Norges Bank Occasional Papers no. 35, Oslo, 2004.

Herranz-Locan, Alfonso. “The Spanish Infrastructure Stock, 1844-1935.” *Research in Economic History* 23 (2005): 85-129.

Herranz-Locan, Alfonso. “Railroad Impact in Backward Economies: Spain, 1850-1913.” *Journal of Economic History* 66 (Dec. 2006): 853-881.

Ichiro, Kakizaki. *Laying the Tracks: the Thai Economy and its Railways 1885-1935*. Kyoto University Press, 2005.

Jondrow, J., K. Lovell, I. Materov, and P. Schmidt. “On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model,” *Journal of Econometrics* 19 (1982): 233-238.

Kumbhaker, Subal and C.A. Knox Lovell. *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press, 2000.

Leacy, F. H. *Historical Statistics of Canada*. Ottawa, 1983.

Leunig, Tim Leunig, Timothy. "Time is Money: A Re-Assessment of the Passenger Social Savings from Victorian British Railways," *The Journal of Economic History* 66 (Sept. 2006): 653-673.

MacKinnon, Mary and Chris Minns. “The Costs of Doing Hard Time: A Penitentiary based Regional Price Index for Canada, 1883-1923,” *The Canadian Journal of Economics* 40 (May 2007): 528-560.

MacKinnon, Mary. “New Evidence on Canadian Wage Rates, 1900-1930,” *The Canadian Journal of Economics* 29 (Feb. 1996): 114-131.

Maddison, Angus. *The world economy: historical statistics*. Paris: Development Centre, OECD, 2003.

Margo, Robert A. , “Hourly and weekly earnings in selected industries and for lower skilled labor: 1890–1926.” Table Ba4314-4319 in *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*, edited by Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright. New York: Cambridge University Press, 2006. http://dx.doi.org/10.1017/ISBN 9780511132971.Ba4214-4544

Mesch, Michael. *Arbeiterexistenz in der Spatgrunderzeit*. Europaverlag Wien, 1984.

Millward, Robert. *Private and Public Enterprise in Europe*. Cambridge: Cambridge University Press, 2004.

Minster for Home Affairs, Australia. *Year Book of the Commonwealth of Australia,* 1910

Mitchell, B.R. *International Historical Statistics: Africa, Asia, & Oceania, 1750-1988*. New York: MacMillan, 1995.

Mitchell, B.R. *International Historical Statistics: The Americas, 1750-2000*. New York: Palgrave MacMillan, 2003.

Mitchell, B.R. *International Historical Statistics: Europe, 1750-1988*. New York, MacMillan, 1992.

Mulhall, Michael. *Dictionary of Statistics*. London: Routledge, 1892.

Parisio, Lucia. “A comparative Analysis of European Railway Efficiency: a cost frontier approach, *Applied Economics* 31 (1999): 815-823.

Raper, Charles Lee. *Railway Transportation: a History of its Economics and its relation to the State. New York*, Knickerbocker Press, 1912.

Republic of Argentina, *Ministerio de Obras Publicas. Estadistica de Los Ferrocarriles en Explotacion*, Various Years.

Scholliers, Peter. "A century of real industrial wages in Belgium, 1840-1939." In Scholliers, P. and Zamagni, V., *Labour's Reward: Real Wages and economic change in 19th and 20th century Europ*e, New York: Edward Elgar, 1995.

Siegenthaler, Hansjoerg. *Historische Statistik der Schweiz Chronos*, Zuerich, 1996.

Studer, Roman. "When Did the Swiss Get so Rich? Comparing Living Standards in Switzerland and Europe," Forthcoming *Journal of European Economic History*.

U.S. Department of Commerce. *Statistical Abstract of the United States*, various years.

Vernon-Harcourt, Leveson Francis. *Rivers and Canals*. Oxford: Clarendon, 1896.

Williamson Jeffrey G. "The Evolution of Global Labor Markets since 1830: Background Evidence and Hypotheses," *Explorations in Economic History* 32 (April 1995): 141 196.

