Banco de México

Documentos de Investigación

Banco de México

Working Papers

N° 2014-14

Railroads and Economic Growth: A Trade Policy Approach

Fernando Pérez-Cervantes
Banco de México

June 2014

Documento de Investigación
2014-14

Working Paper
2014-14

Railroads and Economic Growth: A Trade Policy Approach*

Fernando Pérez-Cervantes†
Banco de México

Abstract: What was the impact of railroads in the output of the United States during the 19th century and how can a New Trade model help answer this question? In order to respond I follow three steps. First, I construct a new digital railroad data set and pair it with geographic and topographic features of the U.S. territory to estimate travel times between every pair of U.S. counties for every year between 1840 and 1900. Second, I use these results, together with a Ricardian model of trade and U.S. county output data from the 19th century, to estimate county gains from trade using a fixed-point algorithm. Third, I estimate counterfactuals with the railroads built up to a certain year. My estimates suggest that there was a lot of migration anticipating railroad construction and not the other way around. However, leaving all factors of production fixed, if the railroads were made suddenly unavailable in 1890 there would have been a 9.6 % reduction in output, but in 1900, after the financial crisis, the impact would have been less than 9 %.
Keywords: Railroads, Trade, Gravity Models.
JEL Classification: F10, F14, N71.

Resumen: ¿Cuál fue el impacto de los ferrocarriles en el producto de Estados Unidos durante el siglo XIX, y cómo puede un modelo de New Trade ayudar a responder esta pregunta? Para responder sigo tres pasos. Primero construyo una nueva base de datos digital de trenes, y le añado características geográficas y topográficas de Estados Unidos para estimar tiempos de traslado entre todos los pares de condados de Estados Unidos para cada año entre 1840 y 1900. En segundo lugar, uso estos resultados junto con un modelo ricardiano de comercio y datos de producto de los condados de Estados Unidos en el siglo XIX, para estimar la ganancias del comercio a nivel condado usando un algoritmo de punto fijo. Tercero, estimo contrafactuales con los trenes construídos hasta cierto año. Mis estimaciones sugieren que hubo mucha migración anticipando la construcción de ferrocarriles, y no al revés. Sin embargo, dejando todos los factores de producción fijos, si los ferrocarriles quedaran súbitamente no disponibles en 1890 habría habido una reducción de 9.6 % en el producto, pero en 1900, después de la crisis financiera, el impacto habría sido menor al 9 %.
Palabras Clave: Ferrocarriles, Comercio, Modelos Gravitacionales.

---

*I am very grateful to Thomas Chaney, Robert E. Lucas, Jr., and Samuel S. Kortum for their constant support, and many helpful comments and discussions. I have also received valuable feedback from Lorenzo Caliendo, Rick Hornbeck, Julene Iriarte, Kerem Kosar, Ralph Ossa, Todd Schuble, and all seminar partici- pants in the International Economics and Capital Theory Working Groups at the University of Chicago. I thank Chris Winters for all his help at the Regenstein Library Map Room, and Steve Mohr for setting up Sparta and Acropolis, the computers that made all this possible in less than a lifetime.

† Dirección General de Investigación Económica. Email: fernando.perez@banxico.org.mx.

1 Introduction

The year 1840 saw a total of 455 miles of railroads opened in the United States. This investment equaled 1% of the total U.S. value of output that year,[1] an addition of almost 18% to the total railroads built up to that time.[2] In 1887, the peak year of railroad construction, nearly 12,000 miles were built, at a cost of around 3% of GNP.[3] By the end of the 19th century, GNP per capita grew by a factor of 3.97 compared to 1840. In this paper, I do an extensive compilation of almost 118,000 miles of railroad-openings and convert it into digital form. I use detailed U.S. county-level census data from the 19th century (around 500,000 observations), a simple structural model of trade with 3109 trading regions, and geographic and topographic characteristics of every county and waterway in the United States to make a quantitative assessment about the U.S. railroad network: I estimate the impact on output if they became suddenly unavailable. I follow three steps to get to this calculation:

1. I find travel times for every pair of counties for every year between 1840 and 1900. This result does not require any economic model, and uses only the fact that optimal travel times satisfy the triangle inequality.

2. I use a Ricardian model to find the gains from trade as the railroad network continued to grow. This analysis, which uses a fixed-point algorithm, works for a large family of

[1]Total value of output (both at the county and at the national level) will be denoted as GNP henceforth to make it comparable with other results in the literature, and because large extensions of land that belonged to more than one county were counted as being in only one county. This model, which uses an immaterial single factor of production to produce a homogeneous final good, is equivalent to other output specifications such as GDP. But extending the analysis to more than one sector or with more factors of production would disentangle the two measures, so I keep GNP to make this simple model and its own natural extensions comparable. The total value of output in this paper is the sum of the total value of agricultural output and total value of manufacturing output. See Donaldson and Hornbeck (2012) for a similar analysis that includes the cost of land.

[2]In 1840, railroad construction cost an average of $25,000 per mile (Stover (1999)), and the value of output was $1.13 billion (Haines et al. (2005)). I measure length of construction using standard GIS geodistance calculators, using a data set of opened railroads in 1840. This cost per mile was inexpensive compared to the railroad's European counterparts. By the same date, Europe had less than half the miles of railroads at more than the total cost of the entire 3,000 miles of American railroads (Stover (1997)).

[3]The cost of railroad construction was 4% of 1880's GNP and 2.5% of 1890's GNP (Haines et al. (2005)).

preferences and trade model specifications, and works without bilateral trade data, which is usually incomplete or unavailable to researchers as is the case in this paper.

3. I build a simple counter-factual scenario in which railroads are unavailable but all the other means of transportation are.

Railroad construction does not increase output because it is an input of production, but rather because it is a reduction in trade costs.[4] My model assumes all factors of production are fixed, and so trade-cost reduction changes relative prices and patterns of specialization until factor markets clear. The introduction of railroads reduced trade barriers; thus the decision regarding whether to build these railroads became a de facto trade-policy problem.

Railroads are not the usual type of trade-barrier reduction. First, the cost of implementing them is high. In 1840, a 1000-mile railroad to connect New York and Chicago would have cost $25 million, the equivalent of 2.2% of the United States' GNP, or 13% of New York state's GNP, or 96% of Illinois's GNP. My estimates suggest that the present discounted value of the increase in output that this railroad would have produced is 1.3% of the GNP of the United States.[5] Modern trade-policy problems (even involving extremely distant and small regions) do not typically have such large costs of implementation, even if negotiations are included.

Because agents re-optimize travel routes, railroads construction (unlike tariff reduction) is a trade-barrier reduction that preserves the triangle inequality: the construction of a single road linking two regions in the United States weakly reduces trade costs between all regions in the United States. For example, between the years 1865 and 1869, the Union Pacific Railroad between Omaha and San Francisco was built. This railroad thus became part of the optimal route to transport goods from East to West, and reduced the length of a trip between Seattle

[4]See Glomm and Ravikumar (1994), Munnell (1992), and Winston (1991) for examples in which infrastructure changes factor productivity.

[5]Assuming population shares fixed and a discount factor of 2.5%. See Section 5 for details on these calculations.

and New York from 52 to 17 days, even when Seattle was 700 miles away from the nearest railroad to the East in 1869.[6]

I study railroad network design as a special type of trade-policy problem, and because I do not have bilateral trade-flows data, I use computational image processing and a fixed-point algorithm to back out the values of trade shares for all the counties in the United States, for every census between 1840 and 1900. The values of these trade shares and their responses to additional trade cost reductions allow me to conclude that the social gains from the actual railroads equaled 9% of GNP in 1900.

I build on seminal previous work to estimate both the counter-factual output of removing the railroads, and I am able to compare it to other results in the literature.[7]

Fogel (1962) conducted a classic first-order estimate of the effects of railroads on the economy, by finding the increase in transportation costs that removing the railroads produces. He made exhaustive accounting on freight and inventory costs and found that the existence of the railroads in 1890 saved the U.S. economy 2.7% of its GNP had only wagons and boats been able to transport the same amount of output. He also studied other counter-factuals, such as the impact of a series of canals in the Midwest and the improvement in the . Previously, in Fogel et al. (1960), an exhaustive cost/benefit analysis of the construction of the railroads to the Pacific Ocean was done, finding that this railroad was built too early. This paper also makes use of optimal routes using combinations of wagons, boats, and railroads to approximate trade costs.

Donaldson (2008) uses an Eaton and Kortum (2002) model to estimate the outcome of removing railroads and of adding 40,000 kilometers of railroads in India that were designed but never built. He found that the arrival of the railroads produced an average 16% increase in real output in the districts that became part of the network. He also found that the railroads that

[6] See Section 3 for details on these calculations.

[7] See Mitchell (1964), Stover (1997), Holmes and Schmitz Jr (2001), Summerhill (2003), Herrendorf, Schmitz, Teixeira, and of Minneapolis. Research Dept (2009) for a nice analysis of the impact of railroads on productivity, competition, and on specific industries.

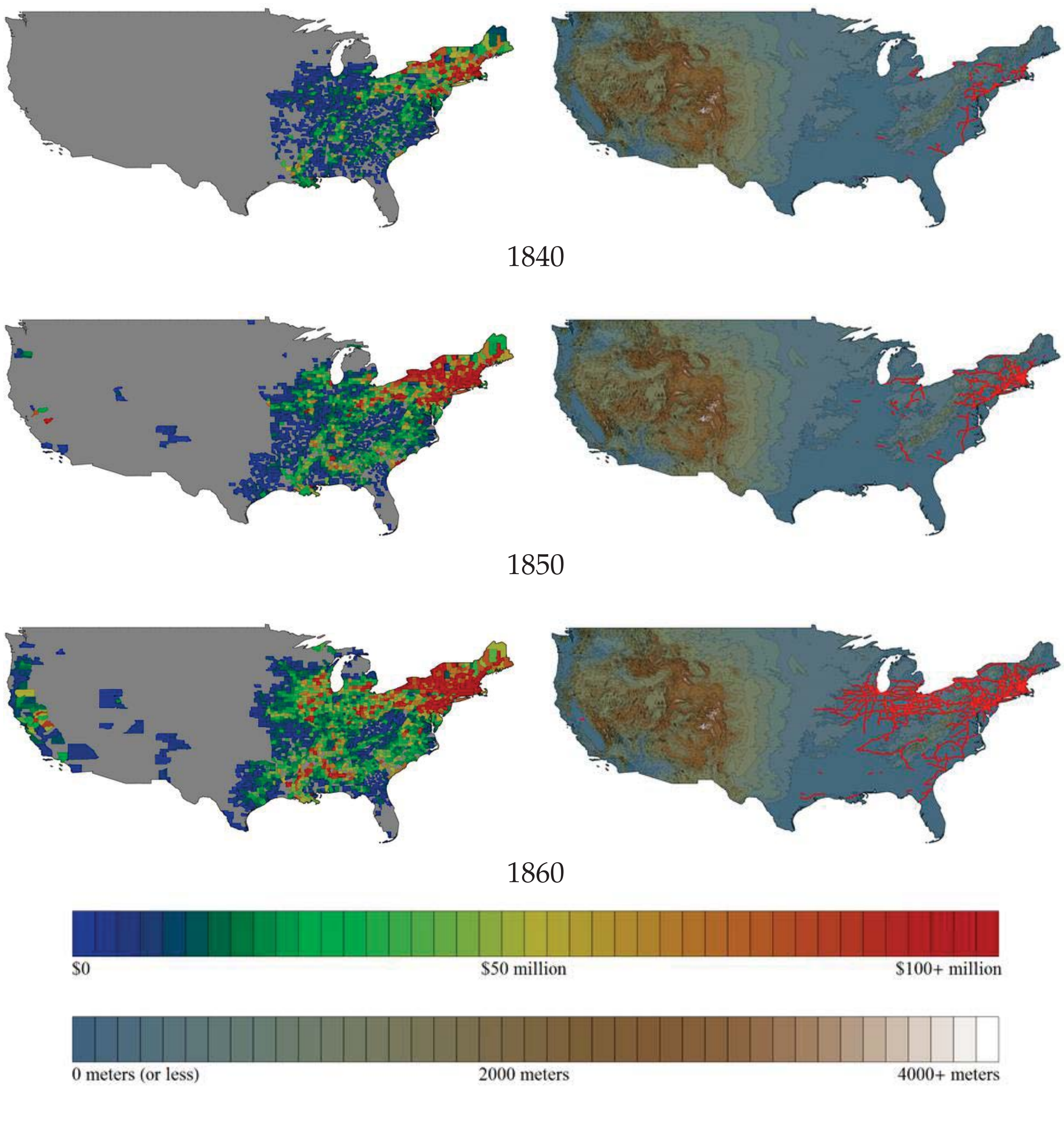
were not built would have not improved output significantly. More recently, Donaldson and Hornbeck (2012) estimate a similar counter-factual as Fogel (1962) for the United States, but with a Ricardian model of trade and allowing for migration. They estimate that removing the railroads in 1890 would decrease GNP by 6.3%. This paper extends this literature by exploring the normative implications of these policies.

Section 2 of the paper explains how the planning of the U.S. railroad system could be seen as a centralized trade-policy problem. Section 3 describes how I obtained optimal travel times between every pair of U.S. counties for every year between 1840 and 1900. Section 4 uses the results of the previous section and, together with a Ricardian model of trade and U.S. county output data from the 19th century, describes how to estimate the gains from trade due to the railroad. Section 5 will describe the counter-factual analysis. Section 6 summarizes the results and concludes. Appendix A contains a description of the computational image-processing methods used to identify trade costs, and Appendix B contains the proofs of the well-known results of the Eaton and Kortum (2002) model.

