CAGE

Transportation Costs in the Antebellum U.S. A New County-Level Dataset with Time, Region, and Direction Specific Freight Rates, 1820–1860

CAGE working paper no. 736

December 2024

Alexander Klein,
Peter Matthews

WARWICK
THE UNIVERSITY OF WARWICK

UKRI Economic and Social Research Council

Transportation Costs in the Antebellum U.S. A New County-Level Dataset with Time, Region, and Direction Specific Freight Rates, 1820–1860[1]

Alexander Klein

Department of Economics, University of Sussex Business School

University of Sussex, UK, CAGE, CEPR

Peter Matthews

Kent Interdisciplinary Centre for Spatial Studies

University of Kent, UK

Abstract

We construct county-to-county transport cost data set for each decade between 1820 and 1860 in the United States using time-, region-, and direction of transport specific freight rates and the historical transport networks. We document several stylized facts about the effects of canals and railways on the average county-to-county transport cost, market access, and the role of new transportation network in the shaping the direction of domestic trade. We show that by 1860, the canals and railways led to the shift of the highest market access region from the Atlantic coast and Mississippi region to the Midwest and the Great Lakes region, and their absence would have increased the transport costs by more than sixty percent in the Northeast and by almost fifty percent in the South. In addition, by 1840, canals had substantially lowered the costs of transporting goods from the Midwest to the east, making the northern route cheaper than the original route via the Mississippi River and the Atlantic coast.

Keywords: transportation, canals, early railways, freight costs, market access

JEL: N71, N91, R40

[1] We are grateful for the financial support of CAGE, University of Warwick. We also acknowledge helpful comments from Dan Bogart, Phil Hoffman, Jean-Laurent Rosenthal, and the seminar participants at Caltech, and the University of Southern Denmark, Odense. All errors strictly are ours. Contacts: alexander.klein@sussex.ac.uk, petergerardmatthews@gmail.com

Section I: Introduction

We construct a data set of county-to-county transport costs in the United States for every decade between 1820 and 1860. The data set is based on a time and region-varying multimodal historical transport network, and detailed multimodal freight rates which differ over time, across regions and with the direction of transport. These are used to calculate the minimal cost path between the counties, providing us with transport costs (in cents per ton-mile) that capture the antebellum construction of canals, railways, expansion of navigable portion of rivers, as well as changes in the regional freight rates.

We build on the work of Atack (2015, 2016, 2017) and Donaldson and Hornbeck (2016) which we advance by introducing time-, region-, and direction-specific freight rates into the calculation of the minimal cost paths. The main purpose of the paper is threefold: (i) provide details of how this data set was constructed, (ii) present a set of stylized facts about the time- and spatial evolution of the county-pair transport costs in the antebellum United States, and (iii) assess the effects of canals and railways on the county-to-county transport cost, and relate them to the changes in the antebellum domestic trade patterns.

In recent decades, economic history has seen an increasing use of geo-spatial methods to calculate region-to-region transport costs, allowing thus to gain deeper understanding of their socioeconomic effects, and revisit old historical debates.[2] Various approaches are used: indicators of the accessibility to transport network, physical characteristics of the network such as length, or algorithms to calculate the minimal cost path between the regions. Especially the minimal-path algorithms allowed scholars to calculate the transport costs more precisely and provide monetary measures – usually in a domestic currency per ton-mile – of the costs of inter- and intra-regional trade. Methodologically, the paper falls into this strand of research.

The calculation of region-to-region, and in our context county-to-county minimal transport cost, consists of three main components: (i) digitized transportation infrastructure, (ii) Dijkstra algorithm, and (iii) cost components for each part of the infrastructure, such freight costs, or/and transshipment costs, which enter the Dijkstra algorithm. Studies have usually used time-varying transport infrastructure but applied cost components which were fixed over time and across regions.[3] This approach captures an *extensive margin* of transport costs changes – changes driven by the physical nature of the transport network. It offers an econometric advantage too by limiting the scope of endogeneity of the transport costs (Donaldson and Hornbeck 2016, page 817). Changes in the transport network were not the only factor driving the changes in the transport costs though. Technological and organizational advances, among others, affected the

---

[2] E.g. Pontarollo et al (2020), Hornung (2015), Berger (2019), Esteban-Oliver et al (2024), Alvarez-Palau et al (2020), Zimran (2020), Donaldson and Hornbeck (2016), Hornbeck and Rotemberg (2024).

[3] Donaldson and Hornbeck (2016) and Hornbeck and Rotemberg (2024) use time- and region-invariant freight rates; Similarly, Zimran (2020) also uses time- and region-invariant freight rates but introduces three different rates for canals and two for navigable rivers; Alvarez-Palau et al (2024) use time-varying freight rates.

freight rates charged by the transport companies, hence county-to-county transport costs too. We call it an *intensive margin* of the transport costs.

The data set we have constructed captures both the intensive and extensive margin of county-pair transport costs. It is intended to shed light on the role of transport network construction such as canals and railways, and the changes in the freight costs charged on each transport mode in the antebellum US. Furthermore, including the intensive margin allows us to capture transport cost changes when the transport infrastructure was completed.

We quantify several facts about the changes in the county-to-county transport costs between 1820 and 1860 and in the location of transport-cost advantaged regions. Two important geographical shifts occurred between 1820 and 1860. First, the region with the lowest average transport costs relative to the rest of the US shifted from East South Central to East North Central. Second, the location of the highest market access changed from the region around the Ohio, Mississippi Rivers and Atlantic coast in 1820 to the regions of Midwest between north of the Ohio River to the Great Lakes in 1860. The physical network of canals and railways and the respective freight rates were, by and large, equally important in explaining the transport cost changes. The counterfactual analysis reveals that removing canals and railways would have led to substantial increase average county-to-county transport costs, in some regions over 60 percent. We also shed new light on the redirection of antebellum domestic trade due to canals and railways by calculating the cost advantage that these new modes of transportation offered.

The structure of the rest of the paper is the following. Section II provides a historical context by discussing the major models of antebellum transportation. Section III presents the methodology and the data sources. Section IV discusses the main trends, spatial patterns, decomposition and counterfactual analysis. Section V shows how our data set can contribute to the debate about the domestic trade patterns in the antebellum US, and section VI concludes.

Section II: Historical Context

The westward expansion of settlement beyond the original thirteen states was one of the defining characteristics of both political and economic development of the United States before the Civil War (e.g. Howe 2009, Taylor 2021). To bridge the gap between the populated eastern seaboard and largely unsettled but fertile western regions, a transportation network had to be designed and built to overcome distance, inaccessibility, and natural barriers. Canals and railways transformed and redirected regional trade from West–South–East direction to West–East direction, and so often steal the glory. But it would be wrong to neglect the contribution of navigable rivers to antebellum trade and technological and organizational advancements in river transport. Indeed, in the trans-Appalachian West, freight was carried primarily by the western river.[4] Contemporaries stressed the role of river steam navigation as one of the main colonization forces of the West (Hall 1848), and riverside towns and cities grew faster than their land-locked peers until freed by railways.[5]

Designing, approving, financing, and constructing canals and railways was a decades long endeavour, often referred to as the ‘internal improvement’ era.[6] Although the ultimate outcome was an impressive network of canals and railways linking the original thirteen states with the expanding western frontier, the process was anything but simple, linear, or well-coordinated. The first attempt to build canals at the end of the 18th century failed for all intents and purposes, leaving behind barely 200 miles of canals. Efforts continued in the first decades of the 19th century, but the political process, legal and constitutional wranglings prevented the optimal interstate and interregional cooperation, and the canals and early railways happened largely in the sphere of private companies and individual states.[7] But they did happen, and this paper focuses on their effect on transportation costs. The rest of this section provides historical context to the construction of the transportation network and freight rates, to guide construction of the new county-pair transport costs data set and the interpretation of the main patterns emerging from this data set.

A: Navigable rivers

During the early ante-bellum period, western settlement moved mostly along the Mississippi river and its tributaries. The most important transfer of this region to the nascent United States happened in 1803 with the Louisiana Purchase. Fur traders and trappers were soon replaced by settlers who, after crossing the Appalachian Mountains to Pittsburgh, purchased supplies, a family flatboat and when the water level was high enough, they floated to their destinations. River navigability was crucial.[8] And it was changing over time as efforts were made to improve their capacity to deliver goods and passengers, as seen on Map 1, which shows the network of navigable rivers between 1820 and 1860.

There were three main river-transport modes: keelboat, flatboat, and steamship. Each had their heydays, experienced changes in freight rates, competed, but also coexisted alongside each other. The flatboat was one of the principal modes of downstream transportation on the Mississippi river before steamships (Clark 1966, Haites, Mak, Walton 1975). Even after the introduction of steam shipping, flatboats on the Mississippi river kept arriving in New Orleans until 1850s.[9] Flatboating included other rivers as well: in

---

[4] Haites, Mak, Walton (1975), pp. 4-5. Trans-Appalachian West includes all or part of the states of Ohio, Indiana, Illinois, Michigan, Wisconsin, Iowa, Missouri, Kentucky, Tennessee, Arkansas, Mississippi, Louisiana.

[5] Twelfth Census of the United States taken in the Year 1900, Volume 1, Table 6, pp. 430-33. Also see e.g. Mahony (1990) on the role of rivers in the settlement of the Midwest.

[6] An excellent monograph on the history of the antebellum internal improvement programs is Larson (2001).

[7] Minicucci (2004). This does not mean that in the antebellum decades, the federal government was not involved in infrastructure projects at all. For example, construction of lighthouses was a federal endeavor.

[8] See Atack (2015) for detailed discussion of the navigability of US river system.

[9] In 1808, 1,049 flatboats arrived in New Orleans, 1,287 in 1826, 1,365 in 1835, 2,792 in 1846-7, and 1,047 in 1852-3 (Haites, Mak, Walton 1975, Table 4, page 21)

1847 3,336 flatboats arrived at Cincinnati.[10] Flatboat freight rates declined dramatically in both nominal and real terms. On the route from Louisville to New Orleans, the average flatboat freight rate declined from $16.55 per ton in 1810-19 to $5.55 per ton by 1850-60.[11] The major reasons were shorter return journey, and river improvements as flatboating was a hazardous business and greater safety reduced passage time, even made it easier and less costly for flatboating to operate at night. For example, on the Louisville-New Orleans route, the passage time declined from 30-35 days during the early 1800s to 18-24 days in the 1840s.[12] All these factors helped lower flatboating freight rates, enabling them to co-exist along steamships until the late 1850s.

As for keelboats, estimates suggest that on all western rivers combined, the peak number of keelboats in operations was only about three hundred (Baldwin 1941, page 181). Keelboating was a long and arduous journey with a round trip from Pittsburgh to Louisville taking nearly two months and boats made typically three round trips a year.[13] The freight rates of keelboat shipping did decline by early 1820s from $5.00 per hundred pounds to $1.25 per hundred pounds.[14] However, it was not enough when they faced competition from steamships whose low freight rates practically ended the use of keelboats on the major trunk routes, leaving them confined to the shallow waters of the upper Mississippi, the Ohio river, the Missouri river and their tributaries.

Attempts to introduce steam driven transportation on the western rivers go back to the early 18th century. However, early steamships, being structurally similar to ocean vessels, were poorly suited to the shallow waters of western rivers. Therefore, a series of modifications and improvements to design and technology was needed to generate the productivity advances behind their commercial dominance over keelboats and flatboats. Once the phase of experimentation and learning was over, the steamship became the dominant form of river-transport, as documented by Table A4 in Appendix, which shows an initial rapid rise in steamship transportation, followed by a maturity stage from cc 1840s.

Steamships led to a substantial decrease in freight rates and passenger rates.[15] The sources of the freight rate decline were numerous, including increase in the ratio of carrying capacity to tonnage, the fall of insurance costs, and shorter journey times. In total, the duration of a return voyage from New Orleans to Louisville in 1860, excluding the number of days in ports, was only about forty percent of that required in 1815-19. Steamboat journey time also declined elsewhere: New Orleans to St. Louis decreased from

[10] Haites, Mak, Walton (1975) Table 4, page 21.
[11] The patterns of freight rates in flatboating were influenced by seasons and a detailed discussion about it is offered in Haites, Mak, Walton (1975), pp. 82-87.
[12] Haites, Mak, Walton (1975), page 79, footnote 11
[13] Haites, Mak, Walton (1975), page 19
[14] Haites, Mak, Walton (1975), page 40
[15] We document them in Appendix, Table A5.

an average of 25 days in 1815 to three days in 1860; and Louisville to Cincinnati fell from, on average, 40 hours in 1819 to 15 hours in 1840.[16]

Steam boating was not just restricted to major western rivers such as Mississippi and Ohio, often called trunks. Tributaries saw the introduction of steam-powered vessels too, though later, and the freight rates were three to four times higher than the usual ones to New Orleans.[17] Since steamboats on tributaries provided less dramatic cost saving than keelboating or flatboating, all three river transportation technologies could coexist in the ante-bellum times.[18]

B: Canals

Water transportation was further extended by canals. Connecting natural waterways or parallel to a single stream to avoid obstructions, canals created a network of routes allowing more effective transportation. The first canals were built in the late eighteenth and early nineteenth centuries and by 1860 over 4,000 miles of them criss-crossed the north-east, what would be later known as Midwest, and some parts of the south.[19] Map 2 shows the canals system for each decade between 1830 and 1860, and Tables 1 and 2 break down the canal system by regions, mileage, and states.[20] Several facts stand out. First, even before the opening of the Erie Canal, canal building had occurred in coastal places such as New Orleans, eastern coasts along New England and the South, although they were short and designed for ships. Second, Pennsylvania had the longest network of canals – 1,105 miles – followed by the state of New York – 935 miles – and Ohio – 879 miles. Third, geographically, the Middle Atlantic region dominated the network, with almost fifty percent of total canal mileage, followed by the Midwest with thirty-two percent, the South with thirteen percent, and New England states with about five percent.