Table 1: Ownership Patterns across Countries: 1880 and 1910

	1880		1910	
Country	Fraction of miles nationalized	Fraction of miles constructed by State	Fraction of miles nationalized	Fraction of miles constructed by State
Russia	0.00	0.04	0.19	0.44
Norway	0.00	0.94	0.00	0.84
Sweden	0.00	0.33	0.02	0.30
Holland	0.01	0.57	0.07	0.52
Belgium	0.24	0.44	0.43	0.50
France	0.02	0.07	0.12	0.10
Switzerland	0.00	0.00	0.53	0.06
Spain	0.00	0.00	0.00	0.00
Italy	0.22	0.22	0.35	0.45
Japan	0.00	1.00	0.55	0.35
US	0.00	0.00	0.00	0.00
Argentina	0.01	0.48	0.01	0.24
Britain	0.00	0.00	0.00	0.00
Germany	0.13	0.51	0.26	0.68
India	0.17	0.34	0.22	0.67
Canada	0.00	0.16	0.00	0.07
Austria	0.01	0.07	0.55	0.31
Australia	0.02	0.92	0.02	0.86

Sources: Bogart (2009)

Table 2: Geographic Variables

Country	Av. Elevation	St. Dev. Elevation	Skew Elevation	% of land area with 100km of coast	Miles navigable rivers and canals per square mile
Russia	266.85	294.1	2.04	15.46	17.97
Norway	400.6	313.39	1.03	67.45	1.59
Sweden	228.85	197.25	1.38	44.27	2.14
Netherlands	6.65	13.6	4.16	75.12	212.01
Belgium	115.8	127.47	1.19	38.56	94.55
France	244.6	305.43	2.54	35.71	37.78
Switz.	883	485.14	0.29	0	2.5
Spain	508.85	330.37	0.92	42.21	5.54
Italy	377.75	416.26	1.56	82.13	11.28
Austria	683	469.27	0.73	1.47	62.18
Japan	278.5	295.94	1.87	99.47	3.22
US	515.1	501.25	1.02	15.24	14.51
Argentina	522.25	781.08	2.33	12.1	1.97
Britain	116.7	128.22	2.32	94.27	31.42
Germany	186.1	174.43	1.07	17.97	81.82
India	268.3	448.49	3.27	15.27	3.28
Australia	207.25	153.82	1.3	19.75	1.36
Canada	333.5	325.86	2.04	29.48	0.34

Sources: see text.

Table 3: Inefficiency Estimates across Countries in the 1880s, 1890s, and 1900s

	Inefficiency score			Efficiency rankings		
Country	1880-89	1890-99	1900-12	1880-89	1890-99	1900-12
Belgium	0.105	0.094	0.051	7	14	1
India		0.253	0.052		17	2
France	0.128	0.079	0.056	8	10	3
Germany	0.138	0.095	0.059	10	15	4
US	0.062	0.094	0.073	3	13	5
Spain	0.136	0.082	0.075	9	11	6
Australia			0.080			7
Norway	0.086	0.092	0.088	6	12	8
Japan		0.075	0.093		8	9
Argentina		0.052	0.093		5	10
Britain	0.066	0.059	0.095	4	6	11
Russia		0.078	0.096		9	12
Sweden		0.043	0.104		3	13
Italy		0.036	0.118		1	14
Switz	0.045	0.068	0.119	2	7	15
Austria		0.043	0.129		2	16
Netherland	0.027	0.232	0.132	1	16	17
Canada	0.075	0.051	0.147	5	4	18
average	0.109	0.113	0.067			
correlation of inefficiency, 1880s, 1890s					-0.4	
correlation of inefficiency, 1890s, 1900s						-0.26

Notes: the estimates are based on the cost function specification described in equation (1) and formula (2).
Sources: see text.

Table 4: Summary of Inefficiency estimates under Alternative Specifications

	Baseline model	Alternative specifications								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	10
Average	0.092	0.092	0.047	0.048	0.088	0.046	0.077	0.033	0.076	0.044
St. Dev.	0.055	0.055	0.014	0.015	0.049	0.013	0.069	0.015	0.065	0.015
Min	0.017	0.017	0.02	0.019	0.018	0.021	0.015	0.013	0.015	0.017
Max	0.324	0.322	0.097	0.144	0.337	0.098	0.457	0.211	0.349	0.167
correlation with baseline model		0.996	0.965	0.966	0.988	0.943	0.961	0.873	0.964	0.778

Notes: Specifications 1, 2, and 3 do not assume a homogenous cost function. Model 2 adds the average haul and average trip. Model 3 adds the log of train miles divided by rail miles and drops the log of rail miles. Models 4, 5, and 6 assume a homogenous cost function. Model 5 adds the average haul and average trip. Model 6 adds the log of train miles divided by rail miles and drops the log of rail miles. Models 7, 8, and 9 not assume a homogenous cost function and replace the normal-half normal with the exponential, gamma, and truncated normal distributions. Model 10 use a translog cost function.