2 Historical Background

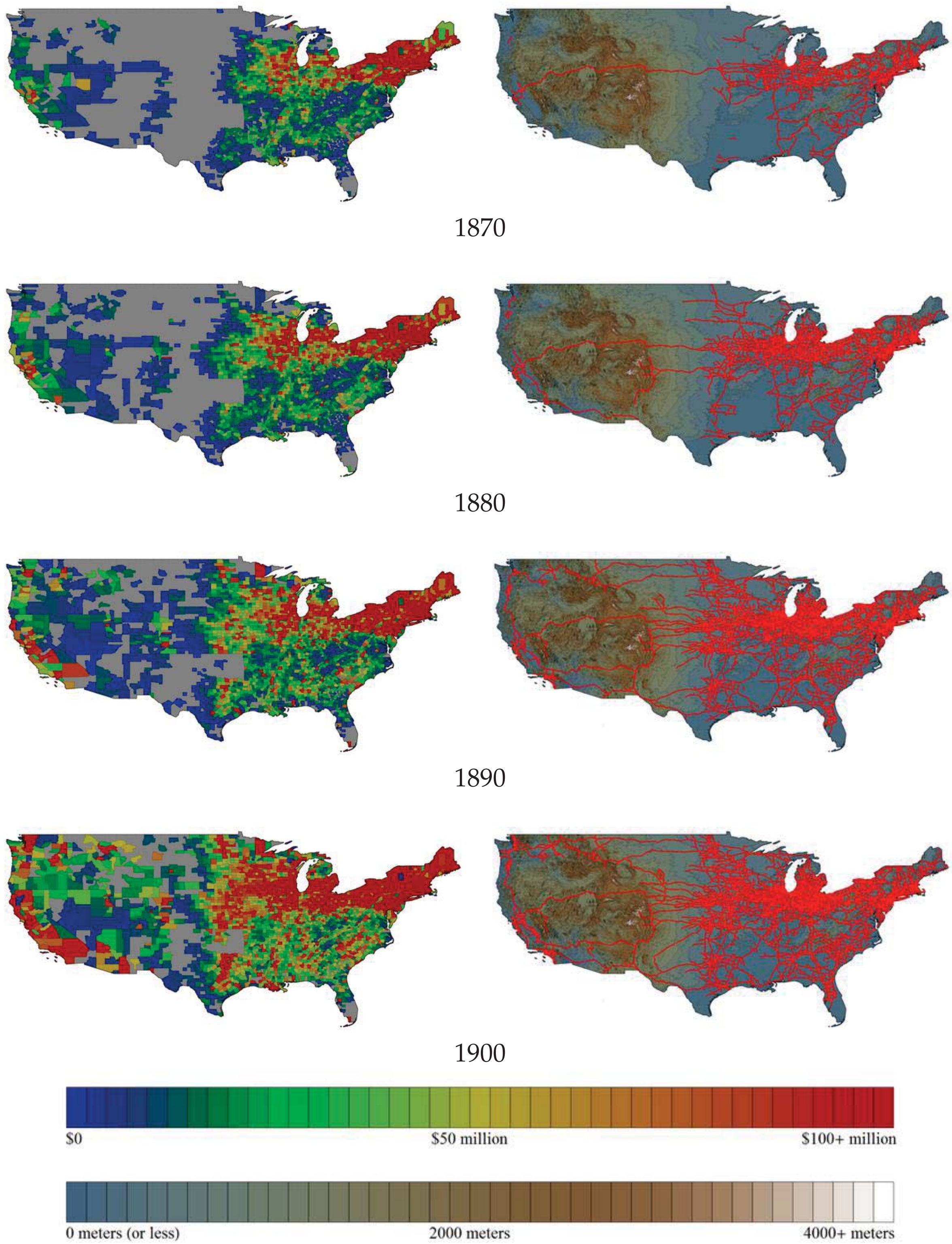
Between 1840 and 1900, the railroad network in the United States grew at an average annual rate of 6.47%, resulting in just over 130,000 miles of built railroads (118,000 of them opened and operating by the year 1900). By contrast, real GNP grew 4.9% and population 2.52% per year on average over this period, and the U.S. territory tripled in size. Figure 2.1 and Figure 2.2 show the expansion over time of this network of opened railroads, some geographic traits of the U.S. territory, and the evolution of the county-level GNP in the United States. Growth in output and railroad construction are clearly related. From these figures, we can see short railroads were usually built where output was large, over flat areas, and avoiding mountainous regions.

Figure 2.1: GNP and railroad construction 1840-1860



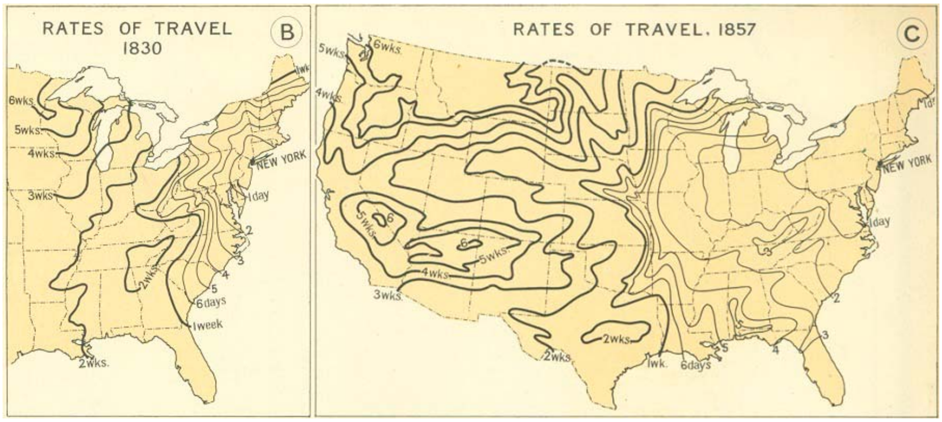
Note: Left: County Value of Output (GNP) in 2010 U.S. dollars during census. Gray areas have no data for that census. Right: Railroad openings in the United States up to each date, and altitude of the U.S. territory.

Figure 2.2: GNP and railroad construction 1870-1900



Note: Left: County Value of Output (GNP) in 2010 U.S. dollars during census. Gray areas have no data for that census. Right: Railroad openings in the United States up to each date, and altitude of the U.S. territory.

Figure 2.3: Time required to travel from New York to various western locations



Note: Images from the Atlas of the Historical Geography of the United States (Paullin and Wright (1932))

The construction of the railroads in the United States, beginning during the 1830s is the most important part of a sequence of actions by large East Coast cities to promote trade with less populated Western regions. Figure 2.3 shows archival data from Paullin and Wright (1932), and pictures the evolution of travel times from New York City to the rest of the territory of the United States. It is clear that railroads influenced travel times, and that regions left far away from the construction were left out of the optimal routes to New York. Price data from Carter, Gartner, Haines, Olmstead, Sutch, and Wright (2006) show that shorter travel times give less resources dedicated to transportation (which implies cheaper costs per ton-mile), but other important margins were also improved, such as the volume and weight of the freight, and the fact that railroads could be used all year long.

However, the construction of the railroads was not a centralized decision. These actions involved mainly canals and railroads financially backed by state governments, and soon the race to be the "first commercial city in the world" became fierce (Stover (1997)). State governments went to extreme measures to keep up the pace in this race. The actions taken during the 1830s to have access to cheaper agricultural goods from farther distances did not go over well with all citizens, particularly large-scale producers in upstate regions far away from the construction projects.[8] The implementation of state taxes to pay for construction did not have popular approval either, even from the people benefiting from the construction.

To avoid raising taxes, state governments designed a system of "benefit taxation" (a property tax that was calculated after changes in property value due to railroad construction) and "tax-less finance" (which involved taxes only if both the banks and the canal/railroad company became insolvent).[9] This system worked for a while, but the drought and recession of 1839 caused a financial disaster of such magnitude that by 1842, eight states (and the Florida Territory) were in default, and three other states were about to default. Most local Constitutions were amended (and written all over again for 12 states) to limit the ability of the states to subsidize infrastructure and lend their credit, and to prevent states from entering long-term financial commitments without raising taxes.[10] The new laws shut down state financing of large infrastructure projects, and left the federal government as the most important (if not the only) source of collateral for those projects. Starting in 1840, any large U.S. railroad investment required financial guarantees to foreign investors that only the federal government could provide. A large number of corporations were created and the total capital raised for railroad construction in the 19th century reached the equivalent of 70% of GNP by 1900. Such events could not have happened had the government not been the ultimate source of collateral. If the corporations wanted to be able to pay for the debt acquired for the railroad construction, they were benefited if railroads were located in areas of high trade volume. If the federal government was responsible for repaying any failed projects, it benefited from the growth created by railroads.

Railroad construction is not a homogeneous trade cost reduction, which means that the benefits are unequally spread among the trading regions. In fact, some regions can be negatively affected if trade costs are reduced far away.[11] One of the most important problems that emerges

[8] For example, when the Erie Canal was built (northwest of New York City), the farmers on Long Island (east of New York City) opposed to the project, as documented by English (1996).

[9] The terminology is from Wallis, Sylla, and Grinath (2004).

[10] See Wallis (2004) for details on the specific changes to each state's constitution regarding this matter.

[11] This is a typical result in most trade models. Donaldson (2008) estimates that railroad construction in neighboring districts in India reduced output by 4% in the average district in India.

in the literature of network formation is the presence of externalities. The next sections will describe the three steps I follow to find the effects of new railroads on the U.S. counties and in the U.S. economy as a whole.

3 Obtaining Optimal Travel Times

This section describes the computational methods used to find the optimal travel times between every pair of U.S. counties for every year between 1840 and 1900 using any combination of wagons, boats, or railroads. Subsection 3.1 describes how I reduced the size of the problem so a modern computer can solve the problem. Subsection 3.2 and Subsection 3.3 describe how I obtained travel times between counties.

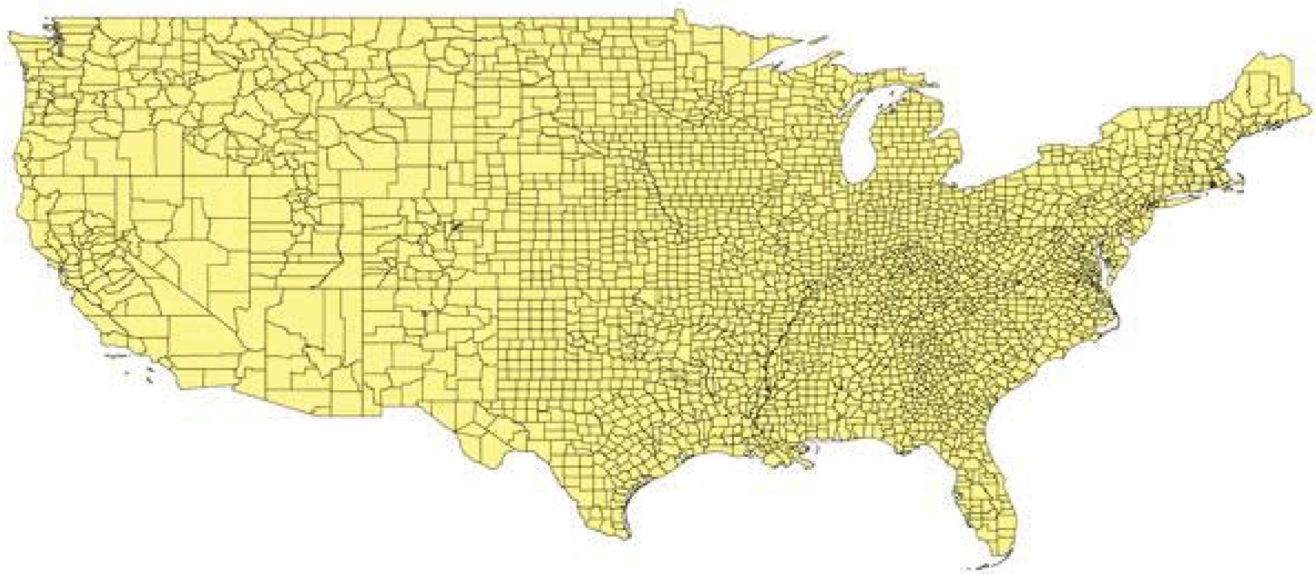
3.1 Simplifying the problem

The problem needs to be simplified in many aspects. First, I abstract from the fact that the U.S. territory is continuous, and assume trade will occur between isolated points. In particular, the set of points that I choose are the 3109 current county population centroids of the continental United States.[12] I assume every county $n \in \{1, 2, ..., 3109\}$ in the continental United States is a small open economy. Because detailed county data are only available from the census (which occurred every decade) I only study seven time periods: t={1840, 1850, 1860, 1870, 1880, 1890, 1900}.

I use the geographic characteristics of the 3109 present counties of the continental United States, which are shown in Figure 3.1. For the purposes of modeling transportation costs δ_{ij}, I will assume the present population centroids of each county are points on a two-dimensional space (latitude and longitude).

[12] See Donaldson and Hornbeck (2012) for an interesting calculation of intra-county trade costs, a type of analysis that treats the territory more like a surface than a set of points.

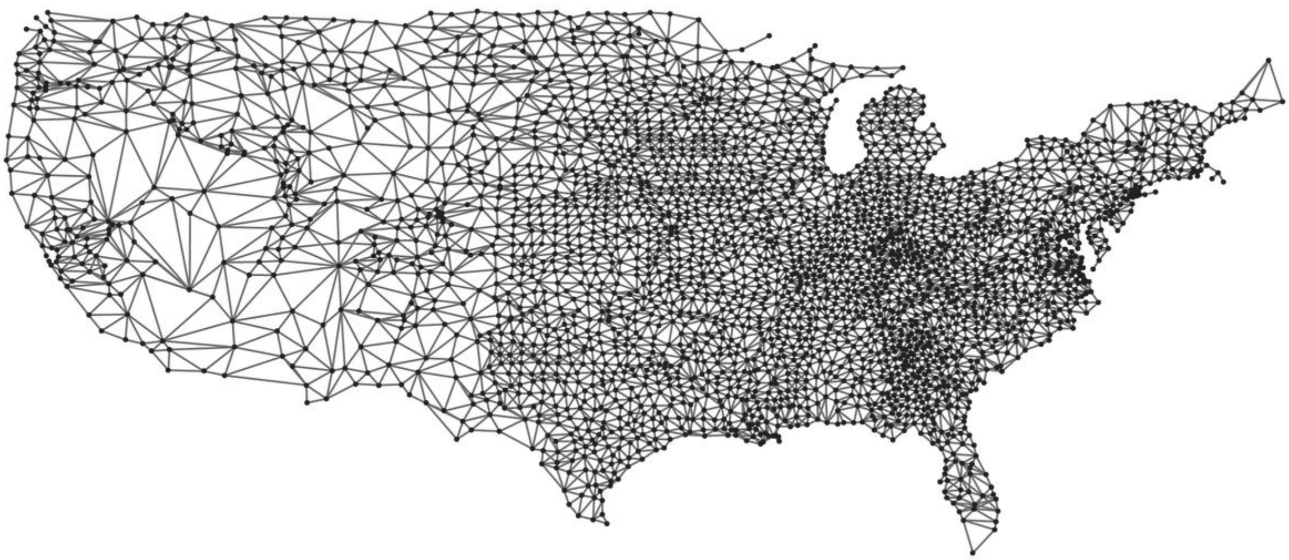
Figure 3.1: The 3109 present counties of the continental United States



I reduce the size of the problem by assuming the 3109 isolated points can only be joined via direct links with neighbors. A link can only go between two present neighboring county maps, and the distance between two counties will be the geodesic distance between its population centroids. Two counties that are not neighboring will have to use more than one link to be connected. The distance between counties is defined as the geodistance between points, to adjust for the fact that because the Earth is a sphere, and the United States is in the northern hemisphere, distances measured in northern regions could be over-estimated if we instead used a Euclidean norm using latitude and longitude.