Canal construction went through three cycles as measured by canal investment.[21] The first cycle lasted from 1815 to 1834, peaked in 1828, completed 2,188 miles and invested over 58 million of current dollars. Canals built in this phase included the Erie and Champlain canals in New York, the Mainline of the Pennsylvania system, and the eastern trunk line of the Ohio state network, the Ohio and Erie Canal. Furthermore, several private or mixed canal enterprises were built: the Blackstone, the Farmington and the Hampshire and Hampden in New England; in the middle Atlantic states these included the Delaware and Hudson, the Morris, the Delaware and Raritan, the Schuylkill Navigation, the Union, and the Chesapeake and Delaware. In the second cycle, from 1834 to 1844, the Erie Canal was enlarged, Ohio extended the Miami Canal north of Dayton, Indiana connected Lake Erie and the Ohio River, Illinois

---

[16] Dixon (1909): table called 'Average time of steamboats between points named', page 29.
[17] Table A6 in Appendix shows typical freight rates between Louisville and selected tributaries, and between Louisville and New Orleans, in the 1840s.
[18] Haites, Maik, Walton (1975), page 55
[19] Report on the Canals of the United States (1883), Table 1, Table 2.
[20] The total canal mileage in Table 1 is lower than in Table 2. It is because Table 1 is based on digitized historical maps and Table 2 on the US Census report. The difference is only 6 percent.
[21] This discussion is based on Segal (1961) in Goodrich ed. (1961), chapter 4.

connected Lake Michigan with Illinois river, Maryland invested into the improvements of Chesapeake and Ohio Canal, and Virginia into the canal between Richmond and Buchanan. The third wave, from 1844 to 1860, saw the improvements and extension of the existing projects such as the Erie Canal enlargement, the Wabash and Erie Canal, the completion of the work on the James River Canal, and of the Black River and Genesee Value canals in Virginia and New York respectively.[22]

The effect of canal system on the transportation costs needs to be assessed carefully.[23] Undoubtedly, the freight rates declined. For example, in 1817, the average freight rate for shipments between Buffalo and New York using wagons and Hudson River was 19.12 cents per ton-mile. Once the Erie Canal was completed, this rate fell to 1.68 cents per ton mile and by mid 1850s, the average freight rate between Albany and Buffalo was one cent.[24] The Erie Canal on its own was a striking success and the most important interregional mode of transport in the antebellum period. Its traffic, which consisted mainly of agricultural commodities, increased from 54,000 tons in 1836 to 1.9 million in 1860.[25] On the other hand, the Erie Canal proved to be the only financially successful venture, with others failing to develop a large volume of traffic, so that in 1859, the Erie Canal accounted for more than thirty percent of the total US canal transportation services, but the Pennsylvania Canal system, a distant second, barely ten percent, and the rest split among the remaining thirteen canal systems.[26] Nonetheless, the canal system led to a significant decline in the transportation costs and many view canals as an important factor that changed the patterns of domestic trade.[27]

C: Railways

The emergence of a new mode of transportation does not usually happen as a clean break from the past. Instead, the new mode coexists alongside the old one for some time. Railways were no different. Their relative merits were discussed already in the 1820s along with canals, especially in Massachusetts and Pennsylvania. Pennsylvania chose to build the Mainline System of canals while Massachusetts and Baltimore, Maryland chose railways (Vance 1995). Baltimore's choice was driven by the decision of Maryland state to promote the Chesapeake and Ohio Canal which disadvantaged Baltimore since it redirected trade away from the city to Alexandria and Georgetown. Thus, Baltimore secured a charter in 1827 for what would become Baltimore and Ohio Railroad. Its planned Ohio River terminus was not reached until 1853, but the project served as impetus and incentive to others for the following reasons (Fishlow 2000). First, railways held the prospect of more rapid and convenient transport. Second, despite still being subject to some geographical constraints, they offered more freedom in choosing the location

[22] Goodrich ed. (1961), page 172, Table 1 and pp. 183-205; Report on the Canals of the United States (1883), pp. 2-30.
[23] For a critical discussion of the research on the US canals, see e.g. Ransom (1964), and Shaw (1984).
[24] Goodrich ed. (1961), pp. 227-228.
[25] Goodrich ed. (1961), page 229
[26] The calculation is based on Fishlow (2000), Table 13.3, page 562.
[27] Clark (1966), chapter X, Pred (1980), chapter 3.

than canals. And third, unlike waterborne transport, they offered transport all year long, including the winter months.

The 1830s witnessed the first round of railway construction with over 3,000 miles of railways laid by 1840. As we can see on Map 3, the emerging network had two directions. Railways in the east-west direction included the ill-fated Erie Railroad of New York venture, the Western Railroad in Massachusetts, two Pennsylvania lines and the original Baltimore and Ohio Railroad. The second direction was north-south along the Atlantic coast. The regional distribution of mileage in Table 6 shows that the South had forty-two percent of US railroads while the Northeast had fifty-eight percent. The depression from 1839 to 1843 slowed down construction, pausing the expansion of the east-west railways before they reached their western terminus, though western interior lines, which relied on state assistance, were more affected than those in New England. Lines in the South suffered the same fate as the western lines, as is evident from Map 3 and Table 3. Nonetheless, railway construction continued, and total mileage still increased between 1840 and 1850 from about 3,100 to over 8,500 miles. An important feature of early railways was their termination at canals, navigable rivers, the Great Lakes, or coastal ports. Only later, once proven financially successful, did railways venture out on their own, beyond being waterway feeders.

After the sluggish 1840s, construction accelerated and by 1860, the total railway mileage almost quadrupled to over 30,000 miles. The regional difference, which began to emerge in the 1840s, became more pronounced as the Northeast region now had more than sixty-two percent of the railways network, and the South only about thirty-three percent. This regional disparity is clearly visible in Map 3. The Midwest was the principal region of rail construction and Chicago emerged as a crucial railway center and node of ten different railway lines spanning over 4,000 miles. Initially, the freight rate differential between canals and railways favoured canals, but by 1860 the difference had been dramatically reduced, largely due to the increased factor productivity (Fishlow 2000). Although the freight costs advantage of canals, and waterways in general, persisted, railways offered other advantages, including greater speed, all-year-long transportation, and less transshipment, which ultimately allowed railways to succeed canals as the main mode of transportation.

D: Antebellum transportation network

Before canals and railways, three natural gateways linking the West with the eastern seaboard could be outlined.[28] The first, the northern gateway, runs from the Great Lakes eastwards, either along the Hudson or Mohawk River valleys or down the St. Lawrence River and New York City is the eastern terminus. The second, the north-eastern gateway, links Pittsburgh and Wheeling on the Ohio river to Baltimore and Philadelphia. The third, the southern gateway, was the Mississippi river and New Orleans as the main southern port.

[28] We follow Kohlmeier (1938) and Haites, Mak, and Walton (1975).

Map 4 puts together all modes of transportation in year 1840 and 1860, showing the major transportation developments for each gateway. The northern gateway saw, among others, the construction of the Erie Canal, and one of the early railroads – the New York Central and New York and Erie railroads. Among the major developments in the north-eastern gateway were the Pennsylvania Canal system linking Pittsburgh and Philadelphia, and the completion of the Baltimore and Ohio railroads reaching Pittsburgh and Wheeling. The major event in the southern gateway was the introduction and development of steamship on the Mississippi river. The antebellum transportation network and its changes have been subject to numerous studies.[29] In the rest of the paper we will examine it from the point of its effect on the county-to-county transport costs.

Section III: Methodology and Data

In this section we discuss the methodological approaches, and the data sets we use in this paper. It consists of two main parts. In the first part, we present the methodology which calculates the county-pair lowest-cost transport costs for each decade between 1820 and 1860 and in the second, we discuss the data sources.

A: Methodology

The calculation of the lowest-cost route between the pair of counties *o-d* (origin-destination) at time *t* requires two inputs. The first input is the network of available transportation routes between counties *o-d* which existed in year *t*. The transportation network consists of nodes and arcs where the nodes are points in space, and arcs are means of transportation available in year *t*. Arcs include all available modes of transportation, for example navigable rivers, canals, or rails. The second input is the cost of transporting freight along each arc in each year *t*.

The calculation of minimal transport cost between counties *o-d* in year *t* consist of three steps: (i) construction of transportation network in year *t* where *t={1820, 1830, 1840, 1850, 1860}* with all available modes of freight transport, (ii) calculation of the cost of freight transport along the arcs between the pair of counties *o-d* at time *t*, hence historical freight rates for each of the existing transport modes, (iii) applying Dijkstra algorithm to calculate the lowest-cost route between the pair of counties *o-d* in year *t*.

Input 1: Transportation network We create a US historical transportation network by combining digitized maps of all modes of transportation which exited in year *t* (hence in years 1820, 1830, 1840, 1850, and 1860). Specifically, we use historical transport routes of navigable rivers, canals, and railways as digitized by Atack (2015, 2016, 2017, and the subsequent updates) and wagon, lakes, and ocean shipping lines created by Donaldson and Hornbeck (2016). Furthermore, we follow the approach of Donaldson and Hornbeck (2016) to create linkages between different network components to account for the transshipment costs that occurred when freight was being transferred to or from one of six modes of

[29] For example, Goodrich ed (1961), Fishlow (1965), Clark (1966), Pred (1980), Atack (2010), Atack and Margo (2011), Zimran (2020).

transportation: navigable rivers, lakes, canals, ocean, or rail. Then, we connect it to individual counties and calculate the average transport costs between the counties as the transport costs between the geographical centroids of each county pair.

Connecting each county centroid to the historical routes of navigable rivers, canals, and railways presents a challenge. In the antebellum US, wagons played an important role in transporting freight from a place in a county to the nearest navigable river, canal, or railway (e.g., Larkin, 1988; Mahoney, 1990). The challenge we face is that the nearest distance of the county geographical centroid to the river, canal or railway is usually different from the average distance that wagons travel to deliver freight. Fogel (1964) was among the first to recognize the importance of measuring the within-county distance and suggested splitting each county into grids and calculating the average of the nearest distances from each grid. We follow the procedure developed by Donaldson and Hornbeck (2016) which offers a refined way of connecting county centroids to the nearest transport mode. Specifically, within each county, we have created 200 random points, calculated the distance from each point to the nearest river, canal, or rail respectively, and taken the average of the distances. The result is a historical transportation network in the benchmark years of 1820, 1830, 1840, 1850, and 1860. We should note at this point that nodes and arcs are time- and region-varying because of the construction of canals and railways, improvements in rivers, and the westward expansion of the United States.

Input 2: Historical freight rates Our main principles in deciding what historical freight rates to use were threefold: (i) provide time variation of mode-specific freight rates, (ii) provide regional variation of mode-specific freight rates, (iii) provide direction-specific freight rates. These principles are grounded in the historical realities of antebellum freight transportation, which we have discussed in Section II. The early decades of the 19th century saw the introduction of steamships, which crucially changed transportation on navigable rivers and lakes: it lowered the freight rates and reduced travel time for downstream and especially upstream return journeys. As a result, freight rates were direction specific. Furthermore, the reduction of freight rates due to technological advances in steam shipping and the time lag in the spread of this technology, especially on the tributaries of the Mississippi and Ohio rivers, introduced time and regional variation in the freight rates. The decades between 1820 and 1860 witnessed the construction of the canal system, which was not uniform in time or space, introducing thus time- and regional-variation in the canal freight rates as well. The emergence of the railway system was, similarly to the canals, an endeavor that varied greatly by decades and across regions, thus affecting freight rates over time and across regions.

Transport cost calculation We combine the historical transportation network with the freight rates, use the Dijkstra algorithm, and calculate the minimal transport costs between each county pair in the benchmark years. The result is a data set of county-pair minimum transportation costs for freight in each benchmark year. The total number of county-pair observations depends on what we consider the land of the United States. Historically, the boundaries of the United States between 1820 and 1860 were shifting south and westward. This expansion took many forms, and usually before an acquired piece of land became a member of the Union, it was organized as a territory. However, many parts of the future United States were being settled years before they were formally organized as a Territory. This poses a challenge for us: should we be strictly historical and consider only states and territories present in the United States in each benchmark year, or should we consider all forty-eight contiguous states, hence also those parts which, while not being part of the mainland United States, were being settled or annexed? Fortunately, the historical boundaries of the United States in the 1820-1860 decades are nested within the boundaries of forty-eight contiguous states; thus, calculating the county-pair transport costs for all forty-eight contiguous states automatically calculates the county-pair costs within the historical boundaries of the United States. It is then straightforward to select only those parts of the United States that the researcher is interested in. Calculating transport cost outside the official borders of the United States into the unorganized territories also offers an opportunity to estimate the cost of the freight transport that settlers had to incur during the decades of westward settlement.

Limitations There are a few limitations to this data which need to be recognized. First, there is no economies of scale or congestion effect in freight transport. We do not include different railway gauges, which require additional costs to modify railway cars, locomotives, and tracks (Taylor and New, 1956). Second, since we focus on the average transport costs, we assume away the seasonal variation of the freight rates on waterborne transport. Third, even though we take into account the direction-specific freight rates on rivers and lakes, the freight rates might be direction-specific on railways too due to back-haul transport. And fourth, we do not consider pre-1820 turnpike construction in the New England and Middle Atlantic states. Instead, we follow Donaldson and Hornbeck's (2016) approach to wagon transport that we described earlier in this section. Overall, we expect that the county-to-county transport costs will be robust to the inclusion of these additional costs, especially over large distances. It is the transport costs over short distances that might be affected but even there, we do not expect that any extra cost on railways or seasonal spike on waterway freight rates would make wagon transport a preferred choice for the cost of wagon transport is larger by an order-of-magnitude.

B: Data: historical freight rates

In this section, we present the freight rates that enter the calculations of the minimal-cost route. We will first discuss the data sources and then present them by mode of transportation. We leave the technical details in the Appendix.

The historical freight rates were collected or calculated from several sources. Freight rates for rivers come from the detailed and authoritative study on Western River transportation by Haites, Maik, and Walton (1975), which, using many historical sources, provides a quantitative assessment of the transportation sector on navigable rivers in the antebellum decades. This is accompanied by the classical studies of

Hunter (1949) on the steamboats on Western rivers, and by Taylor (1957) on the transportation sector in general. Furthermore, we use a recent study by Binder (2011) on the antebellum transportation sector, and the US Census report on the transportation sector by Newcomb and Ward (1901). Freight rates for canals were collected and calculated from historical sources, including von Gerstner (1842-3), Annual Report of the State Engineer and Conveyor of the State of New York (1854), US Census Report on Canals (1883) and Aldrich Report (1893). These were complemented by the studies of Ransom (1967) and Scheiber (1969). Railway freight rates come from the study of von Gerstner (1842-3), who assembled vast historical sources about the early railways, especially in the 1840s. The Census report by Newcomb and Ward (1901) and Fishlow (1965) provide data for the remaining years. Freight rates on the Great Lakes come from Newcomb and Ward (1901) and Binder (2011), as well as coastal shipping and wagon transport from Binder (2011) and Donaldson and Hornbeck (2016).