Table 5: Parameter Estimates for Selected Cost Function Models

Variable	(1)	(2)	(3)
Ln passenger miles	0.408	0.407	0.382
	(0.0455)	(0.045)	(0.061)
Ln ton miles	0.348	0.349	0.431
	(0.046)	(0.046)	(0.038)
Ln rail miles	0.281	0.257	
	(0.095)	(0.096)	
Av. Trip length		-0.003	
		(0.0018)	
Av. Haul		-0.394	
		(0.984)	
Ln (train miles/rail miles)			0.003
			(0.062)
Ln Coal prices	0.162	0.153	0.152
	(0.038)	(0.038)	(0.04)
Ln Unskilled wages	-0.177	-0.182	-0.198
	(0.076)	(0.075)	(0.076)
Ln Construction cost per mile	0.45	0.444	0.342
	(0.061)	(0.061)	(0.069)
Ln Skilled wages	0.33	0.33	0.325
	(0.071)	(0.071)	(0.073)
Year	0.66	0.054	0.191
	(0.326)	(0.325)	(0.347)
Year-squared	-0.00002	-0.00002	-0.00005
	(0.00009)	(0.00009)	(0.00009)
Country fixed effects	Yes	Yes	Yes
N	402	402	376
Log-Likelihood	386	388	360
Sigma v	0.062	0.061	0.086
Sigma u	0.117	0.117	0.058

Notes: the dependent variable is the Ln total expenses. Standard errors are in parentheses.

Sources: see text.

Table 6: Determinants of Inefficiency: Models without Country fixed effects

variable	(1)	(2)	(3)	(4)	(5)
fraction nationalized	0.052	0.069	0.062	0.078	0.014
	(0.018)***	(0.02)***	(0.022)***	(0.025)***	(0.007)*
fraction constructed by state	0.008	0.013	0.016	0.011	0.0005
	(0.009)	(0.009)	(0.009)*	0.011	(0.003)
Average elevation		0.00008	-7.00E-03	-5.00E-03	0.0006
		(0.00002)***	(3.00E-03)**	(3.00E-03)*	(0.0009)
St. dev. Elevation		-0.00005	-5.00E-05	-6.00E-05	-3.39E-06
		(0.00003)	(3.00E-05)*	(3.00E-05)**	(0.00001)
Skew elevation		0.022	0.023	0.0221	0.0023
		(0.004)***	(0.004)***	(0.005)***	(0.0015)
% land area within 100km		6.30E-06	-0.00002	-8.56E-06	0.00003
		(0.0001)	(0.0001)	(0.0001)	(0.00003)
Rivers and canals per sq. mi.		-4.08E-06	-0.00005	-0.00004	1.48E-06
		(0.00007)	(0.00007)	(0.00008)	(0.00002)
Year			-0.0017	-0.00125	0.00004
			(0.0006)***	(0.0007)*	(0.0002)
Year*average elevation			3.79E-06	3.14E-06	-3.40E-07
			(1.59E-06)**	(1.65E-06)*	(4.78E-07)
Gdp growth t				-0.049	-0.012
				(0.062)	(0.018)
Gdp per capita t-1				3.16E-06	8.00E-07
				(3.67E-06)	(1.06E-06)
Average age of track				-0.0009	-0.00018
				(0.0005)*	(0.0001)
constant	0.083	0.0293	3.407	2.425	-0.0529
	(0.004)***	(0.0155)*	(1.18)	(1.32)*	(0.384)
N	402	402	376	402	402
R-square	0.03	0.1	0.12	0.13	0.02

Notes: the dependent variable is the Ln total expenses. Standard errors are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% levels respectively.

Table 7: Determinants of Inefficiency: Models with Country fixed effects

variable	1	2	3
fraction miles nationalized	0.194	0.187	0.027
	(0.03)***	(0.029)***	(.009)***
fraction constructed by state	-0.572	-0.57	-0.032
	(0.093)***	(0.09)***	(0.028)
Gdp growth t		0.004	-0.003
		(0.06)	(0.018)
Gdp per capita t-1		6.00E-05	6.12E-06
		(1.00E-05)***	(3.75E-06)
average age of track		0.001	0.0004
		(0.002)	(0.0006)
year	-0.0007	-0.003	-0.0006
	(0.0003)***	(0.001)***	(0.0004)
country fixed effects	yes	yes	yes
N	402	402	402
R-square (within)	0.14	0.2	0.03

Notes: the dependent variable is the Ln total expenses. Standard errors are in parentheses. *, **, *** indicates significance at the 10%, 5%, and 1% levels respectively.