I constructed a set of neighboring counties for every county in the continental United States. This set of neighbors reduced the number of potential ways of directly joining counties by one order of magnitude, an average of six neighbors per county in continental United States. I have 9092 pairs of neighbors and the longest link is 283 miles long. Without the neighbors-only restriction, the number of neighboring pairs would be quadratic in counties, not to mention the real world ambiguity of having a 2580 mile link between San Francisco and New York that does

Figure 3.2: Population Centroids and County Neighbors



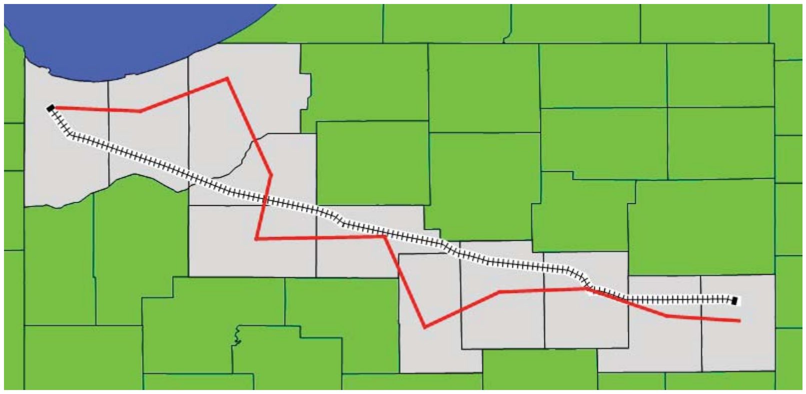
Note: Population centroids of the 3109 continental U.S. counties and the associated neighboring county pairs.

not touch any other county in the middle. This simplification turns out to be useful. The size of the problem has become manageable for a modern computer, because I choose between 9092 links instead of 4.6 million.

3.2 Combining data on railroad construction and county data

I have compiled, from different sources, an extensive dataset of the railroad construction projects in the United States from 1827 to 1900. The dataset consists of geospacial data of 12,800 railroad projects. A project consists of a series of ordered points indicating the latitude and longitude of the railroads in the continental United States. For each project I also have the year it opened. This classification chopped original railroads into short segments. Geospacial data for aban-

Figure 3.3: Simplification of railroads



Notes: Black-dotted: real railroad. Gray: the set of counties that the railroad touches. Red: simplified railroad connecting all the counties the railroad touches.

doned, two-way, or incomplete railroad construction were not available for my data set and add up to just over 9% of the total mileage of my data set, up to 1890.[13]

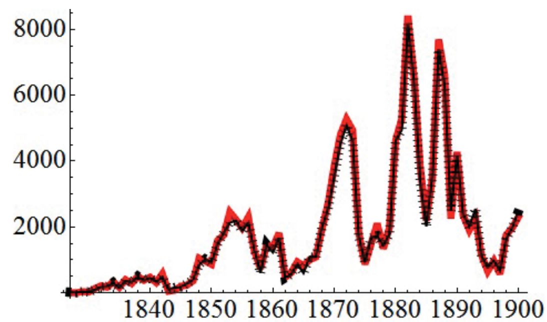
Using computational image-processing,[14] I can determine the counties through which each railroad project goes. Using this information, I create a simplified railroad data set containing the set of links that constitute the shortest way to connect all the county centroids that the railroad project traverses, with the restriction that only neighboring links can be used to build the railroad.[15] By construction, a project that belongs to only one county will be dropped, because by definition, counties are connected with themselves. As Figure 3.4 shows, the removal of projects and the mismatch of lengths due to centroid simplification add up to very small errors, both in the cross section and in the cumulative data. The most important information is preserved. I then repeat the process of data simplification with rivers, oceans, and canals, as seen in Figure 3.5. In this model, all water routes will follow coastlines and I am assuming that travel times upstream and downstream are the same.

[13]There were 129,774 miles of constructed railroads up to 1890 according to Depew (1968) vs. 117,708 miles of opened railroads in my data set, a difference of 9.3%

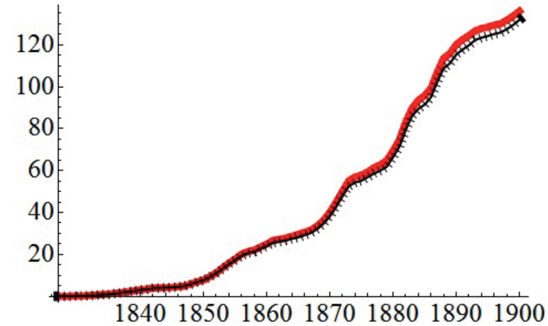
[14]See Appendix A for details on the procedure.

[15]This set of links is called the minimum spanning tree of the county population centroids. See Figure 3.3 for details on this procedure.

Figure 3.4: Real and simplified data over time



Miles of railroads opened each year



Accumulated Opened Miles (Thousands)

Notes: Red: simplified data (3,343 lines, over a finite set of nodes). Black-dotted: real railroad data (12,800 railroads over a continuum space),

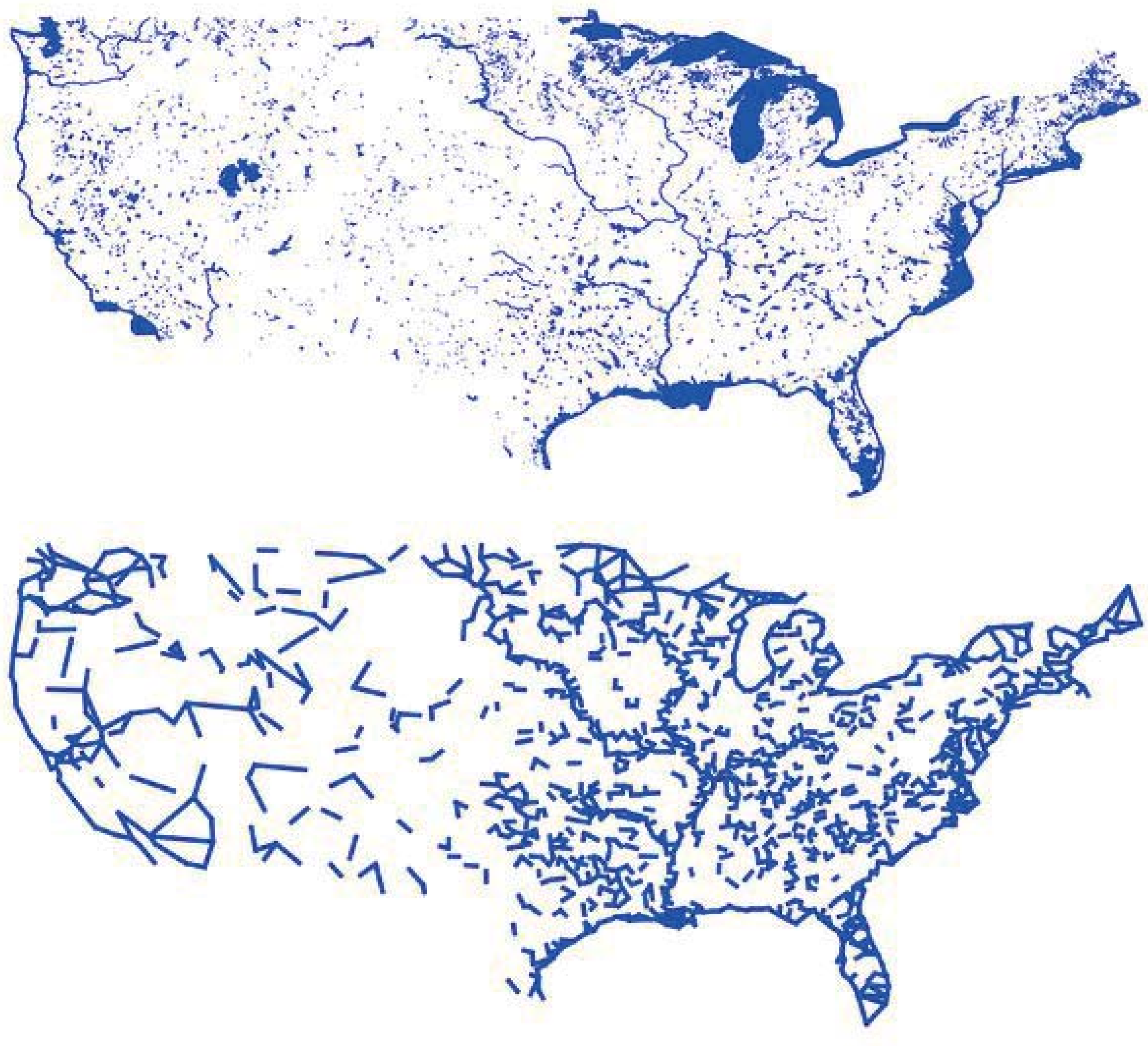
Labeling train transportation as faster than water transportation, which in turn was faster than wagon transportation, Figure 3.6 shows the fastest way to connect two adjacent counties. Note that using direct links to go from one county to its neighbor (especially the red ones near black ones) may violate the triangle inequality in travel time; that is, two adjacent counties can be connected more quickly via a third county using trains than directly but using horse-pulled wagons.

3.3 Using Dijkstra's algorithm to find optimal travel times

I will define the matrix of optimal travel times for every decade, $d_{ij}^{(t)}$, as the fastest time to go from county j to county i using any combination of wagon, boat, and railroad, for every pair $(i, j) \in \{1, 2, ..., 3109\} \times \{1, 2, ..., 3109\}$ available during that time. I obtain the elements of the matrix $d_{ij}^{(t)}$, which satisfy triangle inequality by construction, as follows: links that only use horses and wagons are calibrated to have a speed of 30 miles per day.[16] Boat transportation is

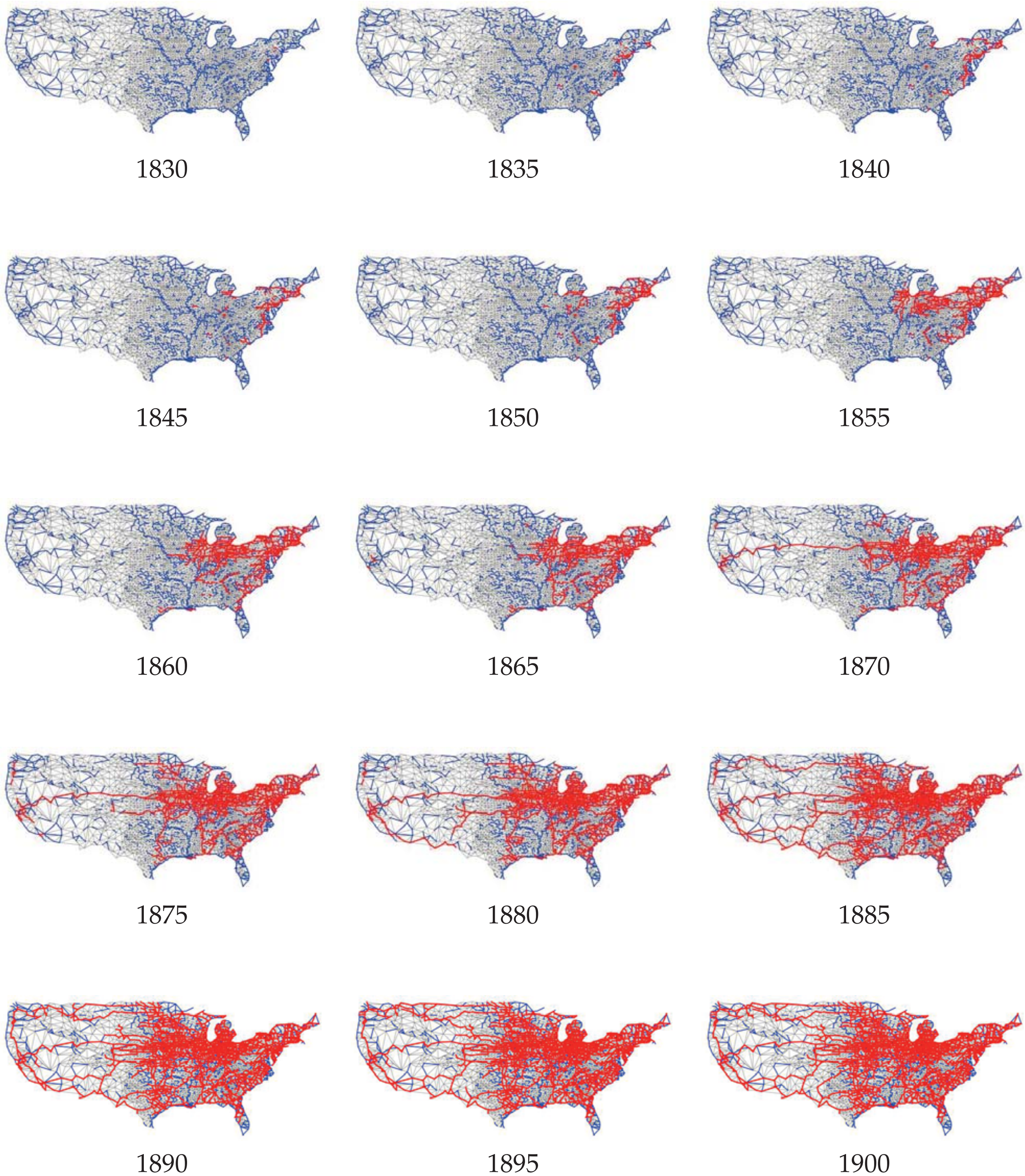
[16]According to Stover (1997), a wagon pulled by four horses, from New York City to Philadelphia (90 miles) in three days. It became known as the "Flying Machine".