Table 4 depicts the time and regional variation of freight rates on navigable rivers and the Great Lakes and freight rate for a route from New York City to New Orleans. We also provide the breakdown between downstream and upstream freight rates, an essential consideration because the advent of steam-powered boats allowed upstream transportation. We see that the freight rates between 1820 and 1860 declined substantially. The most significant decrease in the freight rates was on the Hudson River (85%), followed by Mississippi (70%) and the transportation on the Great Lakes (between 72% and 79%). Major tributaries and coastal shipping also saw the freight rates drop by about 46% and 48%, respectively. There is a considerable regional variation, though, with the average rates on the tributaries being considerably higher than on the Mississippi or Ohio rivers. For the rivers where no freight rate data was available, we proxied freight rates by assuming that they were inversely proportional to river discharge, up to a point, and that the percentage change in rates between decades was the same as that for rivers with known freight rates[30].

Table 5 shows the freight rates on canals between 1830 and 1860. Again, we report freight rates for individual canals whenever data makes it possible. If there was no historic data available for a canal, we gave the canal with a freight rate based on the rate associated with nearest canal features with known freight rate.[31] Overall, canal freight transport experienced a substantial decline between 1830 and 1860, in the case of the Erie Canal, as much as 64%. Freight rates varied by region, with the canals in Pennsylvania having higher freight rates than the canals in Ohio or the Erie Canal.

The average railway freight rates by region are depicted in Map 5 The freight rates declined over time, and we can observe substantial regional variation. In 1840, New England had the lowest average freight rate, while Southeast Central had the highest. By 1860, New England had lost its position as the lowest railway freight rate region, which is attributable to the relatively low railway density, especially in Maine,

[30] The appendix contains more detail on the imputation procedure when river freight rates were missing.
[31] Appendix contains detailed discussion how we implemented them in Dijkstra calculation and the imputation procedure when the canal freight rates were missing.

as seen on Map 3. The underlying data for region is in the Appendix (Table A1-A3 and Map A1) together with the technical details how we used in Dijkstra calculation. We see in Map 5 that the regional breakdown follows state borders in 1840 and 1850 whereas different regions are displayed in 1860. Freight rate data for 1840 are for individual railway lines which were geographically confined within the state borders. For 1850, we follow Fishlow (1965) who assigned rail freight rates to regions. As the railway system expanded, the state or regional character of the early railway system gave way to a railway system where rail companies expanded beyond states and regions. Therefore, in 1860 the freight rates are depicted by the areas served by specific railway companies rather than by states or regions.

Section IV: County-pair transport costs

In this section, we present the new data set. First, we discuss the main trends and spatial variation; second, we analyze the contribution of the transportation system and freight rates to the changes in county-pair transport costs. The transportation cost changes between 1820 and 1860 reflect both the construction of new modes of transportation, such as canals and railways, and freight rate changes, due to both the new modes of transportation, and the changes in the technology and organization of freight transport (Chandler, 1977; Vance, 1995; Fishlow, 2000; Usselman, 2002). We decompose these into the changes in the transportation network – external margin – and the changes in the freight rates – intensive margin. Furthermore, we will assess the importance of canals and railways by performing a counterfactual analysis, calculating minimal county-pair costs without these modes of transport.

A: County-pair transport costs: main trends and spatial variation

Before we present the main results, a brief note about the historical boundaries of the United States is needed. As we have discussed earlier, the shifting boundaries and westwards territorial expansion of the United States throughout the nineteenth century presents a challenge in how to treat the United States territorially. Therefore, while we present transportation costs for the entire United States, we have also explicitly distinguished between regions formally organized as states or territories and those not.

Transport from a county to the rest of the United States Table 6A shows the average transportation costs from the county of origin *i* to all other counties in the United States in each decade between 1820 and 1860 and broken down by the regions of origin.[32] Several facts stand out. First, the region with the lowest average transport cost to the rest of the US in 1820 was East South Central, which is unsurprising since it was the region with the Mississippi River. By 1860, however, it lost this position to East North Central, a region with a dense railway and canal network. Second, the average transport costs from a county to the rest of the US declined rapidly: almost sixty-two percent between 1820 and 1860. The rate of this decline varied over time. The largest declines occurred in the first two decades (25% and 25.7%

[32] We use arithmetic averages, meaning every county has an equal weight. An alternative is a weighted average with population weights, which would weight the destination counties by their relative population size. Each of these averages offers important information, and which is preferred is context-specific.

respectively), before slowing down in the 1840s, and then peaking again, albeit at a lower magnitude, in the 1850s (19.2%). Within the regions organized as states or territories in Panel A, the greatest declines happened in the East North Central, commonly known as the Midwest, and Middle Atlantic regions: 72 and 71 percent, respectively. The decline of the average transport cost from a county to the rest of the US also happened in the South, and it was not trivial: ranging between 67 and 68 percent, if we consider only the part fully organized in either states or territories.

In Panel C, we compare the regional transport costs relative to the US and between two broader regions – the Northeast and the South. Two important facts stand out. First, the transportation costs in the Northeast had declined by twenty percent in the antebellum decades relative to the US average, from 85 to 65 percent, while those in the South declined by about thirteen percentage points – from 82 to 69 percent. Second, the transportation costs in the South relative to the Northeast changed – small in magnitude but important in direction. In 1820, the average transport costs from the counties in the South to other counties in the US were four percent lower than in the Northeast (row 13). This difference had reversed by 1860: the average transport costs from the counties in the South to other counties in the US were six percent higher than in the Northeast.

Within regions Table 6B shows the same summary statistics as Table 6A, but for the average transport costs from a county *i* to the other counties within the same region. Several facts stand out. The region with the lowest average within-region transport costs is New England, followed by Middle Atlantic and East South Central. The primacy of New England is not surprising, given the length of navigable rivers and the construction of early canals. By 1860, this primacy is lost to the Middle Atlantic region though. The decrease of the within-region county-to-county transportation costs between 1820 and 1860 was, similarly to the results in Table 6A, substantial in all regions. Unlike Table 6A, however, the ratio of within-region costs between the South and Northeast was very large (row 10). In 1820, transporting goods within the South was 26 percent more expensive than transportation in the Northeast region. By 1860 this gap had widened, with transport costs within the South being 79 percent higher than in the Northeast.

Spatial variation We use market access measure to show a spatial aspect of antebellum transportation network. We use the approach of Donaldson and Hornbeck (2016) and with the parameters from Hornbeck and Rotemberg (2024) where $MA_i = \sum_j Pop_j \, {\tau_{ij}}^{-\theta}$ defines the market access of county i, Pop_j is the population of county j, ${\tau_{ij}}^{-\theta}$ is trade cost function and θ is trade elasticity parameter.[33] The parameter τ_{ij}, also called 'iceberg trade cost' is defined as ($\tau_{ij} = 1 + t_{ij}/P$) where t_{ij} is the transport cost between county i and j and P is the average price per ton of transported goods. We use the parameters from Hornbeck and Rotemberg (2024) who estimated the value of 3.05 for θ and 38.7 for *P*.

[33] Literature also uses the term 'market potential'.

To assess the occurrence and type of spatial dependency at the overall level in market access values, we calculated Global Moran's indicator (Moran 1948) in 1820, 1840, and 1860. This measure tells us whether market access values are distributed randomly, clustered, or dispersed. The resulting positive and significant Global Moran's indicators are indicative of spatial dependency in the form of clustering: across all years, counties with similar market access tend to occur close together. The contribution of individual counties to this clustering is illustrated by calculating the Local Moran's indicator (Anselin 1995) for each county, identifying certain counties as part of significant high-high or low-low clusters (when a county with relatively high or low market access is surrounded by other counties with similarly high or low market access). For a county to be in high-high cluster, the similarity between its market access and that of its neighbours is significantly greater than would be expected if market access was distributed randomly. We show the Local Moran's indicator overlaid with the transportation network in Map 6.[34]

While market access exhibited a clustered distribution in all years, the location of the clusters of high market access change from 1820 to 1860. In 1820 these clusters are concentrated around the Ohio, Mississippi Rivers and Atlantic coast. In 1860, the clustering of high market access extends across the Midwest, north of the Ohio River to the Great Lakes, while the extent of the cluster along the Atlantic coast is reduced. Clusters of counties with low market access were in west / central states particularly in Wyoming, Utah, Colorado, New Mexico, and west Texas.[35]

Overall, the changes in county-to-county transport costs produced two important geographical shifts in the regions' importance between 1820 and 1860. First, the region with the lowest average transport costs to the other part in the US shifted from East South Central to East North Central. Second, the location of the highest market access changed from the region around the Ohio, Mississippi Rivers and Atlantic coast in 1820 to the region north of the Ohio River to the Great Lakes in 1860.

B: Quantitative importance of transport infrastructure: canals and railways

Decomposition Altogether, over 13, 000 miles of canals and 42,000 miles of railroads were built by 1860. There is little doubt that their effect on the county-to-county transport costs was profound. Furthermore, the differences in canal and railway network density between the Northeast and the South affected the cost of moving goods in each region. However, we should not omit the freight rates' importance as a source of declining county-to-county transport costs. The freight rates charged by steamboats, canal, and railway companies reflected, among others, technological and organizational advances made by them in the antebellum decades.[36] We view the construction of the canals and railways as an *extensive margin* of county-to-county transport cost decline and the changes in the freight rates charged to deliver goods as an

[34] For the cluster analysis using local Moran's I, the threshold for significance was calculated by applying a False Discovery Rate correction to a 95% confidence level (Caldas et al 2006).
[35] The spatial pattern in 1820 is qualitatively similar to Zimran (2020), Appendix, Figure A.3.
[36] The importance of other factors than the transportation infrastructure is discussed in Fishlow (1965, 2000), Haites, Mak, and Walton (1975).

intensive margin. We focus on the transport cost changes between 1820 and 1860: between the year with no systematic canal network and no railways and the year just before the Civil War. We can apportion the transport cost changes between these two effects with the following decomposition:

$$\Delta(TC) = \left[(TI_t \times FR_t) - (TI_{t_0} \times FR_t)\right] + \left[(TI_{t_0} \times FR_t) - (TI_{t_0} \times FR_{t_0})\right]$$

where t_0 denotes the beginning of the period and t the end (in our case 1820-1860), and TI denotes the transportation infrastructure consisting of all modes of transportation that existed at time t_0 or t and FR stands for 'freight rates' at time t_0 or t. We will call the first bracket on the right-hand side as 'transportation infrastructure effect' or 'extensive margin', and the second as 'freight rate effect' or 'intensive margin'.

Table 7 reports the results of this decomposition. As in the previous tables, we consider two broad regions: those which consisted of states or territories (Panel A), and those which still contained unorganized territories (Panel B). The table is also split into two vertical panels: 'Transport costs Δ1820-1860' which presents the data entering the decomposition, and 'Decomposition' which reports the results. Row (1) shows the results for the entire United States. We see that the average county pair costs declined by 15.5 cents. Of these, the changes in the transportation infrastructure account for 48.5% and the changes in the freight rates for 51.5%. The regional breakdown in Panel A shows that the transportation system changes (such as canal and railway construction) account for over 50 percent of the transport cost decline in Middle Atlantic, East North Central, and East South Central with the rest of the decline being due to freight rate declines. This suggests that the contribution of the newly constructed physical transport routes of canals and railways and the freight rates charged by canal and rail companies was quite even handed. The difference between them is, however, substantial in Panel B: the region which still contained unorganized territories. This is not surprising. Although the United States was moving its western frontier towards the Pacific Ocean, large parts had to wait until the post-bellum decades for railway construction. Some parts saw the arrival of railways by the 1860s, such as West South Central, generating clearly visible effects on the average county-pair changes (row (7)), but the transport cost declines in most regions in Panel B were largely due to declining freight rates.

It might be argued that the reason why the transportation infrastructure is not the dominant factor in row (1) – the entire United States – is because it includes the unorganized territories with almost no canals or railways. If we consider only the parts consisting only of either states or territories, as shown in the Appendix, Table A7, not surprisingly the transportation infrastructure plays more prominent role, accounting for over half the county-pair transport cost declines in all regions but New England. But again, the difference between 'transport infrastructure' and 'freight rate' effect is not substantial, except maybe for East North Central and East South Central, where transportation infrastructure contributes ten percentage points more to the decline of the average county-pair transport costs than 'freight rate effect'.

Overall, we can conclude that the contribution of the freight rates and transportation infrastructure is not hugely dissimilar and both factors are quantitatively important.

Counterfactuals We have established the overall quantitative importance of the transportation infrastructure in the previous section. To determine how much canals, railways, or both, contributed to the decrease of county-pair transportation costs, we perform a counterfactual analysis. Specifically, we construct county-pair transport costs in three different counterfactual scenarios: (i) no canals were built, (ii) no railways were built, and (iii) no canal and no railways were built. We then calculate the percentage difference between the 'actual' costs and the 'counterfactual' costs to assess the impact of canal and railway building on transport costs.

Table 8 presents the average actual (Panel A) and counterfactual (Panel B) county-to-county transport costs in each decade between 1820 and 1860 and under all three counterfactual scenarios. Starting with the year 1830, we see that the absence of canals would have meant 3.2 percent higher county-to-county transport costs. This is a rather modest effect, but this number reflects the fact that we consider *all* counties in the United States, while canal building was largely a regional enterprise. If we were to zoom into the region where canals were being built, we could expect the effect of canals to be bigger – something we will do below. After 1830, the effect of canals on transportation costs increased and by 1860, the average county-to-county transport costs across the entire United States would have been ten percent higher without canals (Panel C). The impact of railways, as shown by the second counterfactual scenario shows was initially very small: in 1840 and 1850 the county-to-county transport costs would have been only 1.2 and 2.5 percent lower, reflecting rail's modest beginnings. This changed after 1850 and on the eve of the Civil War, the lack of railways would have raised average county-to-county transport costs by almost twenty percent. And finally, the third counterfactual scenario excludes both canals and railways. Their contribution in 1840 and 1850 is, unsurprisingly, quite modest, reaching ten percent by 1850. However, in 1860, the average county-to-county transport costs would have been almost thirty percent higher if neither canals nor railways existed.