Figure 3.5: Simplification of the navigable water data set



Notes: Top: Navigable water transportation in 1830. Bottom: Simplified dataset of navigable water transportation in 1830

Figure 3.6: Fastest route to connect two adjacent counties using only one link.



Notes: Black links: horse-pulled wagon. Blue links: boat. Red links: railroad.

calibrated to travel 120 miles per day.[17] Train speed is calibrated to 350 miles per day.[18] Assuming no switching costs, I use Dijkstra's algorithm to obtain the length of the fastest route between county centroids, which gives a matrix of 3109 by 3109 freight times, each entry being the solution of an optimization problem with 9092 unknown variables. Figure 3.7 shows one row of the matrix of $d_{ij}^{(t)}$, the row corresponding to New York County (Manhattan). Also note the 1860 sub-figure in Figure 3.7 appears to the casual eye very similar to Stover (1999), who digs into an enormous amount of archival data on route logs. This sub-figure also almost matches the most important cities in the raster data from Paullin and Wright (1932). This similarity indicates switching transportation did not represent a large hold-up problem in terms of time, and that travel speeds were homogenous throughout the territory.[19]

4 Calibrating Gains from Trade

This section sets up the model of trade, and explains an innovative method for parameter identification (which can be extended to a large family of trade models without loss of generality) to obtain gains from trade of every county for every census year between 1840 and 1900.

4.1 The Trade Model

The representative consumer in county n at time t has CES preferences for consumption over a continuum of goods $c_n^{(t)}(j)$ with $j \in (0,1)$ and constant elasticity of substitution $\sigma > 1$:

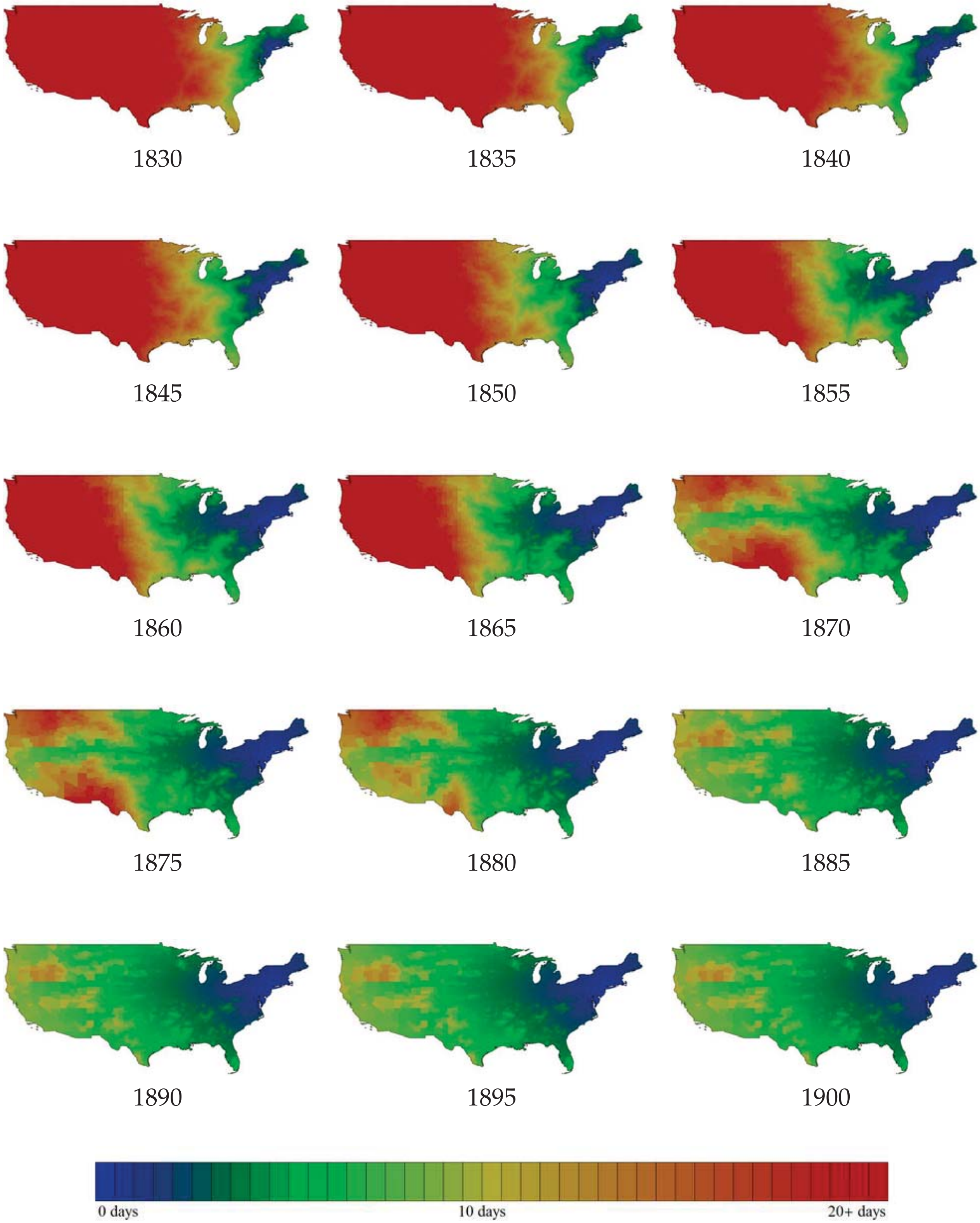
$$U_n^{(t)} = \left(\int_0^1 c_n^{(t)}(j)^{\frac{\sigma-1}{\sigma}} dj \right)^{\frac{\sigma}{\sigma-1}}$$

[17] As documented in Depew (1968).

[18] Trains traveled at a maximum at 20 miles per hour during most of the 19th Century, and the average shipment that involved exclusively train transportation traveled between 300 and 400 miles per day (documented by Stover (1999)).

[19] See Fogel (1962) and Donaldson and Hornbeck (2012) for estimates on transshipment costs when different types of transportation methods are used.

Figure 3.7: Length of time to travel from New York City



Notes: Destinations are every population centroid in the continental United States, using the fastest combination of wagon, boat, and railroad in the simplified dataset.

The representative consumer in county n has income $Y_n^{(t)}$ and spends it all in the county where she lives. The unit price of each good $c_n^{(t)}(j)$ in county n at time t is $p_n^{(t)}(j)$. The environment will be perfectly competitive. The representative consumer in county n will only purchase from the cheapest source of good j.

Goods are traded between counties at a cost. Goods that need to travel farther from source to destination will have a larger per-unit trade cost. Let $\delta_{ni} \geq 1$ be the iceberg cost of shipping goods from county i to county n. It equals the units of good j that must be shipped from location i in order for one unit of j to arrive to location n. I normalize $\delta_{nn} = 1$ for all $n = \{1, 2, ..., 3109\}$. Iceberg costs satisfy triangle inequality; that is, $\delta_{ni} \leq \delta_{nm}\delta_{mi}$ for any value of m.

Let $\kappa_i^{(t)}(j)$ be the cost in county $i = \{1, 2, ..., 3109\}$ to produce good $j \in (0, 1)$ at time t. Then by arbitrage, the per-unit price of good j in location n will equal $p_n(j) = \min_i \left\{\kappa_i^{(t)}(j)\, \delta_{ni}^{(t)}\right\}$.

County n has a fixed endowment of the single factor of production L_n. Every county has, for every good $j \in (0, 1)$, productivity draws which have a Frechet distribution with productivity parameter T_n and dispersion parameter θ.[20] All computational methods developed in this section (i.e. before the counter-factuals are defined) can relax the identification of this input of production. Specifying whether a single factor or multiple factors of production produce output is not necessary, as long as the production function is assumed to be Cobb-Douglass with constant returns to scale.

The representative consumer in county n has income $Y_n^{(t)}$ and solves:

$$\max_{c_n^{(t)}(j)} U_n^{(t)}$$
$$\text{s.t.} \quad \int_0^1 p_n^{(t)}(j)\, c_n^{(t)}(j)\, dj \leq Y_n^{(t)}$$

[20]See Eaton and Kortum (2002) for details on this technology specification.

the solution is $c_n(j) = \frac{\left(p_n^{(t)}(j)\right)^{-\sigma}}{\int_0^1 \left(p_n^{(t)}(x)\right)^{1-\sigma} dx} Y_n^{(t)}$. In equilibrium, factor markets clear. Because of lack of data, I assume trade balance. Finding the factor returns that balance trade is equivalent to solving for the factor-clearing conditions. In equilibrium, county n's income equals the sum of all 3109 counties' expenditures in goods from n:

$$w_n^{(t)} L_n^{(t)} = \sum_{i=1}^{3109} \frac{T_n^{(t)} \left(w_n^{(t)} \delta_{in}^{(t)}\right)^{-\theta} w_i^{(t)} L_i^{(t)}}{\sum_{m=1}^{3109} T_m^{(t)} \left(w_m^{(t)} \delta_{im}^{(t)}\right)^{-\theta}}$$

Even though income is an endogenous variable, I can condition on it and obtain the equilibrium by finding the source effects $S_n^{(t)} = T_n^{(t)} \left(w_n^{(t)}\right)^{-\theta}$ that satisfy the system of equations, given income $Y_n^{(t)} = w_n^{(t)} L_n^{(t)}$:[21][22]

$$Y_n^{(t)} = \sum_{i=1}^{3109} \frac{S_n^{(t)} \left(\delta_{in}^{(t)}\right)^{-\theta} Y_i^{(t)}}{\sum_{m=1}^{3109} S_m^{(t)} \left(\delta_{im}^{(t)}\right)^{-\theta}} \tag{4.1}$$

In the next subsection, I will explain how to obtain the source effects $S_n^{(t)}$ that clear factor markets from 19th century U.S. census data without trade data but with income and geographic data on railroads, rivers, oceans, and canals.

4.2 Calibrating the parameters of the model

I model iceberg costs between two counties to be linear in the length of time the freight took to travel between the two counties, and to be relative to the unit price of the good to be transported. The functional form of iceberg costs is

$$\delta_{ni}^{(t)} = 1 + \lambda d_{ni}^{(t)}$$

[21]The term "source effect" comes from Eaton and Kortum (2002)

[22]See Appendix B for details on these calculations

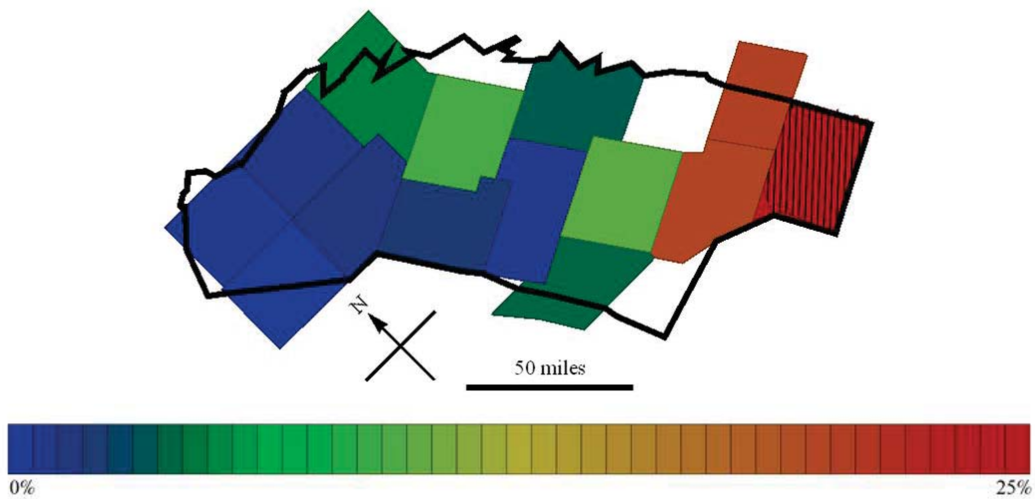
where λ is a constant representing the percentage costs of an additional day of transportation per unit of any good compared to its source price, and d_{ni} is the time in days for the fastest route from county i to county n. The parameter λ was calibrated to be equal to 0.05 from archival data on freight costs from the Midwest to the East Coast taken from Depew (1968), using my own calculations on travel times from the results of Subsection 3.3, and using spot prices from the "Wholesale Prices of Selected Commodities: 1784-1998" table in Carter, Gartner, Haines, Olmstead, Sutch, and Wright (2006).

Given the functional form of $\delta_{in}^{(t)}$, and because $\lambda = 0.05$, the remaining parameters to be calibrated are $S_n^{(t)}$ for all t and n, as well as θ. From census data (corresponding to t={1840, 1850, 1860, 1870, 1880, 1890, 1900}), I have the value of agricultural and manufacture output (I call this value GNP) of every county in the United States. These values are obtained at the county level. The data are presented in current dollars, and were calculated using local prices, except in the few cases where regional prices were used instead (see Carter, Gartner, Haines, Olmstead, Sutch, and Wright (2006) for the few exceptions).

To let present counties and old counties match geographically, I assume counties that had no data collected in the census had no factors of production. For example, take Milam county in Texas over time, pictured in Figure 4.1. In 1860, a total of 14 counties had more than 50% of their territory in what used to constitute Milam in 1850. In 1860, Milam produced 24% of the total value of output of those 14 counties, and its two immediate neighbors another 20% each. It is likely, then, that in 1850 most of the value of output of Milam county was produced in a short vicinity of what constituted Milam in 1860. In 2010, the territory of Milam almost coincides with the territory of Milam in 1860.

This story is repeated throughout the whole territory of the United States over time. The way data is simplified in this paper is as follows: a point located in Milam's 2010 population centroid produced all the output of Milam 1850, the population centroids of 2010 of the 14 counties that used to be Milam produced the output of the counties in 1860 and so on.

Figure 4.1: Milam county over time



Notes: County territory changed over time, but county names tended to preserve the location of the output value. Black polygon: Territory of Milam county in 1850. Colored polygons: territories of 14 counties in 1860, including Milam in red. Black stripes: Territory of Milam county in 2010. Colors of the 14 counties represent the share of their value of output in 1860.