All this is for the entire United States, including unorganized territories in the West. The effects of canals and railways might have been dampened by the fact that they did not include unorganized territories in the West. Indeed, sending goods in 1840 from, for example, Georgia to what would become Nebraska, might have involved railways for part of the journey, but most of it would likely be via wagons and navigable rivers. While canals and railways could have made an impact beyond the regions they were built in, this impact was likely complex, and its proper assessment is beyond the scope of this paper. Nonetheless, we attempt a first look at the complex effects of these new modes of transportation on county-to-county transportation costs by doing two things. First, we examine the counterfactual scenarios at the regional level. Second, we will look at their impact on the minimal routes from the Midwest to the East Coast in Section V.

Table 9 presents the counterfactual calculations for the regions which included states and official territories. It is organized similarly to Table 8 with the actual and counterfactual average county-county transport costs by regions. Starting with scenario 1 – no canals – in Panel A and E we see that the largest effect in 1830 was in the Middle Atlantic region, and it stayed that way until 1860. The magnitude was not trivial – without canals, the county-pair transport costs in this region would have been more than twelve percent higher in 1830 and almost 25 percent higher in 1860. This is unsurprising given that this region had the largest canal network in the country. The effect of canals in the other regions reflect the density of canal network in them – lowest is in the southern regions. Interestingly, in 1860, the effect of canals is similar across the five regions, which we attribute to the growing density of railway network which canals were connected to, thus interconnecting them with each other. Examining scenario 2 – no railways – in Panel B and F, we see that in all five regions, in 1840 and 1850, the contribution is modest, less than ten percent, but increases to more than 35 percent by 1860, reflecting the slow start of railway construction and its rapid increase in the 1850s.

Scenario 3 in Panel C and G – no canals and no railway – represents the 'complete' counterfactual because it considers the newly constructed modes of transport together, thus accounting for their interconnectedness. We see that the effect of canals and railways together is now substantial in the Northeast, ranging between 18 and 38 percent between 1840 and 1850 and 64 and 81 percent in 1860. This contrasts with the southern regions where the contribution of the canals and railways is much lower: between 9 and 13 percent in the years 1840 and 1850 and between 46 and 48 percent in 1860. This difference is partly because the length of railway network in the southern region was not as advanced as in the northeast.

Section V: Antebellum redirection of domestic trade: transport cost view

Scholars have seen the advent of canals and railways as one of the main factors that led to the redirection of the domestic trade from the 'midwest-south-east' corridor with New Orleans as the major port to the 'west-east' corridor with the Great Lakes, canals and later railways as the main modes of transportation.[37] Indeed, Table 10 shows the percentage distribution of shipment from the Trans-Appalachian West to the East and vice versa in the decades preceding the Civil War. We see in Panel A that prior to the opening of the Erie Canals, almost all the trade was carried by the southern gateway, hence the Mississippi river. Indeed, if a farmer in the Trans-Appalachian West wanted to ship agricultural proceeds to the east, the shipment travelled on the Mississippi river to New Orleans from where it was re-exported to the eastern coast. Panel B shows that the southern gateway was also the major route for the products form the east to find their markets behind the Appalachian, though the other two gateways had a more prominent role in delivering goods there than other way round.

[37] Clark (1966), chapter X, Pred (1980), chapter 3.

Table 10 also shows how the development of canals and railroads changed the pattern of trade. The Erie Canal was opened in 1825 and even though the first decade of its operation does not show any impact, this can be due to lack of data rather than lack of operation. Since 1830s, however, the trade was redirected from the southern to the northern gateway. In addition, the opening of railroads in early 1840s redirected the trade towards the north-eastern gateway as well. This does not mean that the Mississippi river lost its function as a significant transport mode, quite the contrary, shipments continued.[38] However, its dominant position was lost as the main port for eastern coast export. Table 11 documents the relative importance of agricultural exports and re-exports via New Orleans in 1839, 1849, and 1860 and the city's gradual decline as the major link to the eastern coast is clearly visible.

Our county-to-county data set allows us to enhance our understanding of the quantitative importance of canals and railways that led to these changes in the domestic trade patterns. Specifically, we can calculate the average transport costs from Midwest to New York City as a major eastern outpost before the canal system was built, and in the year of canal and railway expansion. Table 12 presents the results. Column (2) shows the average county transport costs for a journey from three states respectively – Illinois, Indiana, and Ohio – to New York City via eastern route, which means not taking either the Mississippi, or Ohio River. Column, (3), on the other hand, shows the average county transport costs for a journey from the same states to New York City, but this time by taking the Mississippi and/or Ohio River, thus travelling via New Orleans (the cost breakdown for reaching New Orleans and from New Orleans to New York City is in columns (4) and (5)). Column (6) then calculates the percentage difference between these two costs. We see that in 1820, before the canals were built, the transport costs from the Midwest to New York City was between 30 and 62 percent (depending on the state of journey origin) more expensive by *not* taking the Mississippi and/or Ohio River. This has changed by 1840 when the average transport costs of taking the eastern route decreased dramatically and sending goods to New York City via this route was between 38 to 81 percent cheaper that via the southern route. By 1860, when the railway system expanded, the cost-advantage of the eastern route was even more pronounced: between 65 and 141 percent.

We can further elucidate these changes visually on Map 7 with an example of two counties: Peoria, Illinois, and Franklin, Ohio. In 1820, sending goods to New York City from Peoria, Illinois using the least cost journey consisted of a wagon route to the nearest navigable portion of the Mississippi River, then a river journey down to New Orleans and, a ship journey to New York City. Sending goods to New York City from Franklin, Ohio, was similar. First, a wagon route to the nearest navigable part of the Ohio River, a journey down the Ohio and Mississippi River, and then a ship journey to New York City. In 1860, the cost minimizing route was different in both cases. A journey from Peoria consisted of taking Peoria and Bureau Valley Railroad line, the great Lakes, Oswego Canal, the Erie Canal, and the Hudson River all the way to New Yor City. A journey from Franklin, Ohio, a leg on Cleveland-Columbus-Cincinnati Railroad,

[38] Hunter 1949, pp. 644-645, Haites, Mak, Walton (1975), page 124.

Sandusky-Mansfield-Newark Railroad, the Great Lakes, Oswego Canal, the Erie Canal, and the Hudson River.

Section VI: Conclusion

We have constructed the data set of county-to-county minimal transport costs for every decade between 1820 and 1860. We have done it by advancing the current methodology with the time-, region, and direction of transport freight rates and by doing so, we have introduced the intensive margin of the transport cost changes. We have shown that the contribution of the freight rates and transportation infrastructure to county-pair transport costs was not too dissimilar: both factors were quantitatively important and even when the transportation infrastructure was more important, it was by ten percentage points at largest. The counterfactual calculations confirmed that the antebellum county-to-county transport costs would have been, on average, almost thirty percent higher if neither canals nor railways we constructed and even higher in some of the regions: by more than sixty percent in the Northeast and by almost fifty percent in the South. We have also shed new light on the effects of canals and railways on the redirection of antebellum domestic trade from the southern route via New Orleans to west-east route via the Great Lakes by providing transport costs estimates for each route between 1820 and 1860.

Our guiding principles were to get closer to the historical reality and the future research can incorporate further refinements. For example, we can consider the transport speed, thus the time it takes to deliver goods. This is pertinent when comparing the freight rates on waterways and on railways: while the former was lower, the journey took longer, which was a disadvantage, especially for time-sensitive agricultural products. Future research could also incorporate turnpikes, wagon routes, and precise location of river ports.

References

Aldrich, N. W., Falkner, R. P., & McCain, C. C. (1893). *Wholesale Prices, Wages, and Transportation: Report by Mr. Aldrich, from the Committee on Finance, March 3, 1893* (No. 1394). US Government Printing Office.

Alvarez-Palau, E., Bogart, D., Satchell, M., & Shaw-Taylor, L. (2024). Transport and urban growth in the first industrial revolution. Mimeo, University of Cambridge

Alvarez-Palau, E. J., & Martí-Henneberg, J. (2020). Shaping the Common Ground: State-Building, the Railway Network, and Regional Development in Finland. *Journal of Interdisciplinary History*, *51*(2), 267-296.

Annual report of the State Engineer and Surveyor on the canals of the State of New York 1853-54, Albany, Charles van Benthuysen, Printer to the Legislature, 1854.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, *27*(2), 93-115.

Atack, J. (2015). Historical Geographic Information Systems (GIS) database of Steamboat-Navigated Rivers During the Nineteenth Century in the United States.

Atack, J. (2016). Historical Geographic Information Systems (GIS) database of U.S. Railroads for 1820-1860.

Atack, J. (2017). Historical Geographic Information Systems (GIS) database of Nineteenth Century U.S. Canals.

Atack, J., Bateman, F., Haines, M., and Margo, R. A. (2010). Did railroads induce or follow economic growth? Urbanization and population growth in the American Midwest, 1850–1860. *Social Science History*, *34*(2), 171-197.

Atack, J., and Margo, R. A. (2011). The impact of access to rail transportation on agricultural improvement: The American Midwest as a test case, 1850–1860. *Journal of Transport and Land Use*, *4*(2), 5-18.

Baldwin, L. D. (1941). *The keelboat age on Western waters*. University of Pittsburgh Press.

Berger, T. (2019). Railroads and rural industrialization: Evidence from a historical policy experiment. *Explorations in Economic History*, *74*, 101277.

Binder, J. J. (2011). The transportation revolution and antebellum sectional disagreement. *Social Science History*, *35*(1), 19-57.

Caldas de Castro, M., and Singer, B. H. (2006). Controlling the False Discovery Rate: A New Application to Account for Multiple and Dependent Tests in Local Statistics of Spatial Association. *Geographical Analysis*, 38(2), 180–208

Chandler Jr, A. D. (1977). *The visible hand*. Harvard University Press.

Clark, J. G. (1966). *The grain trade in the old northwest*. The University of Illinois Press.

Dixon, F. H. (1909). *A traffic history of the Mississippi River system* (No. 11). US Government Printing Office.

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

Esteban-Oliver, G., & Martí-Henneberg, J. (2024). The Spanish railway network, 1848–2023. *Revista de Historia Economica-Journal of Iberian and Latin American Economic History*, *42*(1), 153-169.

Fishlow, A. (1965). *American railroads and the transformation of the ante-bellum economy*. Harvard University Press

Fishlow A. (2000). Internal Transportation in the Nineteenth and Early Twentieth Centuries. In: Engerman SL, Gallman RE, eds. *The Cambridge Economic History of the United States*. Cambridge Economic History of the United States. Cambridge University Press, pp. 543-642.

Fogel, R. W. (1964). *Railroads and American economic growth*. The Johns Hopkins Press.

von Gerstner, F. A. R. (1842-3). *Early American Railroads: Franz Anton Ritter von Gerstner's 'Die innern Communicationen'1842-1843*. Stanford University Press 1997.

Goodrich, C., Rubin, J., Cranmer, H. J., & Segal, H. H. (1961). *Canals and American economic development*. Columbia University Press.

Haites, E. F., Mak, J., & Walton, G. M. (1975). *Western river transportation: The era of early internal development, 1810-1860*. The Johns Hopkins University Press.

Hall, J. (1848). *The west: Its commerce and navigation*. HW Derby & Company.

Harris, C. D. (1954). The Market as a Factor in the Localization of Industry in the United States. *Annals of the association of American geographers*, *44*(4), 315-348.

Head, K., & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics* (Vol. 4, pp. 131-195). Elsevier.

Hornbeck, R., & Rotemberg, M. (2024). Growth off the rails: Aggregate productivity growth in distorted economies. Forthcoming in the *Journal of Political Economy*.

Hornung, E. (2015). Railroads and growth in Prussia. *Journal of the European Economic Association*, *13*(4), 699-736.

Howe, D. W. (2007). *What hath God wrought: The transformation of America, 1815-1848*. Oxford University Press.

Hunter, L. C. (1949). *Steamboats on the Western rivers*. Harvard University Press.

Kohlmeier, A. L. (1938). *The old northwest as the keystone of the arch of American federal union: a study in commerce and politics*. Bloomington, Indiana: The Principia Press.

Larkin, J. (1988). *The reshaping of everyday life, 1790-1840*. Harper Perennial.

Larson, J. L. (2001). *Internal improvement: National public works and the promise of popular government in the early United States*. University of North Carolina Press.

Mahoney, T. R. (1990). *River towns in the Great West: The structure of provincial urbanization in the American Midwest, 1820-1870*. Cambridge University Press.

Minicucci, S. (2004). Internal Improvements and the Union, 1790–1860. *Studies in American Political Development*, *18*(2), 160-185.

Moran, P. A. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)*, *10*(2), 243-251.

Newcomb, H. T., & Ward Jr, E. G. (1901). *Changes in the rates of charge for railway and other transportation services*. Washington: Government Printing Office.

Pontarollo, N., & Ricciuti, R. (2020). Railways and manufacturing productivity in Italy after unification. *Journal of Regional Science*, *60*(4), 775-800.

Pred, A. (1980). *Urban growth and city-systems in the United States, 1840-1860*. Harvard University Press.

Ransom, R. L. (1964). Canals and development: A discussion of the issues. *The American Economic Review*, *54*(3), 365-376.

Ransom, R. L. (1967). Interregional Canals and Economic Specialization in the Antebellum United States. *Explorations in Economic History*, *5*(1), 12-35.

Report on the Canals of the United States by T.C. Purdy (1883). In: 1880 Census: Volume 4. *Report on the Agencies of Transportation in the United States, including the Statistics of Railroads, Steam Navigation, Canals, Telegraphs, and Telephones*. Washington: Government Printing Office.

Scheiber, H.N. (1969). *The Ohio Canal Era: A Case Study of Government and the Economy, 1820-1861*. The Ohio University Press.