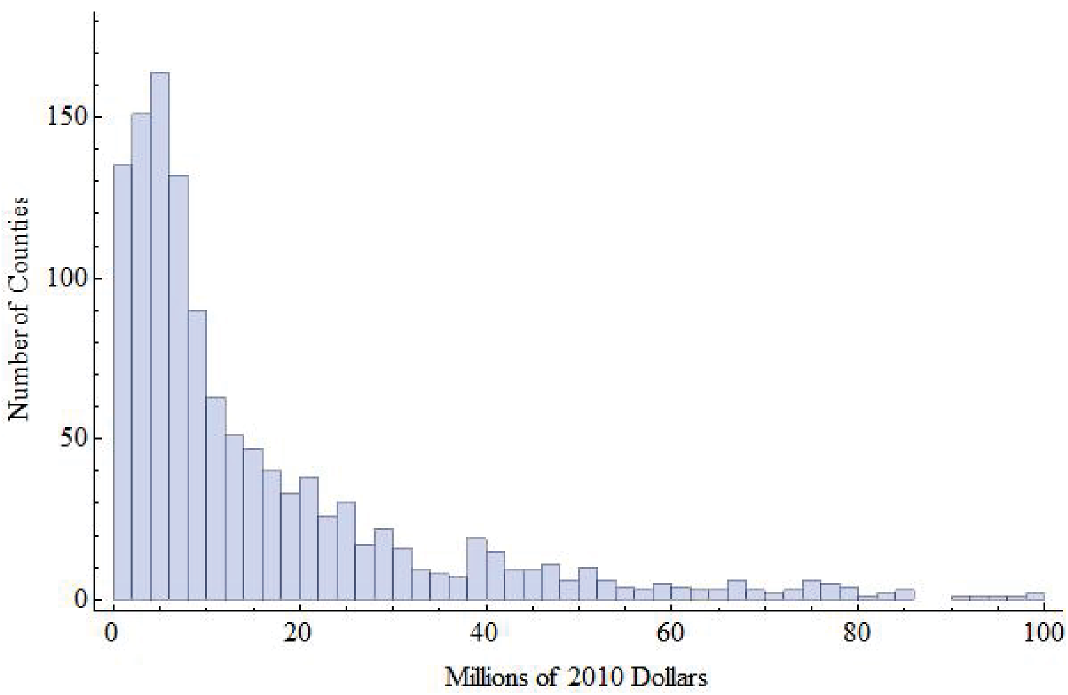
In the census of 1840, I have the value of GNP in 1840 for 1257 U.S. counties, which means I am assigning a value of $Y_n^{(t)} = 0$ to the 1852 counties with no reported income. In 2010 U.S. dollars, among the counties with a positive value of output, the mean county GNP was \$20.11 million and only 30 counties had GNP over \$100 million (Figure 4.2), with the 5 largest accounting for more than 7% of the GNP of the United States by 1840 (Table 1). In the next subsection, I will suppress the time (t) superscript notation, because all the calculations will be static.

Table 1: The five counties with the largest GNP in 1840

County	GNP in 1840
Philadelphia, PA	$517.77
Middlesex, MA	$397.08
New York, NY	$369.72
Worcester, MA	$243.31
Providence, RI	$217.96

Notes: Millions of 2010 U.S. dollars. The 5 counties with the largest GNP add to 7.05% of the GNP of the United States. There are 1257 counties with available data.

Figure 4.2: Histogram of U.S. county GNP in 1840



Notes: Quantities are in millions of 2010 U.S. Dollars. Data are truncated to $100 million for illustrative purposes.

4.3 Identifying Trade Shares

I will obtain the GNP in every county from the census and define it to be Y_n. This value will be taken as given. We have that Equation 4.1 can be expressed in matrix notation as

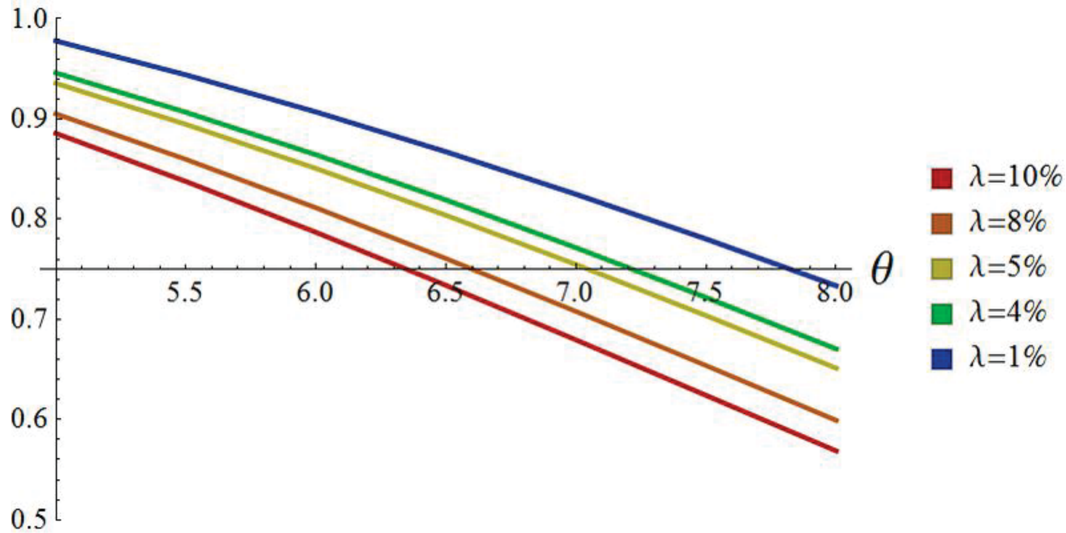
$$\mathbf{Y} = \Pi'\mathbf{Y} \tag{4.2}$$

with $\mathbf{Y} = (Y_1, Y_2, ..., Y_{3109})'$ and Π is a 3109-by-3109 matrix whose position (i, n) is $\pi_{in} = \frac{S_n(\delta_{in})^{-\theta}}{\sum_{m=1}^{3109} S_m(\delta_{im})^{-\theta}}$. The number π_{in}, is the fraction of county i's income spent on goods from county n, also known as county n's trade share on i. We have that Equation 4.2 is a system of 3109 equations and 3109 unknowns: the components of the vector $\mathbf{S} = (S_1, S_2, ..., S_{3109})'$. The rows of Π add to 1 for any value of $\mathbf{S}$. This last fact implies that one equation is redundant: if some vector $\mathbf{S}$ satisfies Equation 4.2, then $\epsilon\mathbf{S}$ with $\epsilon \neq 0$ also satisfies it, because Π has one eigenvalue equal to 1. Normalizing the vector $\mathbf{S}$ to add up to 1 gives a unique solution.

The fact that the system of equations (4.2) has a solution does not mean it is easily attainable numerically. In fact, any numerical method that directly wants to find the solution of the system will tend to be unstable from any initial point because the gradient is almost flat: the denominator in each of the 3109 summands of the 3109 equations is itself a sum of 3109 elements, so, given any initial point, the direct search will have increments that are two orders of magnitude larger than the elements of $\mathbf{S}$ and will tend to get negative values for some elements of $\mathbf{S}$.

Because of this issue, I use an iterative algorithm that is both fast and numerically stable. It is an adaptation of the Alvarez and Lucas (2007) algorithm. To the best of my knowledge, it has never been used in this context, neither in Eaton and Kortum (2002) setting nor in a general gravity setting. The elements π_{in} of the matrix Π can be decomposed into source effects and a function of trade costs, a functional form that appears in a wide range of trade models (such as Armington settings). This decomposition will be key in identifying changes in the equilibrium outcome when only trade costs vary. As explained in Subsection 4.5, the simplicity

Figure 4.3: Value of interstate trade as share of GNP in 1890 as a function of θ and λ



Notes: In 1890, Depew (1968) found this share to be 75%. The value $\theta = 7$ is assumed to be constant for the whole period 1840-1900

of the algorithm and its reach are encouraging for future research in this area, as long as trade costs are assumed to be observed.

4.4 The Iterative Algorithm

The values of the vector **Y** are obtained from the census. The values of the 3109-by-3109 matrix of optimal distances **d**, are obtained from the railroad and geographic and topographic data as described in Subsection 3.3. The value of $\lambda = 0.05$ is calibrated from the Commodity Prices of the Statistics of the United States, as explained before. The rounded up value of the $\theta = 7$ parameter is calibrated with interstate banking data from Depew (1968) and it almost fits the volume of trade of the census of 1890, which was 75% of GNP, as seen in Figure 4.3. Note the value $\theta = 7$ is also obtained in the Eaton and Kortum (2002) literature, where bilateral trade is usually available, but not the trade costs. I will now explain how to obtain the trade shares with an iterative algorithm.

Initialize the iterator $k = 0$ and start with a vector $\mathbf{S}^k$ such that $\mathbf{1}'\mathbf{S}^k = 1$ and $(S_i)^k > 0$ for all i with $Y_i > 0$. A good guess is $\mathbf{S}^k = \frac{\mathbf{Y}}{\mathbf{1}'\mathbf{Y}}$. This vector will uniquely characterize the matrix Π^k.

Then I follow the iterative process until convergence:

$$\mathbf{S}^{k+1} = \mathcal{H}\left(\mathbf{T}^k\right) = \mathbf{S}^k + \eta\left[\mathbf{Y} - \left(\Pi^k\right)'\mathbf{Y}\right] \tag{4.3}$$

with $\eta > 0$ and sufficiently small so as not to have any entries of $\mathbf{S}^{k+1}$ negative for any k. The term $\mathbf{Y} - \left(\Pi^k\right)'\mathbf{Y}$ is a vector of "excess supply", that is, the difference between the observed county income and the total expenditure of other counties in products sourced from some county that is predicted by the vector of source effects. If county n spends more dollars Y_n than what the current vector of source effects $\left(\Pi_n^k\right)'\mathbf{Y}$ predicts, then it must be the case that county n has a larger weight as a source of output. Note the function $\mathcal{H}$ takes the space of vectors of length 3109 that add up to 1, to itself.[23] The fixed point of this iterative process satisfies the balanced trade equilibrium condition and implies that for those source effects, the factor market clears. Numerically, the benefits of railroad construction policy, given county incomes, are independent of the selection of factor of production, even if the factor changes constantly over time. Counter-factual analysis, as the one in Dekle, Eaton, and Kortum (2007)

[23]Proof:

$$\begin{aligned}
\mathbf{1}'\mathcal{H}\left(\mathbf{T}^k\right) &= \mathbf{1}'\left(\mathbf{T}^k + \eta\left[\mathbf{Y} - \left(\Pi^k\right)'\mathbf{Y}\right]\right) \\
&= 1 + \eta\mathbf{1}'\left(\mathbf{Y} - \left(\Pi^k\right)'\mathbf{Y}\right) \\
&= 1 + \eta\left(\sum_{n=1}^{3109} w_n L_n - \sum_{n=1}^{3109}\sum_{i=1}^{3109} \frac{T_n\left(w_n\delta_{in}\right)^{-\theta}}{\sum_{m=1}^{3109} T_m\left(w_m\delta_{im}\right)^{-\theta}} w_i L_i\right) \\
&= 1 + \eta\left(\sum_{n=1}^{3109} w_n L_n - \sum_{i=1}^{3109} w_i L_i \sum_{n=1}^{3109} \frac{T_n\left(w_n\delta_{in}\right)^{-\theta}}{\sum_{m=1}^{3109} T_m\left(w_m\delta_{im}\right)^{-\theta}}\right) \\
&= 1 + \eta\left(\sum_{n=1}^{3109} w_n L_n - \sum_{i=1}^{3109} w_i L_i\right) = 1
\end{aligned}$$

follows trivially once the matrix Π is given and transportation costs do not affect productivity parameters.

4.5 Alternate implementations of the iterative algorithm

This iterative algorithm is not specific to this model. If trade costs are assumed to be observed, then this "excess supply" iterative algorithm to find trade balances without loss of generality. Trade shares are functions of factor rents and productivity in an Eaton and Kortum (2002) setting, but similar equilibrium conditions arise in other model specifications or when other types of data are available for the researcher.

4.5.1 When the only factor of production is observed

For this paper, if labor is assumed as the only factor of production, and if labor in every county is observed, then the factor rent is the GNP per worker, and I can define $S_n = T_n \left(\frac{Y_n}{L_n}\right)^{-\theta}$ as the vector of source effects, and still start with $\mathbf{S}^k$ such that $\mathbf{1}'\mathbf{S}^k = 1$ for $k = 0$. Then the process in Equation 4.3 still converges, and finds normalized values of productivity T_n. Productivity is a sufficient statistic for welfare under autarky, and getting these parameters allows to observe the dispersion of ideas and technology over time, as well as the spatial correlation of productivity shocks on census years, such as droughts or financial crises.

4.5.2 Armington Preferences

Another example in which this setting can be useful is with Armington preferences.[24] For example, an endowment economy-Armington model with N regions and balanced trade with

[24] See Arkolakis, Costinot, and Rodriguez-Clare (2009) for an extensive discussion on trade share equation equivalences among several trade models.

preferences defined by

$$U_n = \left(\sum_{i=1}^{N} \alpha_i^{\frac{1}{\sigma}} c_{ni}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

where σ is the elasticity of substitution and α_i is the preference parameter for goods from region i (common across all regions), the market clearing condition is given by

$$Y_n = \sum_{i=1}^{N} \frac{\alpha_n \left(p_n \delta_{in}\right)^{1-\sigma} Y_i}{\sum_{m=1}^{3109} \alpha_m \left(p_m \delta_{im}\right)^{1-\sigma}}$$

where p_n is the price per unit of endowment good in region n. In that case, the source effect is just $S_n = \alpha_n \left(p_n\right)^{1-\sigma}$ and the iterative equation is

$$Y_n = \sum_{i=1}^{N} \frac{S_n \left(\delta_{in}\right)^{1-\sigma} Y_i}{\sum_{m=1}^{3109} S_m \left(\delta_{im}\right)^{1-\sigma}}$$

Again, the intuition behind the reasons for convergence is similar as the productivity. Fixing the price per unit of endowment goods, an excess supply in a region implies that overall the preference for the goods of that region must be larger than the current vector of preferences predicts.

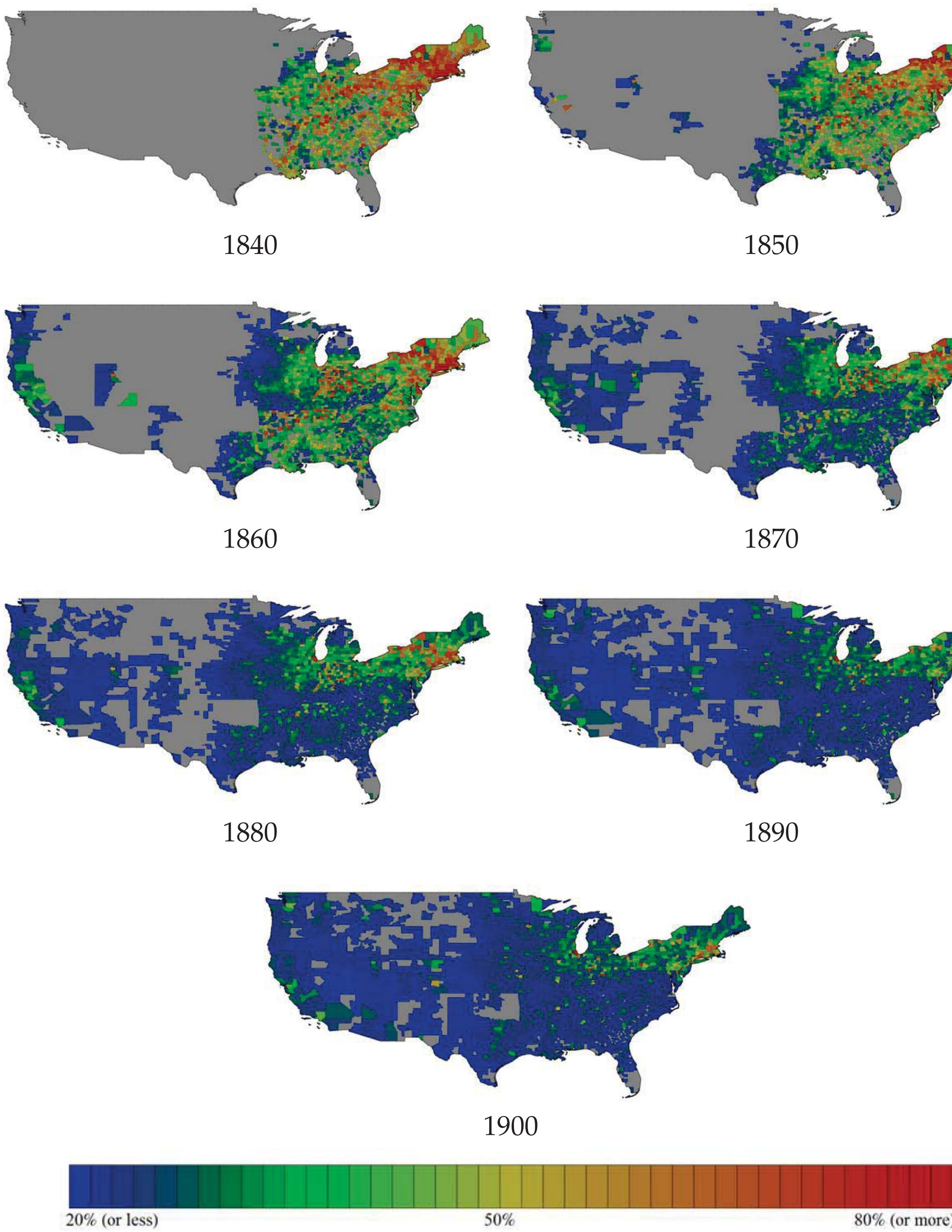
4.6 Using the results of the iterative algorithm

An important contribution of the paper is that it deals with a high dimensional income data set in which trade data do not exist, but are inferred from a model, and the family of models from which it can be inferred is large. The algorithm finds trade shares π_{in} and so finds gains from trade for a general setting of trade models. From Eaton and Kortum (2002) we know that "own shares" define for gains from trade with respect to autarky, and from Donaldson (2008) we know this is also a sufficient statistic for gains from trade in an intertemporal setting. The vector of own shares is π_{nn}, for all $n = \{1, 2, ..., 3109\}$. In Figure 4.4, results of the calculated own-trade shares are pictured for all the existing counties during all U.S. census years between 1840 and 1900, given the railroad construction up to the time of the census. Low values imply large gains from trade. Gains from trade increased through time and spread along the whole U.S. territory. However, the impact of railroads on these gains is not seen here. It can only be seen doing a counter-factual exercise. During the 1840s, trade was expensive and the rich counties of the Northeast could not benefit from trade, because at the time, the regions connected with railroads were all high-income regions too, unable to exploit low wages in other regions. The number of counties with which to trade for a low cost, for almost any county, was small. Note also the Louisiana region did not benefit much as the railroad network kept expanding. They stopped being part of the fastest route from the South to the Northeast, thus stagnating their benefits from the trains in the area. This might be an overstatement of my model, mainly because I assume symmetric travel times over the Mississippi River, which was not the case. Nevertheless, as seen in Figure 2.1 and Figure 2.2, the GNP of the counties in that region decreased substantially as the railroads to the West kept expanding. Being located next to a river gave two advantages to rural counties during the first half of the 19th century: it made land (and labor) productive, and kept trade relatively cheap. As the freight costs were reduced, the advantage of low cost trade vanished, and the relative importance of high productivity vanished compared to the geographic advantage of being part of the optimal transportation routes.

5 Counter-Factuals

I have identified the source effects and the gains from trade due to railroads of every county in the United States for every census between 1840 and 1900 by finding the fixed point in a simple

Figure 4.4: U.S. counties' estimated own share of sales from 1840 to 1900



Notes: Low values indicate large gains from trade. Gray: counties with no data iterative algorithm. Let x' be the equilibrium counter-factual value of some variable x after changes in the railroad network. We know that changes in the railroad network will change the whole matrix if δ_{ni} into δ'_{in} but do not change factor endowments or productivity, so the new factor returns will satisfy:

$$w'_n L_n = \sum_{i=1}^{3109} \frac{T_n \left(w'_n \delta'_{in}\right)^{-\theta}}{\sum_{m=1}^{3109} T_m \left(w'_m \delta'_{im}\right)^{-\theta}} w'_i L_i$$

But production factors are not observable, so now I make use of the Dekle, Eaton, and Kortum (2007) setup to calculate the instantaneous effect of changing the railroad, assuming they do not change county productivity. This is a similar exercise as Caliendo and Parro (2009), who calculate changes in real wages in a multi-sector framework. Define counter-factual proportional change in variable x to be $\widehat{x} = x'/x$. I am assuming that factors are fixed, so I get that $w'_n L_n = \widehat{w}_n w_n L_n = \widehat{w}_n Y_n$. The counter-factual factor returns satisfy

$$\widehat{w}_n Y_n = \sum_{i=1}^{3109} \frac{\pi_{ni} \left(\widehat{w}_n \widehat{\delta}_{in}\right)^{-\theta}}{\sum_{m=1}^{3109} \pi_{nm} \left(\widehat{w}_m \widehat{\delta}_{im}\right)^{-\theta}} \widehat{w}_i Y_i \tag{5.1}$$

whose solution can be easily found with the Alvarez and Lucas (2007) algorithm, up to a normalization. Such normalization is $\sum_{n=1}^{3109} Y_n = \sum_{n=1}^{3109} \widehat{w}_n Y_n$ and is a useful one because the railroad construction cost data is available for every decade, so this cost will be measured, instead of in dollars, as a fraction of the nominal GNP of the United States of the baseline case. All the gains from the railroads will be measured by the percentage real gains in output, which do not depend on the normalization. Also, this process does not depend on the observation of the productivity variables or on the factor of production, so it can be generalized. Note that the price indices in every county will satisfy the equality:

$$\widehat{P}_n = \left(\sum_{i=1}^{3109} \pi_{ni} \left(\widehat{w}_i \widehat{\delta}_{ni} \right)^{-\theta} \right)^{-\frac{1}{\theta}}$$

On the other hand, the output and counter-factual output in county n is given respectively by

$$\begin{aligned} U_n &= \frac{w_n L_n}{P_n} = \frac{Y_n}{P_n} = \Gamma^{\frac{1}{\sigma-1}} \left(\frac{\theta - (\sigma - 1)}{\theta} \right) \left(\sum_{m=1}^{3109} S_m \left(\delta_{nm} \right)^{-\theta} \right)^{\frac{1}{\theta}} Y_n \\ U'_n &= \frac{w'_n L_n}{P'_n} = \frac{\widehat{w}_n}{\widehat{P}_n} \frac{Y_n}{P_n} = \frac{\widehat{w}_n}{\widehat{P}_n} U_n \end{aligned}$$

so, the percentage change in total output in the United States will be given by

$$\frac{\sum_{n=1}^{3109} U'_n}{\sum_{i=1}^{3109} U_i} = \frac{\sum_{n=1}^{3109} U_n \frac{\widehat{w}_n}{\widehat{P}_n}}{\sum_{i=1}^{3109} U_i} = \sum_{n=1}^{3109} \alpha_n \frac{\widehat{w}_n}{\widehat{P}_n}$$

where $\alpha_n = \frac{U_n}{\sum_{i=1}^{3109} U_i}$ is the share of total output of the United States coming from county n. Note that this percentage change is independent of the normalization of the source effects, and that it is not necessary to identify the level of factor endowments per county. In other words, nominal GNP and transportation costs are all that is needed to calculate changes in output due to changes in the railroad network.

In the next subsections I do counter-factuals. In Subsection 5.1, I estimate the impact of recent railroad construction on output by finding the counter-factual output that would have been produced if instead of using the current railroads, the railroad network built up to another year in the past was used instead. In Subsection 5.2 I estimate the upper bound of the impact of future railroads by finding the change in output if the railroads that were constructed in the next decade were made suddenly available at no cost. In Subsection 5.3, I estimate the impact of railroad construction of any year on every census.

Table 2: Output losses for removing the railroads built during the past decade

Census Year	Miles opened in past decade	Percentage Losses
1840	3,013*	0.21%*
1850	4,708	0.35%
1860	16,239	1.43%
1870	14,410	1.91%
1880	28,119	2.04%
1890	48,849	1.82%
1900	16,616	0.13%

Notes: Losses are measured as percentage of GNP. Railroad length is calculated from the geographic data of the railroad openings data set. (*) Output and railroad construction are compared to not having any railroads.

5.1 The impact of recent railroad construction on output

In this subsection, I estimate the impact of recent railroad construction by calculating changes in output if the railroads opened in the past are made suddenly unavailable. I do two counterfactuals. First, I calculate the output reduction if railroads built up to 10 years before every census are removed. The results are in Table 2.

It is interesting to see that the construction of the railroads to the Pacific had a similar effect on output to the railroads constructed in the next decade, which mainly interconnected routes, and a smaller effect than the construction of the other two railroads to the Pacific, during the decade of 1880. As a rough estimate of efficiency, during these 60 years, an average of 3% of GNP was invested on railroads, which would have made most of these decades of constructions non efficient on net output if productivity and factors remained constant for a decade. Note that the increase in construction of the railroads coincides with large gains from the transportation cost reductions associated with them, and that when these gains ended, so did railroad construction.

Table 3: Output losses for removing all the railroads

Census Year	Miles opened	Percentage Losses
1840	3,013	0.21%
1850	7,721	0.62%
1860	23,960	2.34%
1870	38,370	4.76%
1880	66,489	7.67%
1890	115,338	9.61%
1900	131,954	8.97%

Notes: Losses are measured as percentage of GNP. Railroad length is calculated from the geographic data of the railroad openings data set.

Table 4: Estimates for losses for removing the railroads in 1890

Model	Losses
Fogel (1962)	2.7%
Donaldson and Hornbeck (2012)	6.3%
Pérez (2013)	9.61%

Notes: Losses are measured as percentage of 1890's GNP.

In a second counter-factual exercise, I estimate the impact of railroads on U.S. output by obtaining the output reduction if all railroads are removed. The results are in Table 3, where the row corresponding to the census of 1890 is a similar exercise to the classic Fogel (1962) calculation of the benefits of railroad construction (an estimate also done recently by Donaldson and Hornbeck (2012)). Interesting things to be noted include that the railroads were more important for output in 1890 than in 1900, and that removing railroads in any of the two decades of construction of the three railroads to the Pacific (1860 and 1800) has a huge impact on output, which means that there could have been some migration to the west in preparation to the construction, and that this migration could have taken more than a decade to happen. Comparing the output losses of 1890 with similar exercises in the literature in Table 4, this analysis gets a larger estimate for the impact of the railroads. The fact that it is much larger than previous work has to do more with the nature of the nodal location of factors of production in the model, an issue also found in Donaldson (2008) for the case India, where he finds that railroads increased output by 16%. In particular, Donaldson and Hornbeck (2012) also have a Ricardian model of trade, and assume population to be spread throughout the counties, so the benefits of connecting counties to the railroad network are not uniform throughout the territory, which explain why they get a smaller number than this analysis gets.

The Dekle, Eaton, and Kortum (2007) setup allows to measure how much of the change in output can be attributed exclusively to faster transportation routes using railroads (see Figure 5.1). In line with the trade literature, reducing transportation costs imply less gains to rich regions, which are more self sufficient. In Figure 4.4 it is possible to see changes in gains from trade over time, and Figure 5.2 shows that the gains from trade due to railroads are much smaller in regions with larger GNP.