Shaw, R. E. (1984). Canals in the early republic: a review of recent literature. *Journal of the Early Republic*, *4*(2), 117-142.

Taylor, A. (2021). *American republics: A continental history of the United States, 1783-1850*. WW Norton & Company.

Taylor, G. R. (1957). *The transportation revolution, 1815-60*. New York: Holt, Rinehart and Winston.

Rogers Taylor, G., & Neu, I. D. (1956). *The American railroad network, 1861–1890*. Harvard University Press.

Twelfth Census of the United States taken in the Year 1900, Part 1, Volume 1: Population.

Usselman, S. W. (2002). *Regulating railroad innovation: business, technology, and politics in America, 1840-1920*. Cambridge University Press.

Vance, J. E. (1995). *The North American railroad: its origin, evolution, and geography*. The Johns Hopkins University Press.

Zimran, A. (2020). Transportation and health in the antebellum United States, 1820–1847. *The Journal of Economic History*, *80*(3), 670-709.

Tables

Table 1: US Canal System by Regions: 1820-1860.

Year/Region	1820	1830	1840	1850	1860
	Panel A: total canal mileage				
New England	77	189	250	123	46
Middle Atlantic	25	875	2,037	2,134	2,255
East North Central		280	771	1,288	1,368
South Atlantic	92	149	407	506	494
East South Central		4	4	4	4
United States	194	1,496	3,469	4,056	4,168
	Panel B: percentage from US total				
New England	39.8%	12.6%	7.2%	3.0%	1.1%
Middle Atlantic	12.8%	58.5%	58.7%	52.6%	54.1%
East North Central		18.7%	22.2%	31.8%	32.8%
South Atlantic	47.5%	10.0%	11.7%	12.5%	11.8%
East South Central		0.3%	0.1%	0.1%	0.1%
United States	100.0%	100.0%	100.0%	100.0%	100.0%

Source: Atack (2016)

Table 2: US Canal System by US States.

State	Length (miles)	Percentage from U.S. Total
Connecticut	83.5	1.88
Delaware	14	0.32
District of Columbia	1.21	0.03
Florida	10.5	0.24
Georgia	37	0.83
Illinois	102	2.30
Indiana	453	10.20
Louisiana	19	0.43
Maine	20.5	0.46
Maryland	194.5	4.38
Massachusetts	33.65	0.76
Michigan	3.14	0.07
New Hampshire	5.13	0.12
New Jersey	171.02	3.85
New York	935.21	21.07
North Carolina	25	0.56
Ohio	879.25	19.81
Oregon	0.75	0.02
Pennsylvania	1105.57	24.90
Rhode Island	45	1.01
South Carolina	51.35	1.16
Texas	8	0.18
Vermont	1.06	0.02
Virginia	240.06	5.41
United States	4439.4	100.00

Source: Report on the Canals of the United States by T.C. Purdy, US Census 1883, Table 1, pp. 22, 24 and Table 2, pp. 31, 32

Table 3: US Railway System by Regions: 1830-1860

Year/Region	1830	1840	1850	1860
	Panel A: total rail mileage			
New England		520	2,492	3,436
Middle Atlantic	41	1,071	2,338	5,765
East North Central		223	1,280	9,673
South Atlantic	20	1,189	2,011	5,870
East South Central		118	330	3,295
West South Central		30	76	846
West North Central				1,370
Mountain				
Pacific				23
US	61	3,151	8,527	30,277
	Panel B: percentage from US total			
New England		16.5%	29.2%	11.3%
Middle Atlantic	67.5%	34.0%	27.4%	19.0%
East North Central		7.1%	15.0%	31.9%
South Atlantic	32.5%	37.7%	23.6%	19.4%
East South Central		3.8%	3.9%	10.9%
West South Central		1.0%	0.9%	2.8%
West North Central				4.5%
Mountain				
Pacific				0.1%
US	100.0%	100.0%	100.0%	100.0%

Source: Atack (2015)

Table 4: Freight Rates on Navigable rivers, the Great Lakes, and Coastal Shipping 1820s-1860s, cents per ton-mile in 2000 $US.

Transport mode and routes	**1820s**			**1830s**			**1840s**			**1850s**			**1860s**
	Average	Down-stream	Up-stream	Average	Down-stream	Up-stream	Average	Down-stream	Up-stream	Average	Down-stream	Up-stream	Average
Navigable rivers													
Allegheny	1.35			1.22			0.85			0.84			0.72
Barren	0.56			0.50			0.35			0.35			0.30
Cumberland	0.43			0.39			0.27			0.27			0.23
Hudson	0.67	0.24	1.09	0.39	0.18	0.60	0.24	0.15	0.34	0.13	0.12	0.15	0.10
Illinois	0.50			0.45			0.31			0.31			0.27
Mississippi	0.23	0.18	0.28	0.14	0.14	0.14	0.07	0.08	0.07	0.07	0.08	0.06	0.07
Missouri	0.85			0.77			0.53			0.53			0.46
Monongahela	1.04			0.94			0.65			0.65			0.56
Muskingum	0.58			0.52			0.31			0.31			0.27
Ohio	0.23	0.18	0.28	0.14	0.14	0.14	0.07	0.08	0.07	0.07	0.08	0.06	0.07
Tennessee	0.50			0.45			0.31			0.31			0.27
Wabash	0.39			0.35			0.24			0.24			0.21
The Great Lakes													
Buffalo-Detroit	0.54	0.12	0.96	0.39	0.11	0.67	0.25	0.10	0.40	0.14	0.08	0.20	0.15
Detroit-Chicago	0.48	0.18	0.78	0.33	0.14	0.53	0.19	0.10	0.28	0.09	0.07	0.11	0.10
Buffalo-Chicago	0.37	0.14	0.59	0.26	0.11	0.41	0.15	0.08	0.23	0.08	0.05	0.10	0.08
Coastal shipping													
New York to New Orleans	0.28	0.11	0.46	0.24	0.10	0.39	0.21	0.09	0.33	0.16	0.07	0.25	0.15

Navigable rivers:

Sources:

Mississippi, Ohio rivers: Haites, Mak, Halton (1975), Table E-2, page 168 (downstream), Table 7, page 32 (upstream). Hudson river: Binder (2011), Table 2, page 28. All remaining rivers: Hunter (1949), Table 21, page 659. Original freight rates for all rivers except for Hudson were re-calculated into ton-miles; mileage for all rivers except for Mississippi are from Hudson (1949). Mileage for Mississippi river is from Louisville to New Orleans.

Data in 1820-30 were derived using the growth rates of freight rates on Mississippi between 1820-30 and 1830-1840 as reported in Haites, Mak and Halton (1975).

Freight rates on Ohio river are the same as on Mississippi river because Mississippi freights are calculated for a journey from Louisville to New Orleans. Hudson river freight rates are calculated using 1816 and 1853 benchmarks from Binder (2011); linear interpolation was used to calculate the missing data between the benchmarks. Freight rates in 1860s re calculated using 1866 Mississippi river freight rate and the ratio of Mississippi river freight rates in 1850s and 1860s.

The Great Lakes, coastal shipping:

Sources: Binder (2011), Table 2, page 28; H.T. Newcomb, Edward G. Ward Jr: Changes in the Rates of Charge for Railway and Other Transportation Services, U.S. Department of Agriculture, Washington, Government Printing Office, 1901, Table 40, page 57. Donaldson and Hornbeck (2016).

Notes: The data sources provide the freight rates for the years 1816, 1853, and 1890. Linear interpolation between 1816 and 1853 for the freight rates on the Great Lakes, and coastal shipping was used. Final freight rates are arithmetic averages of the freight rates in the relevant years. Freight rate on the Great Lakes in the 1850s are for the years 1850-1853. The freight rate for the Great Lakes in 1860 were calculated using the ratio of the freight rate from Chicago to Buffalo in 1860 to that in 1850s and applying it to the average 1850s freight rates in the other two routes.

GDP deflator:

Source: Samuel H. Williamson, 'What Was the U.S. GDP Then?' MeasuringWorth, 2022.

Table 5: Canal freight rates by transportation mode 1830s-1860, cents per ton-mile in 2000 $US.

Transport mode and routes	1830s	1840s	1850s	1860
Canals				
Erie	0.57	0.21	0.18	0.20
Ohio	0.58	0.37	0.16	0.27
Wabash and Erie			0.31	0.30
Miami and Erie				0.30
Pennsylvania Main Line	0.75	0.40	0.39	0.45
Delaware and Hudson	0.72			
Illinois Canals			0.23	0.30
Schuylkill River (navigation)	0.34			

Canals:

Sources: von Gerstner (1842-43), Ransom (1967), Scheiber (1969); Annual Report of the State Engineer and Conveyor of the Canals of the State of New York, Albany (1854); U.S Congress, Senate: Wholesale Prices, Wages, and Transportation: Report by Mr. Aldrich from the Committee on Finance, Part 1, Washington: Government Printing Office 1893, Table 51 (page 524). Report on the Canals of the United States, US Census (1883)

Notes: ***Erie Canal**:* 1830s: von Gerstner (1842-3), Table 1.21, page 108, figure for 1839; 1840s is the average in the period 1840-1849 calculated from Ransom (1967), Table 8 (page 23) and Table A-3 (page 31); 1850s: Annual Report (1854), Table W, figure for 1853 1860s: Aldrich reports (1893), Table 51 (page 524), figure for 1862

Pennsylvania Canals: 1830s: average freight rate 1838-1840 from Ransom (1967) (ton-miles from Table 9, page 24) and revenues from Report on Canals (1883), page 738(8); 1840s: average freight rate of 1844-46 calculated from Ransom (1967) (ton-miles from Table 9, page 24) and revenues from Report on Canals (1883), page 738(8); 1850s: freight rate for the year 1853 from Annual Report (1854); 1860s: calculated using Erie Canal freight rate and the ratio of Erie to Pennsylvania canal freight rate from 1850s.

Freight rate in the 1850s is based on the route from Philadelphia to Pittsburgh

***Ohio Canal**:* 1830s: von Gerstner (1842-43), page 395, freight rate in 1839; 1840s: the average of 1830 and 1850; 1850s: from Annual Report (1854), Table W, figure for 1853; 1860s: calculated from Scheiber (1969), Table 11.3 and Report on the Canals (1883), page 747(17)

Miami and Erie: 1860s calculated from Scheiber (1969), Table 11.3 and Report on the Canals (1883), page 747(17)

***Delaware and Hudson**:* 1830s: figure for 1839 calculated from von Gerstner (1842-43), revenue from Figure 1.1, page 127, adjustment to subtract rail revenue from page 126, volume of freight from page 125.

Schuylkill River (or navigation): von Gerstner (1842-43), page 555.

***Wabash and Erie**:* 1850s*:* Annual Report (1854), Table W, figure for 1853; 1860 calculated from Scheiber (1969), Table 11.3 and Report on the Canals (1883), page 747(17)

***Illinois Canals**:* 1850s: Annual Report (1854), Table W, figure for 1853; 1860s: calculated using Erie Canal freight rate and the ratio of Erie to Illinois canal freight rate from 1850s.

GDP deflator:

Source: Samuel H. Williamson, 'What Was the U.S. GDP Then?' MeasuringWorth, 2022.

Table 6A: Average county-pair freight costs: United States and its regions in 1820-1860 ($US2000).

Row	From county of origin to the rest of US	**1820**	**1830**	**1840**	**1850**	**1860**	**%Δ 1820-1830**	**%Δ 1830-1840**	**%Δ 1840-1850**	**%Δ 1850-1860**	**%Δ 1820-1860**
(1)	United States	25.1	18.8	14.0	11.8	9.6	-25.0%	-25.7%	-15.2%	-19.2%	-61.8%
	Panel A: US regions which were already organized as states or territories										
(2)	New England	21.4	15.9	11.4	9.0	6.8	-25.8%	-28.1%	-20.9%	-24.8%	-68.3%
(3)	Middle Atlantic	21.1	15.0	10.0	8.3	6.0	-29.1%	-33.0%	-17.5%	-27.3%	-71.5%
(4)	East North Central	21.5	14.3	10.1	8.3	5.9	-33.2%	-29.3%	-18.1%	-28.5%	-72.3%
(5)	South Atlantic	21.3	15.7	11.6	9.6	7.0	-26.4%	-26.2%	-17.2%	-27.2%	-67.3%
(6)	East South Central	19.8	14.2	10.0	8.6	6.2	-28.1%	-29.6%	-14.5%	-27.1%	-68.5%
	Panel B: US regions which contained unorganized territories										
(7)	West North Central	27.8	20.4	14.2	12.4	10.3	-26.6%	-30.4%	-12.9%	-16.5%	-62.8%
(8)	West South Central	24.5	18.6	14.7	12.6	10.9	-24.1%	-20.8%	-14.8%	-12.9%	-55.4%
(9)	Mountain	43.3	37.3	30.1	26.0	23.0	-14.0%	-19.3%	-13.7%	-11.3%	-46.9%
(10)	Pacific	25.0	20.0	15.9	13.1	10.7	-19.9%	-20.6%	-17.6%	-18.0%	-57.0%
	Panel C: Regional ratios of the average regional transport costs										
(11)	Northeast / United States [Regions (2), (3), (4)] / US	0.85	0.80	0.75	0.72	0.65					
(12)	South / United States [Regions (5), (6)] / US	0.82	0.79	0.77	0.77	0.69					
(13)	South / Northeast [Regions (5), (6)] / [Regions (2), (3), (4)]	0.96	0.99	1.02	1.06	1.06					

Source: see the text for detailed description

Notes: this table shows county-pair minimal freight cost by US regions between 1820 and 1860.

Table 6B: Average within region county-pair freight costs in 1820-1860 ($US2000).