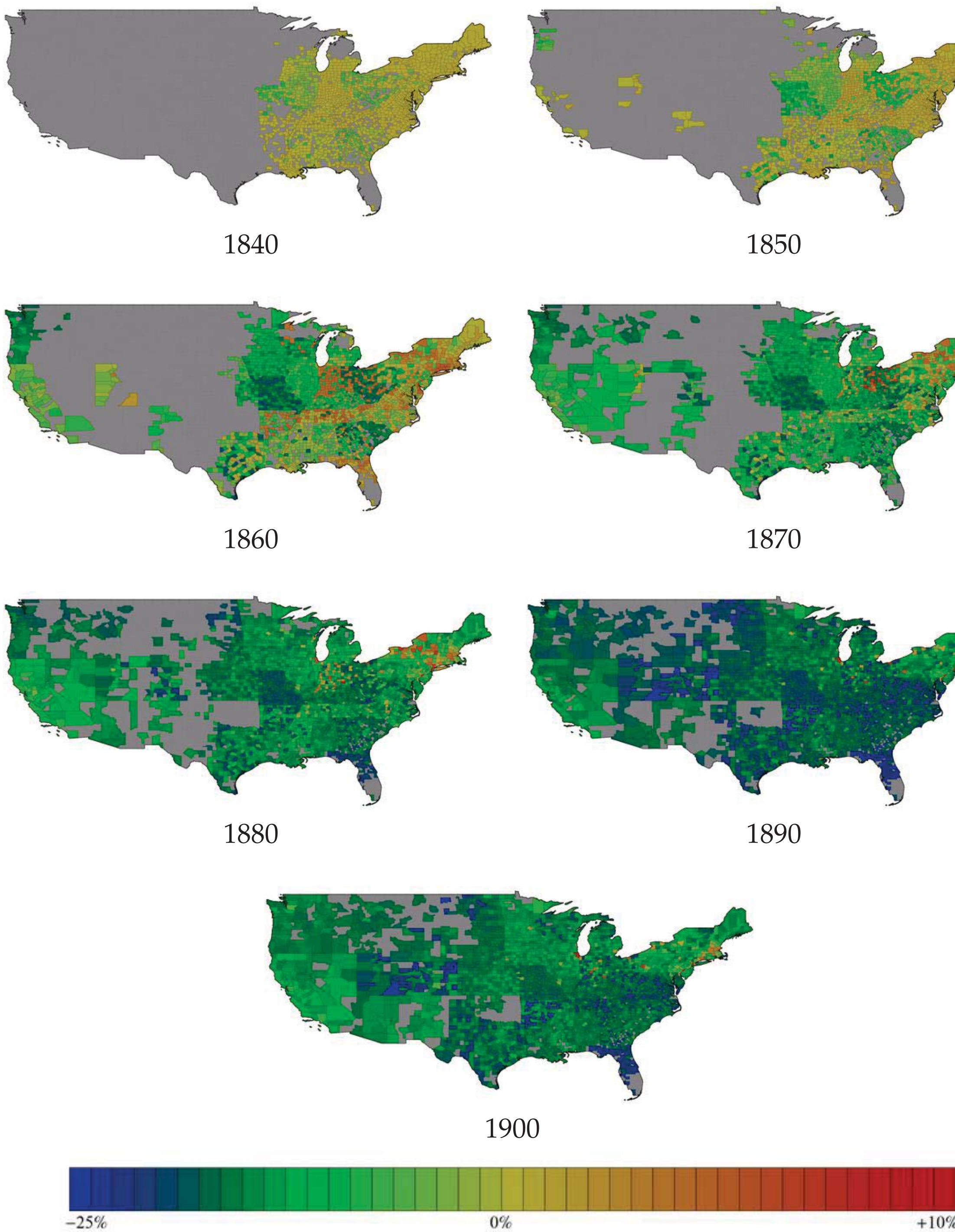
5.2 Calculate the impact of future railroad construction on output

In this subsection, I calculate changes in output if the railroads opened during the next 10 years are immediately made available for no cost, to approximate the upper bound of the impact of railroad construction on current factors of production. The results are in Table 5. The impact of adding the railroads of the decade is smaller than the impact of removing the railroads on the next decade for the six census years studied, and output did not increase by more than 1.2% in any of the cases, which means that the sources of the observed growth were not exclusively transportation costs.

5.3 Change in output for every census year when any railroads are available

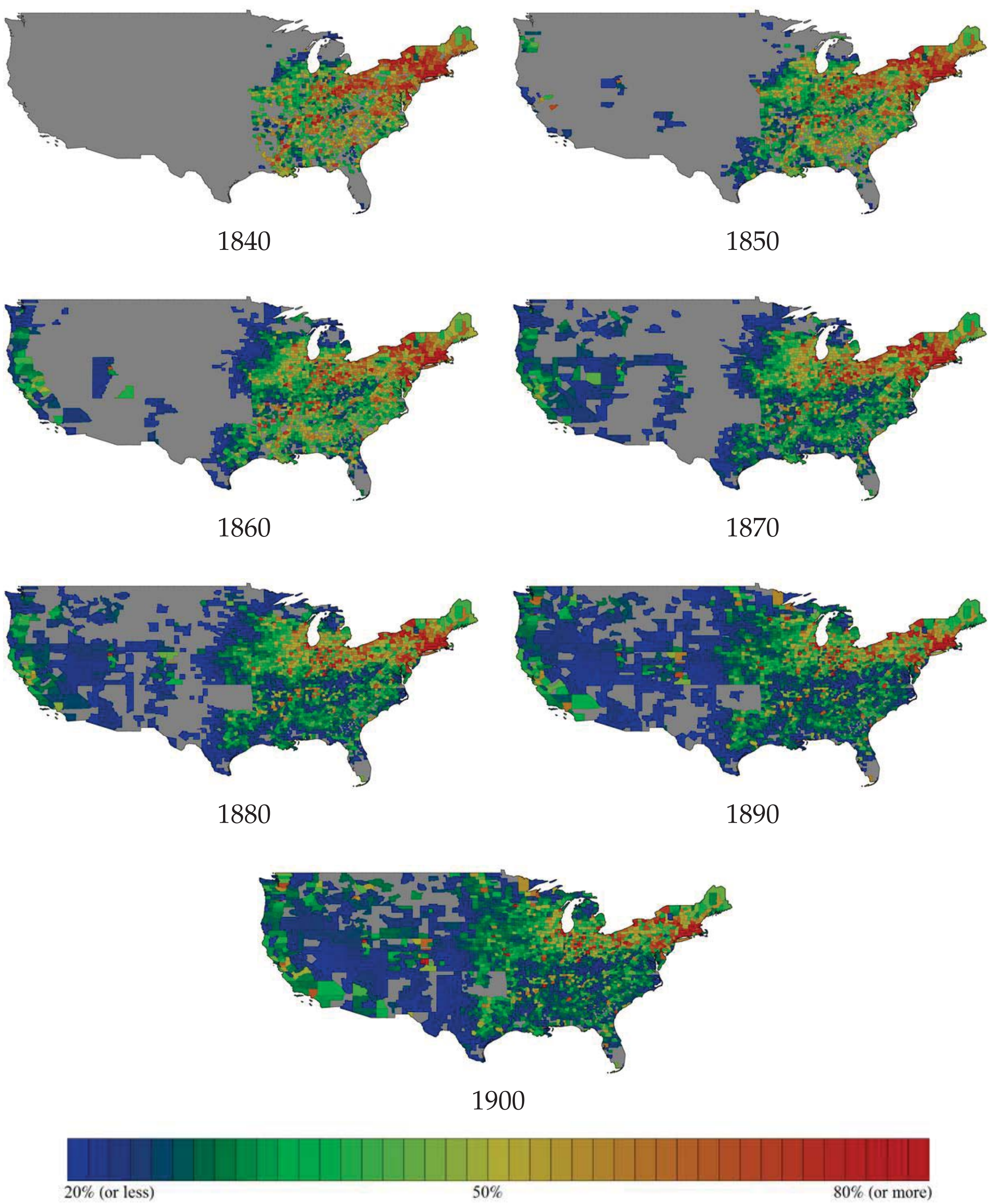
Finally, I want to calculate the change in output if every census year factors of production use the means of transportation available up to any other year, past or future. That is a total of

Figure 5.1: U.S. counties' change in output if railroads were removed, census years 1840-1900



Notes: Percentage change in output if railroads were made suddenly unavailable. Gray: counties with no data. Notice how the crisis during the decade of 1890 reduced the importance of railroads throughout the territory.

Figure 5.2: U.S. counties' estimated own share of sales from 1840 to 1900 if railroads were removed.



Notes: Low values indicate large gains from trade. Gray: counties with no data

Table 5: Output gains for adding the railroads built during the next decade for no cost

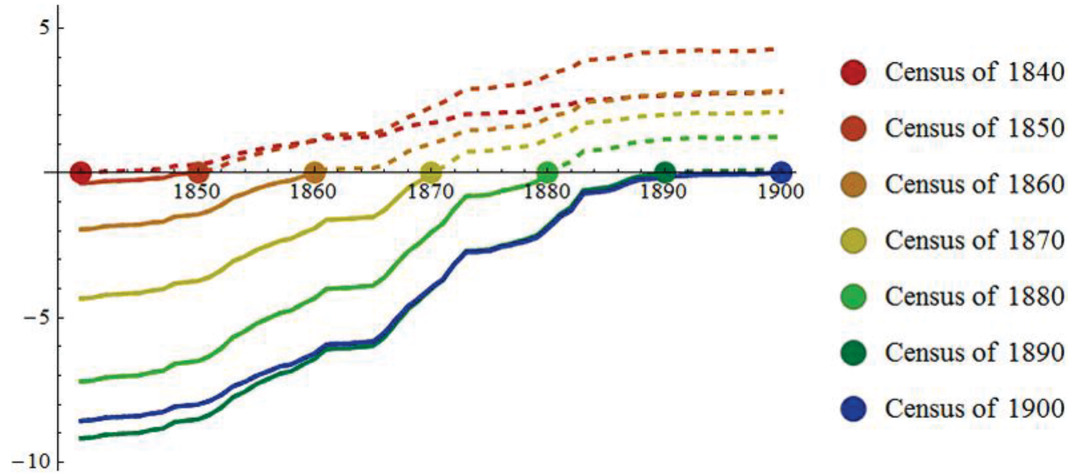
Census Year	Miles opened in the next decade	Percentage Gains
1840	4,708	0.28%
1850	16,239	1.12%
1860	14,410	1.01%
1870	28,119	1.17%
1880	48,849	1.16%
1890	16,616	0.12%

Notes: Losses are measured as percentage of GNP. Railroad length is calculated from the geographic data of the railroad openings data set.

420 counter-factual equilibrium calculations for the 3109 county factor returns and 9.6 million pairs of trade costs. The results of using the railroads in 1840-1900 are in Figure 5.3, with past railroads in solid lines, and future railroads in dotted lines.

There are many things to be noted from Figure 5.3, but I want to highlight two. The first one is that in 1850, once California became part of the territory of the United States, the construction of the railroads to the Pacific, together with all the major network improvements in the Midwest, would have created huge gains from connecting East and West. The state of California had a total of just about 92,000 residents according to the 1850 census (only 0.4% of the 23.2 million people in the United States), but its counties were so highly productive, that reducing trade costs would have had a huge impact on the overall welfare of the East. The project, however, would not have been profitable. If building railroads to the West had cost the same per mile as in the East, just the construction a railroad from Chicago to San Francisco would have cost around 4% of the GNP of 1850. All the other construction in the Midwest during those 20 years adds at least eight times as much to that cost. Also, by the time of the next census, the population had spread so much (without any real connection between the Atlantic and the Pacific) and the productivity of the land had decreased in such a way that the same trade cost

Figure 5.3: Percentage change in output with railroads built up to any year using census-year data



Notes: Factors of production are fixed, and output is calculated using trade shares of present available means of transportation at every census year. The dotted lines represent railroads built in the future.

reduction would have had only a fraction of the impact. This analysis concludes that a railroad to the Pacific would have been a worse investment in the census year of 1860 than in 1850. This conclusion, which uses a lower bound for marginal costs (because it uses the costs of construction only from observed railroads), contrasts with White (2011) and with Fogel et al. (1960), who find that the railroad to the Pacific was constructed at least a couple decades earlier than needed.

The second thing to be noted is that future railroads create much less growth than the very most recently opened railroads. Comparing the census of 1890 and 1900, removing the railroads in 1890 would have created a larger loss than removing the railroads in 1900. That decade, major events happened in the railroad industry. First, the economic and financial crisis known as "The Panic of 1893" stopped the construction of any major new routes. When railroad revenue went down, events like the "Pullman Strike", which happened when wages were lowered, caused railroad workers to collapse the transportation routes west of Michigan. The reliability of the railroad became an issue, and people migrated back to the regions that were less dependent on railroads. During this decade, the automobile was invented, maybe as an answer to the struggle that the transportation industry was starting to create.

6 Conclusion

I have developed a 3 step procedure to find the impact of the railroads on output. Each of the 3 steps is a contribution to the existing literature, and its results can be extended to various areas of research, and generalized to a large family of trade models. The first step –the construction of the dataset– contains precise geographic information of almost 13,000 railroad tracks, the historical value of output for counties in the United States, and information on possible transportation routes to connect counties via wagons, boats, or trains. All this information was simplified at the county level, but it can still be used in its raw form, or simplified into other geographical levels.

The second step, an iterative method to find non-existing trade data using only the nominal GNP and transportation costs, allows for the calculation of counter-factual percentage changes in equilibrium variables, even if the factors of production are not identified and trade data is not available. If factors of production are defined and their endowments are available, then also the original levels can be measured. This is an important steps for studying inter-regional trade, where usually bilateral trade data is not available. The most important requirement in this model is trade costs, which are usually known at the regional level, so this process also allows to increase the amount of trading regions in trade models by an order of magnitude without increasing the computational burden of the problem. Usually, the number of trading regions in models of trade is low because of lack of data, so the iterative method described on this paper allows to get rid of this obstacle.

Finally, I make use of this model and construct counter-factuals, using the actual timeline of railroad construction. By the end of the 19th century, I find that railroads increased output by 9%. Comparing the results to what others have done in the past, the flexibility of the model allows to improve on the estimate of the gains from trade of railroads to be at least 9.6%. Finally, I estimate that the impact of the railroads, letting factors of production fixed, has a stronger impact going backwards in time than forward.

References

ALVAREZ, F., AND R. E. LUCAS, JR. (2007): "General Equilibrium analysis of the Eaton-Kortum Model of International Trade," *Journal of Monetary Economics*, 54(6), 1726–1768.

ARKOLAKIS, C., A. COSTINOT, AND A. RODRIGUEZ-CLARE (2009): "New trade models, same old gains?," Discussion paper, National Bureau of Economic Research.

CALIENDO, L., AND F. PARRO (2009): "Estimates of the Trade and Welfare Effects of NAFTA," .

CARTER, S., S. GARTNER, M. HAINES, A. OLMSTEAD, R. SUTCH, AND G. WRIGHT (2006): *Historical Statistics of the United States: Millennial Edition*. Cambridge: Cambridge University Press.

DEKLE, R., J. EATON, AND S. KORTUM (2007): "Unbalanced trade," *The American economic review*, 97(2), 351–355.

DEPEW, C. (1968): *One Hundred Years of American Commerce 1795-1895*. Greenwood Press.

DONALDSON, D. (2008): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *WP LSE*.

DONALDSON, D., AND R. HORNBECK (2012): "Railroads and American Economic Growth: A 'Market Access' Approach," .

EATON, J., AND S. S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

ENGLISH, W. B. (1996): "Understanding the costs of sovereign default: American state debts in the 1840's," *The American Economic Review*, pp. 259–275.

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

FOGEL, R., ET AL. (1960): *The Union Pacific Railroad: A case in premature enterprise*, vol. 78. Johns Hopkins Press.

GLOMM, G., AND B. RAVIKUMAR (1994): "Public Investment in Infrastructure in a Simple Growth Model," *Journal of Economic Dynamics and Control*, 18(6), 1173–1187.

HAINES, M. R., ET AL. (2005): "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002," .

HERRENDORF, B., J. SCHMITZ, A. TEIXEIRA, AND F. R. B. OF MINNEAPOLIS. RESEARCH DEPT (2009): *Transportation and Development: Insights from the US, 1840-1860*. Federal Reserve Bank of Minneapolis, Research Department.

HOLMES, T., AND J. SCHMITZ JR (2001): "Competition at work: railroads vs. monopoly in the US shipping industry," *Federal Reserve Bank of Minneapolis Quarterly Review*, 25(2), 3–29.

MITCHELL, B. R. (1964): "The Coming of the Railway and United Kingdom Economic Growth," *The Journal of Economic History*, 24(3), 315–336.

MUNNELL, A. H. (1992): "Policy Watch: Infrastructure Investment and Economic Growth," *The Journal of Economic Perspectives*, 6(4), pp. 189–198.

PAULLIN, C., AND J. WRIGHT (1932): *Atlas of the historical geography of the United States*. Carnegie Institution of Washington.

STOVER, J. F. (1997): *American Railroads*. University of Chicago Press.

——— (1999): *The Routledge Historical Atlas of the American Railroads*. Psychology Press.

SUMMERHILL, W. R. (2003): *Order Against Progress: Government, Foreign Investment, and Railroads in Brazil, 1854-1913*. Stanford Univ Pr.

WALLIS, J., R. SYLLA, AND A. GRINATH (2004): "Sovereign Default and Repudiation: The Emerging Market Debt Crisis in the United States," *NBER Working Paper, w10753*.

WALLIS, J. J. (2004): "Constitutions, Corporations, and Corruption: American States and Constitutional Change," Discussion paper, National Bureau of Economic Research.

WHITE, R. (2011): *Railroaded: The Transcontinentals and the Making of Modern America*. WW Norton & Co Inc.

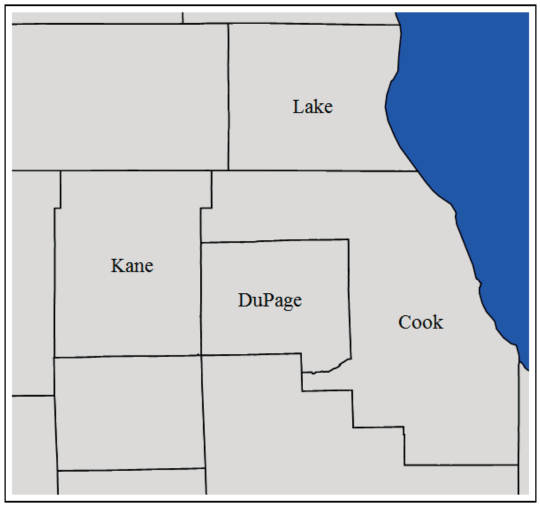
WINSTON, C. (1991): "Efficient Transportation Infrastructure Policy," *The Journal of Economic Perspectives*, 5(1), pp. 113–127.

A Computational Image Processing

A.1 Single Channel Computational Image Processing

The method I used to determine if two counties were neighbors and if a railroad, river, or lake goes through a county was Single Channel Computational Image Processing. This method is simple arithmetic of white light, applied to the images that geographical data create when drawn with the computer.

Figure A.1: How the counties actually look in a regular map



The basic theory is the following. Every gray level of color that is readable by a computer can be expressed as a number $x \in [0, 1]$, where x indicates the level of white light. A value of $x = 0$ is the black color, and $x = 1$ is pure white. Two colors can be added, subtracted and multiplied, with the only restriction being that if the result goes above 1, then it is truncated to 1, and if the result goes below 0, then it is truncated to 0.

In order to determine if two counties are neighbors, what I do is to draw both counties individually in dark gray color ($x = \frac{1}{4}$) with white background ($x = 1$), and add the images. If the result is a completely white image (verified by obtaining the minimum value of color in the drawing to be equal to 1), then the counties are not neighbors. If, on the other hand, it contains at least one pixel different from white (and in this particular case one whose color equals to $\frac{1}{2}$), then the two counties intersect. The procedure needs fairly detailed maps in order to be reliable, that is why the maps from the U.S. Census were used. Figures A.1,A.2, and A.3 show a visual example of how it was determined that two counties in Illinois (Cook and Du Page) are neighbors, while Kane and Lake, also in Illinois, are not neighbors. .

Then I do the same to find the neighboring counties or every railroad, river, lake, and canal in the United States, at every moment in time. The process is extremely efficient, and much

Figure A.2: Image Processing: Cook and Du Page are neighbors

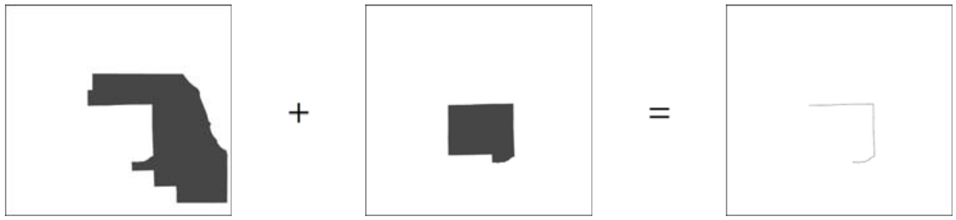
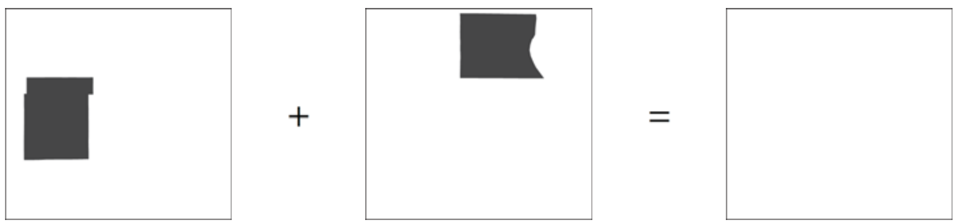


Figure A.3: Image Processing: Kane and Lake are not neighbors



faster than computational geometry methods, such as the Point-In-Polygon algorithm. Figures A.4 and A.5 show an example of a railroad that is added with the map of Cook county in Illinois, and since not all the pixels ad up to 1, it is determined that the railroad touches Cook County.

A.2 Multichannel Image Processing

The method I used to combine the U.S. topological maps and the railroad maps is multichannel computational image processing. The idea is analogous to the single channel image processing, but with three colors instead of one (red, green, and blue instead of white). Every color that is readable by a computer can be expressed as a vector $x = (x_1, x_2, x_3) \in [0,1]^3$, where x_1 indicates the level of red light, x_2 indicates the level of green light, and x_3 indicates the level of blue light. A value of $x = (0,0,0)$ is the black color, $x = (1,0,0)$ is red, $x = (0,1,1)$ is yellow, and $x = (1,1,1)$ is white, just to name a few examples. The same as with single channel image processing, two colors can be added, subtracted and multiplied element by element. If, for each

Figure A.4: A section of the Berkshire Hathaway line, built in 1873

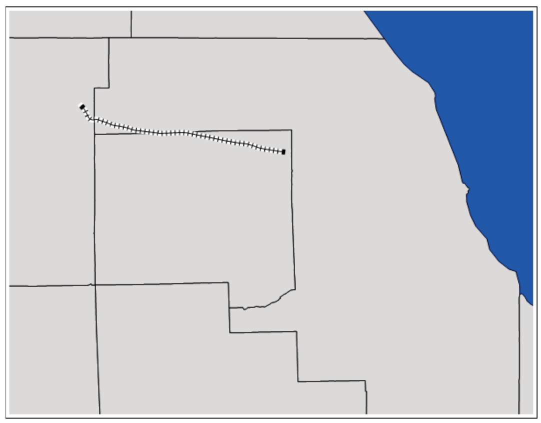
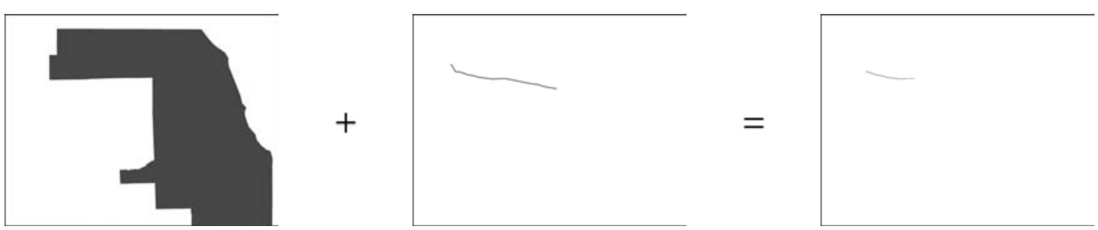


Figure A.5: Image Processing: That section of the Berkshire Hathaway line, built in 1873



Notes: This railroad goes through Cook County.

Figure A.6: Multichannel Image Processing: How the 1870 railroad map was drawn

element, the result goes above 1, then it is truncated to 1, and if the result goes below 0, then it is truncated to 0.

The maps shown in Figure 2.1 are the result of taking a topological map of the United States and erasing all the light that corresponds to the railroads. Then the red lines are added and everything that lies outside of the United States is deleted by adding white light, as seen in Figure A.6.

B Overview of Eaton and Kortum

In equilibrium, the balanced trade conditions are the usual: county n's income equals the sum of all the 3109 counties' expenditure in goods from n. We have that

$$
\begin{aligned}
w_n^{(t)} L_n^{(t)} &= \sum_{i=1}^{3109} \{\text{county } i\text{'s expenditure on goods from county } n\} \\
&= \sum_{i=1}^{3109} w_i^{(t)} L_i^{(t)} \{\text{share of county } i\text{'s expenditure on goods from county } n\} \\
&= \sum_{i=1}^{3109} w_i^{(t)} L_i^{(t)} \int_0^1 \mathbb{1} \{\text{county } n \text{ is the lowest cost provider of good } j \text{ in county } i\} \, dj \\
&= \sum_{i=1}^{3109} w_i^{(t)} L_i^{(t)} \mathbb{P} \left(\frac{w_n^{(t)} \delta_{in}^{(t)}}{Z_n^{(t)}(j)} \leq \min_{m \neq i} \left\{ \frac{w_m^{(t)} \delta_{im}^{(t)}}{Z_m^{(t)}(j)} \right\} \right) \\
&= \sum_{i=1}^{3109} w_i^{(t)} L_i^{(t)} \frac{T_n^{(t)} \left(w_n^{(t)} \delta_{in}^{(t)} \right)^{-\theta}}{\sum_{m=1}^{3109} T_m^{(t)} \left(w_m^{(t)} \delta_{im}^{(t)} \right)^{-\theta}}
\end{aligned}
$$

The second equality comes from the fact that the origin-specific price index component is the same for every destination county. The price paid in county n for goods satisfies the following

Price Index identity:

$$
\begin{aligned}
P_n^{1-\sigma} &= \int_0^1 p_n(j)^{1-\sigma}\, dj \\
&= \sum_{i=1}^{3109} \left(\int_0^1 p_n(j)^{1-\sigma}\, \mathbb{1}\{\, i \text{ is the lowest cost provider of good } j \text{ in } n\}\, dj \right) \\
&= \sum_{i=1}^{3109} \left(\int_0^1 \left(\frac{w_i \delta_{ni}}{Z_i(j)} \right)^{1-\sigma} \mathbb{1}\{ i \text{ is the lowest cost provider of good } j \text{ in } n\}\, dj \right) \\
&= \sum_{i=1}^{3109} \left(\mathbb{E}\left(\left(\frac{w_i \delta_{ni}}{Z_i(j)} \right)^{1-\sigma} \middle| \frac{w_i \delta_{ni}}{Z_i(j)} \le \min_{m \neq i} \left\{ \frac{w_m \delta_{nm}}{Z_m(j)} \right\} \right) \mathbb{P}\left(\frac{w_i \delta_{ni}}{Z_i(j)} \le \min_{m \neq i} \left\{ \frac{w_m \delta_{nm}}{Z_m(j)} \right\} \right) \right) \\
&= \sum_{i=1}^{3109} \left(\Gamma\left(\frac{\theta - (\sigma - 1)}{\theta} \right) \left(\sum_{m=1}^{3109} T_m (w_m \delta_{nm})^{-\theta} \right)^{\frac{\sigma-1}{\theta}} \mathbb{P}\left(\frac{w_i \delta_{ni}}{Z_i(j)} \le \min_{m \neq i} \left\{ \frac{w_m \delta_{nm}}{Z_m(j)} \right\} \right) \right) \\
&= \Gamma\left(\frac{\theta - (\sigma - 1)}{\theta} \right) \left(\sum_{m=1}^{3109} T_m (w_m \delta_{nm})^{-\theta} \right)^{\frac{\sigma-1}{\theta}} \sum_{i=1}^{3109} \left(\mathbb{P}\left(\frac{w_i \delta_{ni}}{Z_i(j)} \le \min_{m \neq i} \left\{ \frac{w_m \delta_{nm}}{Z_m(j)} \right\} \right) \right) \\
&= \Gamma\left(\frac{\theta - (\sigma - 1)}{\theta} \right) \left(\sum_{m=1}^{3109} T_m (w_m \delta_{nm})^{-\theta} \right)^{\frac{\sigma-1}{\theta}}
\end{aligned}
$$

Notice how the average price paid in county n for goods imported from some other county i is just an expected value, it is independent of i and satisfies

$$
\mathbb{E}\left(\left(\frac{w_i \delta_{ni}}{Z_i(j)} \right)^{1-\sigma} \middle| \frac{w_i \delta_{ni}}{Z_i(j)} \le \min_{m \neq i} \left\{ \frac{w_m \delta_{nm}}{Z_m(j)} \right\} \right) = \Gamma\left(\frac{\theta - (\sigma - 1)}{\theta} \right) \left(\sum_{m=1}^{3109} T_m (w_m \delta_{nm})^{-\theta} \right)^{\frac{\sigma-1}{\theta}}
$$

The fourth equality is comes from the Law of Large numbers; and the fifth equality is a property of independent Fréchet random variables.