Row	Within region transport costs	**1820**	**1830**	**1840**	**1850**	**1860**	**%Δ 1820-1830**	**%Δ 1830-1840**	**%Δ 1840-1850**	**%Δ 1850-1860**	**%Δ 1820-1860**
	Panel A: US regions which were already organized as states or territories										
(1)	New England	5.3	4.2	3.4	2.3	2.0	-21.7%	-18.5%	-32.3%	-11.2%	-61.7%
(2)	Middle Atlantic	8.4	5.3	3.1	2.3	1.8	-37.3%	-42.2%	-24.2%	-23.8%	-79.0%
(3)	East North Central	10.4	6.6	4.6	3.2	2.3	-36.6%	-30.6%	-28.7%	-30.4%	-78.2%
(4)	South Atlantic	10.5	8.6	6.6	5.2	4.0	-18.1%	-23.4%	-21.5%	-22.0%	-61.6%
(5)	East South Central	9.7	6.9	4.6	4.1	3.2	-28.9%	-32.7%	-12.4%	-20.9%	-66.8%
	Panel B: US regions which contained unorganized territories										
(6)	West North Central	21.1	16.9	11.6	10.3	8.2	-19.9%	-30.9%	-11.4%	-20.3%	-60.9%
(7)	West South Central	16.4	13.8	12.0	10.3	9.0	-16.3%	-12.8%	-14.2%	-12.3%	-45.1%
(8)	Mountain	29.6	27.6	25.2	21.9	21.1	-6.7%	-8.7%	-13.1%	-3.6%	-28.5%
(9)	Pacific	12.1	11.2	9.6	7.4	6.9	-7.1%	-14.6%	-22.8%	-7.5%	-43.4%
	Panel C: Regional ratios of the average regional transport costs										
(10)	South / Northeast [Regions (4), (5)] / [Regions (1), (2), (3)]	1.26	1.45	1.53	1.76	1.79					
(11)	West / Northeast [Regions (6), (7), (8), (9)] / [Regions (1), (2), (3)]	2.46	3.25	3.98	4.76	5.59					

Source: see the text for detailed description

Notes: this table shows county-pair minimal freight cost by US regions between 1820 and 1860.

Table 7: Decomposition of the changes in the average county-pair transport costs ($US 2000) between 1820 and 1860.

	Transport Costs Δ 1820-1860					Decomposition			
	From a county in the US region to the rest of the US	Year 1820	Year 1860	1820 transp infrastructure with 1860 freight rates	Δ 1820-1860	Transport system effect - *extensive margin*	Freight rate effect - *intensive margin*	Transport system effect - *extensive margin* %	Freight rate effect - *intensive margin* %
(1)	United States	25.1	9.6	17.1	-15.5	-7.5	-8.0	48.5%	51.5%
	Panel A: US regions which were already organized as states or territories								
(2)	New England	21.4	6.8	13.8	-14.6	-7.0	-7.7	47.6%	52.4%
(3)	Middle Atlantic	21.1	6.0	13.8	-15.1	-7.8	-7.3	51.8%	48.2%
(4)	East North Central	21.5	5.9	14.2	-15.5	-8.3	-7.2	53.4%	46.6%
(5)	South Atlantic	21.3	7.0	14.0	-14.3	-7.0	-7.3	49.1%	50.9%
(6)	East South Central	19.8	6.2	13.4	-13.5	-7.2	-6.4	52.9%	47.1%
	Panel B: US regions which contained unorganized territories								
(7)	West North Central	27.8	10.3	19.3	-17.5	-9.0	-8.5	51.6%	48.4%
(8)	West South Central	24.5	10.9	16.8	-13.6	-5.9	-7.7	43.2%	56.8%
(9)	Mountain	43.3	23.0	31.0	-20.3	-7.9	-12.4	39.1%	60.9%
(10)	Pacific	25.0	10.7	16.9	-14.3	-6.2	-8.1	43.1%	56.9%

Source: see the text for details on the data sources

Table 8: Actual and counterfactual average county-pair freight cost ($US2000) 1820-1860.

From county of origin to the rest of US	**1820**	**1830**	**1840**	**1850**	**1860**
	Panel A: Actual county-pair transport costs				
United States	25.1	18.8	14.0	11.8	9.6
	Panel B: Counterfactual county-pair transport costs				
Scenario 1: No Canals		19.4	14.7	12.3	10.5
Scenario 2: No Railways			14.1	12.1	11.5
Scenario 3: No Canal nor Railways			15.0	13.0	12.3
Panel C: Percentage difference between counterfactual and actual county-pair transport costs					
Scenario 1: No Canals		3.2%	5.3%	3.7%	10.1%
Scenario 2: No Railways			1.2%	2.5%	19.7%
Scenario 3: No Canal nor Railways			7.4%	10.0%	28.1%

Source: see the text for details on the data sources

Table 9: Actual and counterfactual average county-pair freight cost by regions ($US2000) 1820-1860.

From a county in the US region to the rest of the US	1830	1840	1850	1860	1830	1840	1850	1860
Panel A: Counterfactual: No Canals					*Panel E: Percentage difference: Panel A vs Panel D*			
New England	17.2	13.2	10.1	8.3	8.2%	15.4%	11.5%	22.5%
Middle Atlantic	16.8	12.6	9.8	7.5	12.2%	25.9%	18.8%	24.7%
East North Central	15.7	11.7	9.1	7.2	9.4%	14.9%	10.1%	20.3%
South Atlantic	16.2	12.2	9.9	8.3	3.4%	5.1%	3.6%	19.7%
East South Central	14.6	10.5	8.8	7.4	2.9%	4.7%	3.5%	19.0%
Panel B: Counterfactual: No Railways					*Panel F: Percentage difference: Panel B vs Panel D*			
New England		11.6	9.6	9.3		1.1%	6.4%	36.0%
Middle Atlantic		10.3	8.6	8.2		2.6%	4.3%	36.5%
East North Central		10.2	8.6	8.1		1.0%	3.5%	36.2%
South Atlantic		11.9	10.1	9.6		2.7%	5.8%	38.5%
East South Central		10.3	8.9	8.5		2.6%	4.5%	36.9%
Panel C: Counterfactual: No Canals, No Railways					*Panel G: Percentage difference: Panel C vs Panel D*			
New England		13.7	11.6	11.1		19.5%	28.3%	62.5%
Middle Atlantic		13.3	11.5	10.9		32.8%	38.3%	81.9%
East North Central		12.0	10.5	9.8		18.4%	25.9%	64.1%
South Atlantic		12.6	10.9	10.3		8.8%	13.5%	48.5%
East South Central		10.9	9.6	9.1		9.0%	11.8%	45.9%
Panel D: Actual freight costs								
New England	15.9	11.4	9.0	6.8				
Middle Atlantic	15.0	10.0	8.3	6.0				
East North Central	14.3	10.1	8.3	5.9				
South Atlantic	15.7	11.6	9.6	7.0				
East South Central	14.2	10.0	8.6	6.2				

Source: see the text for details on the data sources

Table 10: Share of Freight Shipments from and to Trans-Appalachian West 1810-1860 (percent).

Years	Northern Gateway	Northeastern Gateway	Southern Gateway	Years	Northern Gateway	Northeastern Gateway	Southern Gateway
	A: Outbound freight shipment				B: Inbound shipments of selected commodities		
1810-1814	0	0	100	1835	32	23	45
1815-1819	1	0	99	1839	33	28	39
1829-1824	0	0	100	1844	36	24	40
1825-1829	0	0	100	1849	48	18	34
1830-1834	1	0	99	1853	61	11	29
1835-1839	13	4	83	1856-57	52	23	25
1840-1844	28	4	68				
1845-1849	38	4	58				
1850-1854	49	4	47				
1855-1860	48	7	45				

Source: Haites, Mak, Walton (1975), Table 1, Table 2, page 8, 9.
Notes: Panel A: The share of northern gateway in 1825-1835 does not include possible shipments by the Erie Canal. The share of northeastern gateway does not include the Pennsylvania Turnpike and the National Road. However, Haites, Mak and Walton (1975: 8) argue they represent only a small portion of the total amount shipped.

Panel B: Selected commodities include sugar, salt, iron, molasses, coffee, groceries, miscellaneous merchandise.

Table 11: Ante-bellum Western Agriculture Exports via New Orleans, 1839-1860.

Commodity	Western exports shipped via New Orleans (%)			Western exports reexported to the Northeast and abroad via New Orleans (%)		
	1839	1849	1860	1842	1849	1860
Flour	53	31	22	31	22	3
Meat Products	51	50	24	30	33	1
Corn	98	39	19	53	31	2
Whiskey	96	67	40	19	7	1
Total Foodstuffs	49	40	17	31	28	3

Source: Pred, A. (1980), Table A.2, page 177.

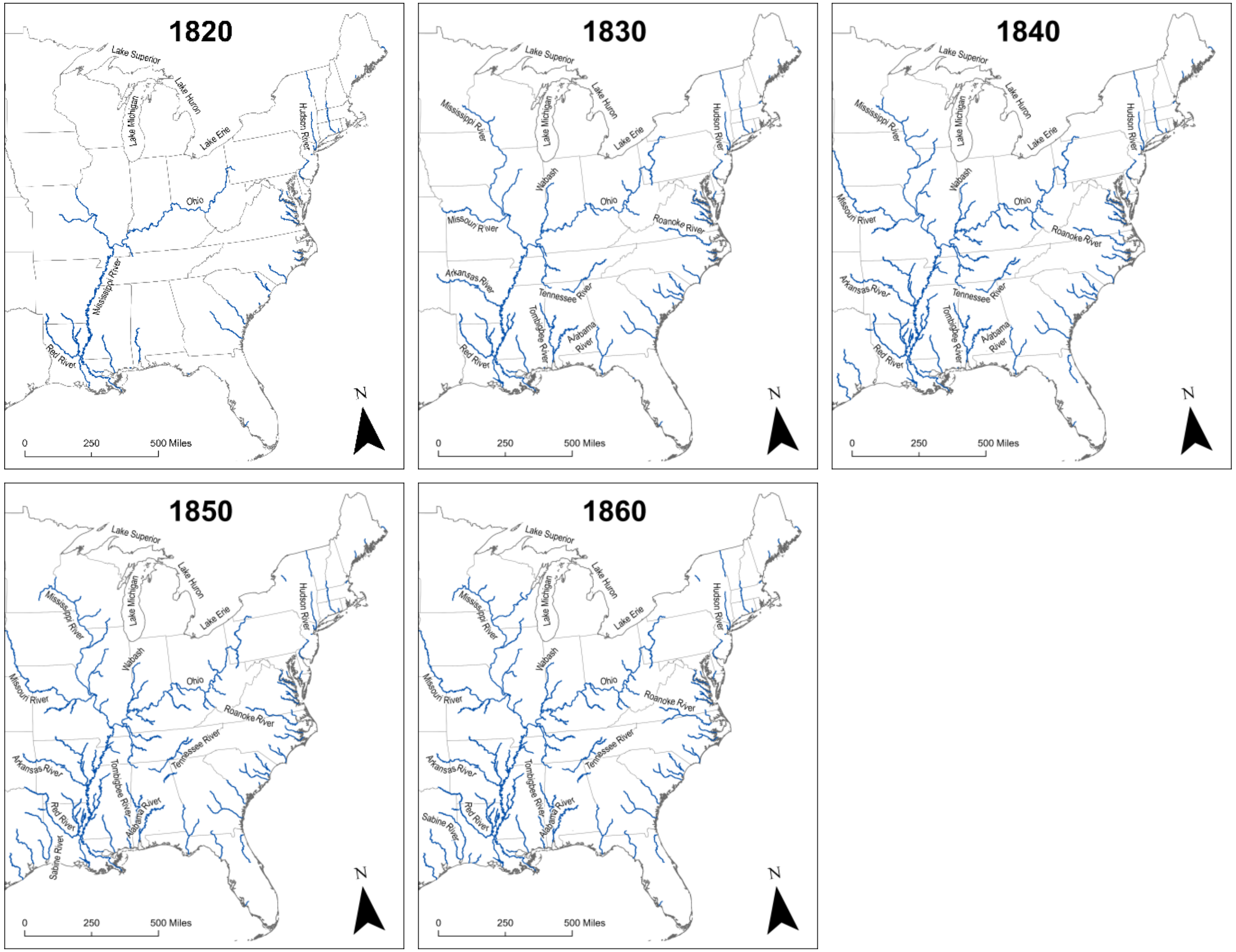
Table 12: Average County-to-County Transport Costs to New York City ($US2000), 1820, 1840, 1860.

(1)	(2)	(3)	(4)	(5)	(6)
States of origin	To New York City using east route (not taking the Mississippi or Ohio Rivers)	To New York City using southern route (via New Orleans)			%Δ [column (2)/column (3)]
		Total	To New Orleans	New Orleans - New York City	
		Year: 1820			
Illinois	37.04	13.74	8.59	5.15	62.9%
Indiana	31.88	15.33	10.18	5.15	51.9%
Ohio	22.13	15.28	10.13	5.15	30.9%
		Year: 1840			
Illinois	5.19	7.19	3.39	3.79	-38.5%
Indiana	5.45	7.85	4.06	3.79	-44.1%
Ohio	3.97	7.20	3.40	3.79	-81.1%
		Year: 1860			
Illinois	3.03	5.01	2.25	2.76	-65.3%
Indiana	2.64	5.21	2.45	2.76	-97.3%
Ohio	2.16	5.21	2.45	2.76	-141.3%

Sources: see the text for details of the data sources
Note: this table shows the average county-pair transport costs from counties in three US states to New York City via two different routes: east route, and southern route via New Orleans.

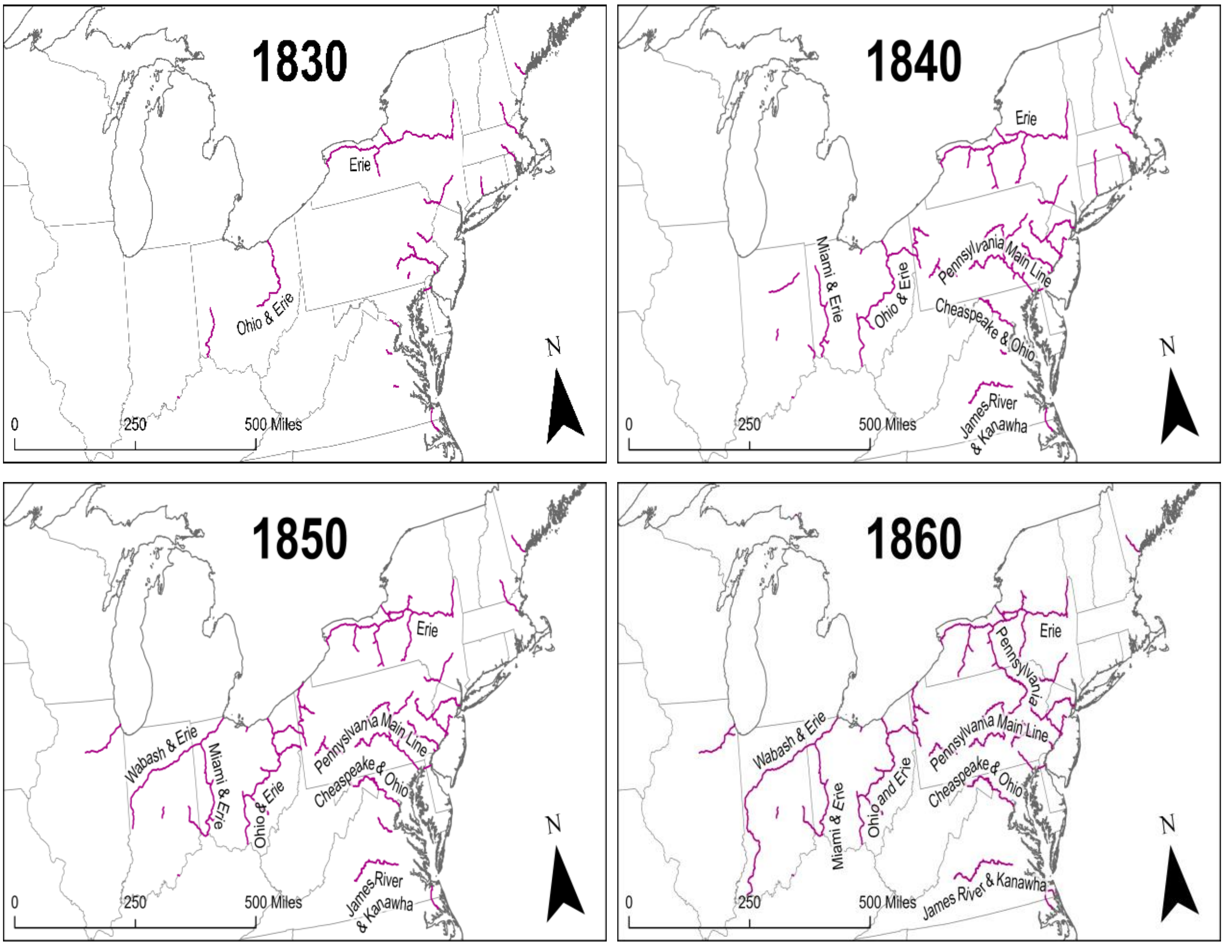
Maps

Map 1: Navigable River System, United States 1820-1860



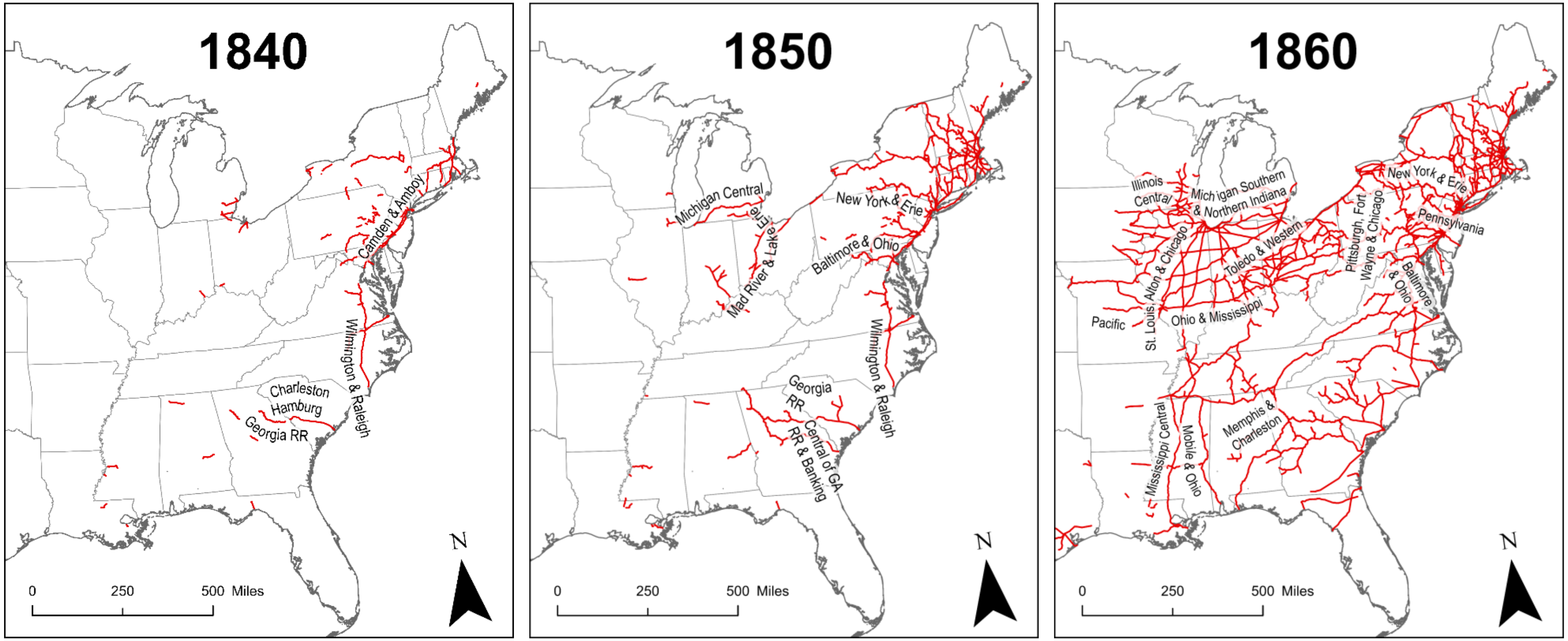
Source: Atack (2015)

Map 2: Canal System, United States 1830-1860



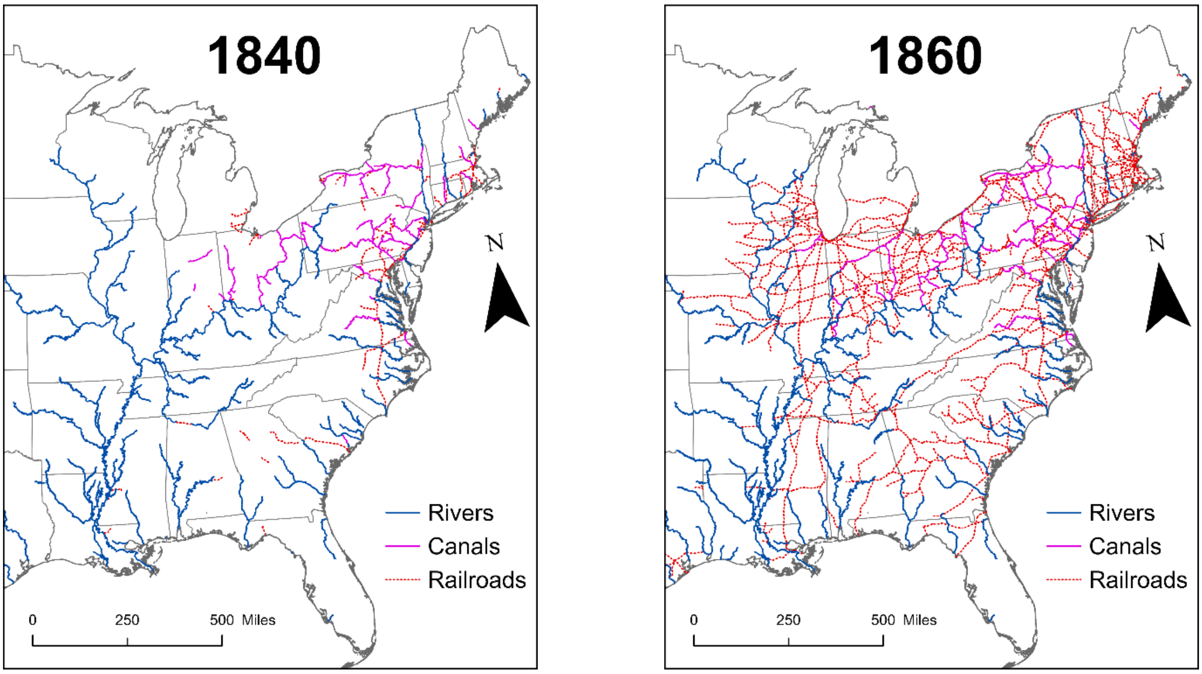
Source: Atack (2017)

Map 3: Railway System, United States 1840-1860.



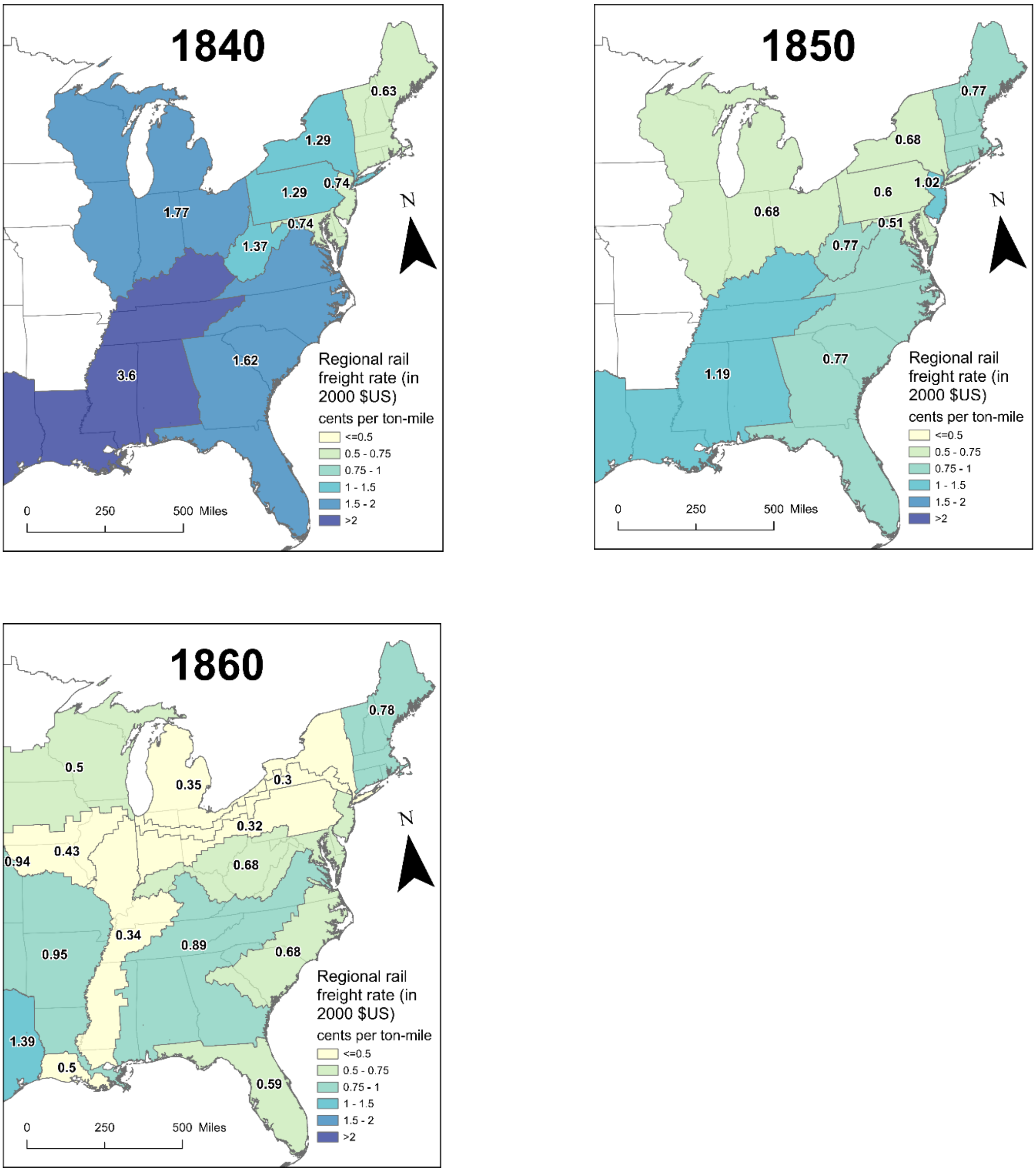
Source: Atack (2016)

Map 4: Transportation System, United States 1840, 1860



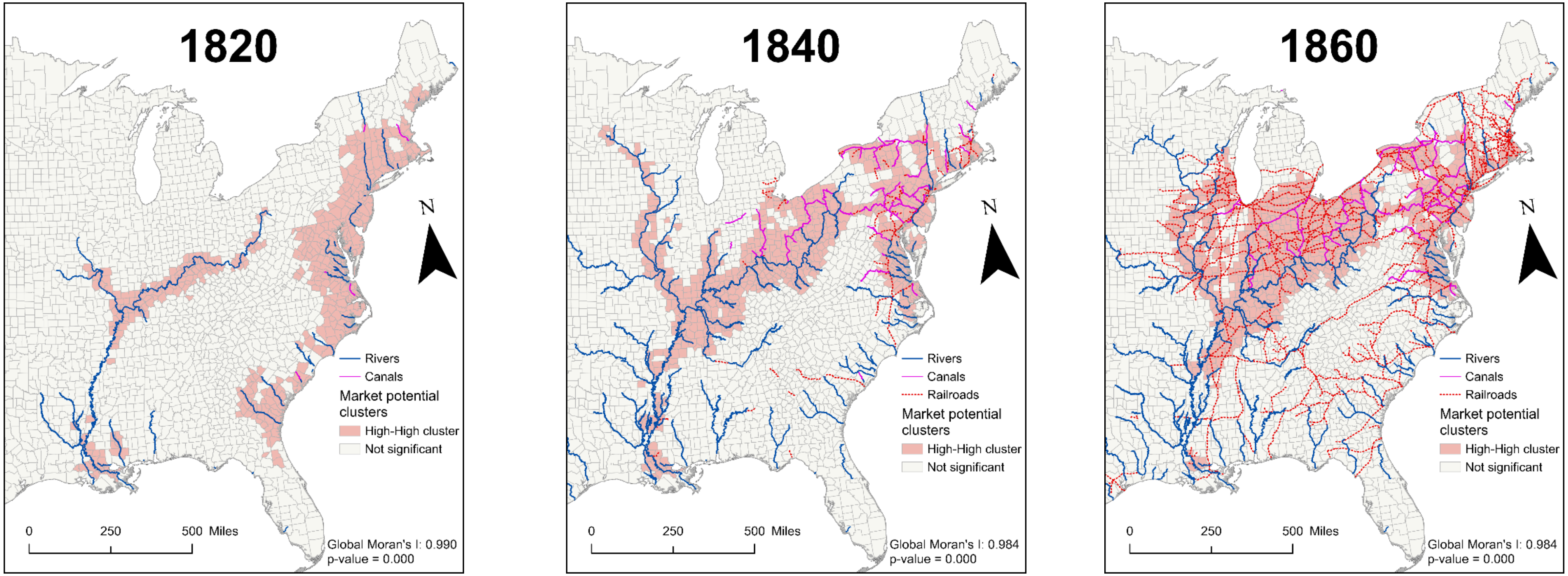
Source: Atack (2015, 2016, 2017)

Map 5: Rail Freight Rates 1840-1860



Source: see the text for details on data sources

Map 6: Market Access 1820, 1840, 1860

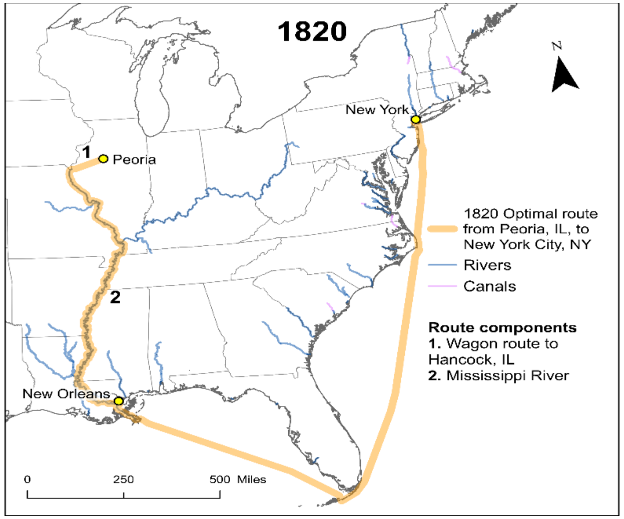


Source: see the text for details on data sources

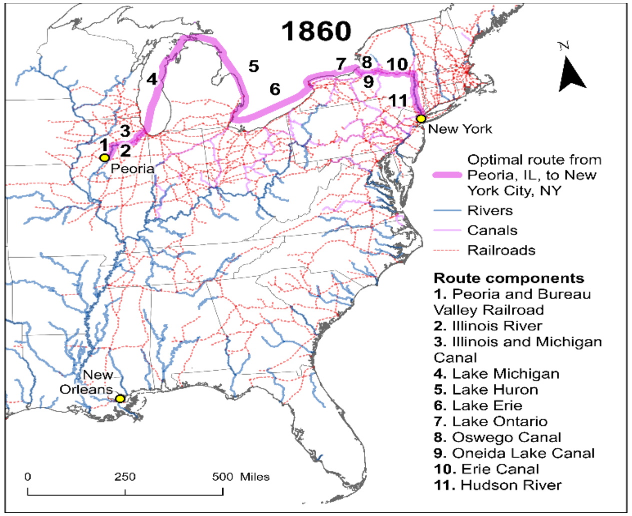
Note: A high-high cluster occurs where counties with relatively high market access are surrounded by other counties with similarly high market access. For a county to be in this cluster, the similarity between its market access and that of its neighbours is significantly greater than would be expected if market access was distributed randomly. the threshold for significance was determined by applying a False Discovery Rate correction to a 95% confidence level.

Maps 7: Cost minimizing routes to New York City in 1820, and 1860

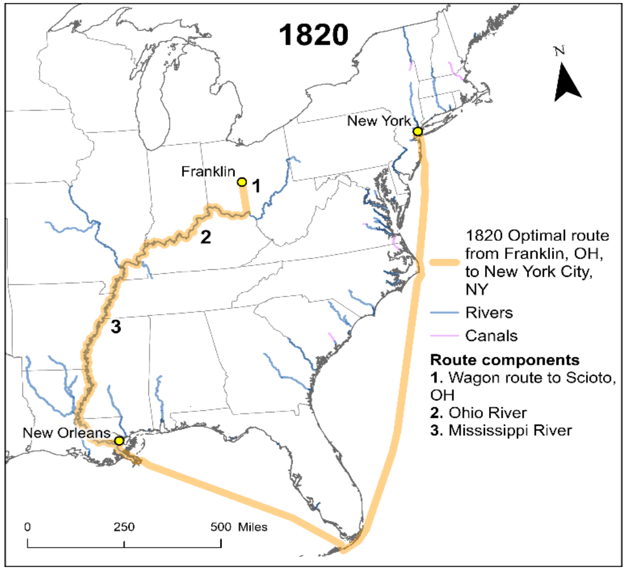
Route: Peoria, Illinois – New York City, 1820



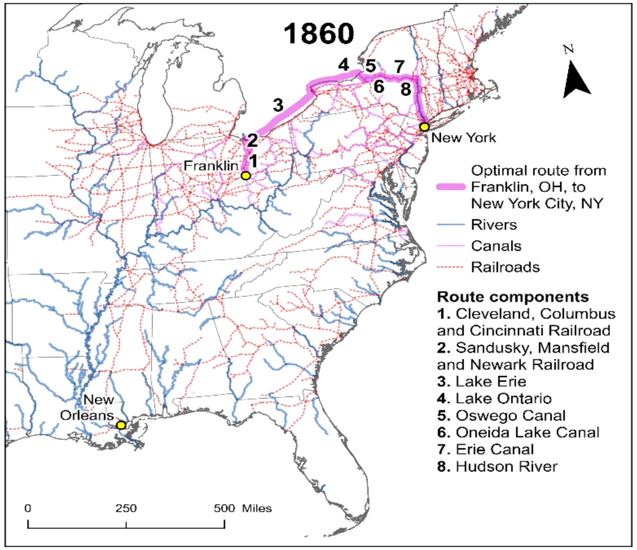
Peoria, Illinois – New York City, 1860



Franklin, Ohio – New York City, 1820



Franklin, Ohio – New York City, 1860



Transportation Costs in the Antebellum U.S. A New County-Level Dataset with Time, Region, and Direction Specific Freight Rates, 1820–1860

Online Appendix

Alexander Klein Peter Matthews

Part 1: Historical freight rates

Rivers Data on the freight rates on Mississippi River 1820-1859 are from Haites, Mak and Halton (1975), both upstream and downstream respectively. They also provide data on Barren Wabash, Tennessee, and Cumberland revers in 1840-1850. Hunter (1949) provides freight rates on Monongahela, Allegheny, Muskingum, Illinois, Missouri rivers in cents per ton mile in 1840-1850, but the calculations were done using railway distances. Therefore, we recalculated them using the length of the respective rivers. Freight rates for other than Mississippi River in the years before 1840 were interpolated using the growth of freight rates on the Mississippi River. Freight rates on the Hudson River come from Taylor (1957) and Binder (2011). Specifically, 1820 freight rates are proxied by 1816 freight rated from Taylor (1957), and the years between 1820 and 1853 were interpolated using 1820 freight rates and 1853 freight rates reported in Binder (2011). Freight rates in 1860 for the Mississippi River come from the government report on freight rates canals, rivers and railways by Newcomb and Ward (1901). The freight rates for the other rivers were calculated using the ratio of 1860 to 1850 freight rates on the Mississippi River and applied to the freight rates in 1850.

Lakes Data for the downstream and upstream freight rates respectively on the Great Lakes in 1816 and 1853 are for the three routes: Buffalo-Detroit, Detroit-Chicago, and Buffalo-Chicago. They come from Binder (2011). Linear interpolation was used to fill in the data in between the benchmark years. Freight rate on Chicago-Buffalo route in 1860 is from Newcomb and Ward (1901). The freight rates for the other two routes were calculated using the ratio of 1860 to 1850 on Chicago-Buffalo route applied to the freight rates in 1850.

Canals The Erie Canal freight rate data in 1830 are from von Gerstner (1842-3). Data for 1840 is the average in the years 1840-1849 calculated from Ransom (1967). Data for 1850 come from Annual Report of the State Engineer and Conveyor of the State of New York (1854), and for 1860 from Aldrich Report

(1893). Pennsylvania Canals freight rate data in 1830 and 1840 were calculated from Ransom (1967) who provides data on ton-miles, and from the Report on Canals (1883) which provide revenue data. Freight rates in 1850 are proxied by the freight rates in 1853 which are from Annual Report of the State Engineer and Conveyor of the State of New York (1854). Freight rate in 1860 is calculated using the ratio of Erie and Pennsylvania Canals freight rates 1850/1860 and applying it to the freight rate of the Pennsylvania Canals in 1850. Ohio Canal freight rate in 1830 is from von Gerstner (1842-3). Data in 1840 is the average of 1830 and 1850. Freight rates in 1850 are proxied by the freight rates in 1853 which are from Annual Report of the State Engineer and Conveyor of the State of New York (1854). Freight rate in 1860 is calculated using data on tonnage from Scheiber (1969) and revenue data provided in the Report on Canals (1883). Miami and Erie Canal freight rate in 1860 is, similarly to the Ohio Canal, calculated using data on tonnage from Scheiber (1969) and revenue data provided in the Report on Canals (1883). Delaware and Hudson freight rate in 1830 are proxied with 1839 freight rates calculated from von Gerstner (1842-3). We have adjusted the original data by subtracting the rail revenue and the volume of rail freight. Schuylkill River freight rate in 1830 is from von Gerstner (1842-3). Wabash and Erie Canal freight rate in 1850 is from Annual Report of the State Engineer and Conveyor of the State of New York (1854). Freight rate in 1860 is calculated using data on tonnage from Scheiber (1969) and revenue data provided in the Report on Canals (1883). Illinois Canals freight rate in 1850 is from Annual Report of the State Engineer and Conveyor of the State of New York (1854). Freight rate in 1860 is calculated using the ratio of Erie and Illinois Canals freight rates 1850/1860 and applying it to the freight rate of the Illinois Canals in 1850. Part 3 of Appendix, section 'Canal Freight Rates' describes the procedure of imputing freight rates for the canals for which we miss this information.

Coastal shipping Freight rate in 1816 is from Binder (2011) and in 1890 from Donaldson and Hornbeck (2016). Linear interpolation is used to fill in the missing data.

Wagon transport Freight rate in 1816 is from Binder (2011) and it is the arithmetic average of four routes: New York City–Philadelphia, Philadelphia–Pittsburgh, Albany–Buffalo, and La Salle, Il–Chicago. Since the freight rates differ only slightly (from 0.308 to 0.383), arithmetic average is a good proxy of weighted average which is infeasible to calculate due to the lack of wagon freight data. Freight rates in 1890 are from Donaldson and Hornbeck (2016). Linear interpolation is used to fill in the missing data.

Railways Feight rates in 1840 are from von Gerstner (1842-3) and Fishlow (1965). Altogether, they provide data for 34 railway lines, and they are listed in Table A1. Freight rates in 1850 are proxied by the freight rates in 1848 which are from Fishlow (1965). They are reported for eight regions/states in Table A2: New England, New York, New Jersey, Pennsylvania, Maryland, South, Southwest, and West. Freight rates for fifteen railway lines in 1860 are from Newcomb and Ward (1901). We used a nationwide map of railways in 1899 from Vance (1995) to allocate US counties into an area according to the dominant railroad company operating in that area. Map A1 presents these areas with the list of dominant railway company and Table A3 shows the freight rates assigned to each area and the congruence between the areas, railway company and freight rates. Part 3 of Appendix, section 'Railway Freight Rates' describes the procedure of estimating freight rates for the railway lines for which we miss this information.

Table A1: Historical Rail Freight Rates 1840, cent per ton-mile ($US2000)

Railway route	Freight rate
Boston & Lowell	0.42
Eastern RR	0.51
Nashua & Lowell	0.51
Nashua & Lowell RR	0.51
Boston & Worcester	0.65
Baltimore and Ohio	0.68
Boston & Maine RR	0.68
Baltimore & Susquehanna	0.86
Boston & Providence	0.86
Providence & Boston	0.86
Mohawk & Hudson	0.93
Rensselear & Saratoga	0.96
Long Island Railroad	0.98
Erie & Kalamazoo	1.03
Ithaca & Owego	1.03
Sarastoga & Schenectardy	1.08
Hudson & Berkshires	1.14
Palmyra & Jackson(burgh)	1.20
Detroit & Pontiac	1.25
Towanda	1.30
Portsmouth & Roanoke	1.37
Winchester & Potomac	1.37
Charleston Hamburg	1.42
Petersburg	1.42
Richmond & Petersburg	1.54
Louisa	1.57
Auburn & Syracuse	1.66
Georgia RR	1.71
Lockport & Niagara	1.71
Richmond, Fredericksburg & Potomac	1.71
Michigan Central	2.05
Madison & Indianapolis	2.57
Tallahassee	3.94
Vicksburg & Jackson	4.28

Source: von Gerstner (1842-43), Fishlow (1965)

Table A2: Historical Rail Freight Rates 1848, cent per ton-mile ($US2000)

Region	Freight rate
New England	0.8
New York	0.7
New Jersey	1.0
Pennsylvania	0.6
Maryland	0.5
South	0.8
Southwest	1.2
West	0.7

Source: Fishlow (1965), page 325

Table A3: Historical Rail Freight Rates 1860, cent per ton-mile ($US2000)

Map reference number	**Railroad company**	Freight rate
1	Chicago, Burlington and Quincy RR: East of Missouri	2.61
2	Southern Railway	5.37
4	Boston and Maine RR	4.63
4	New York, New Haven and Hartford RR	4.77
5	Baltimore and Ohio RR	1.79
5	Chesapeake and Ohio Railway	6.41
8	Erie RR	1.81
9	Illinois Central RR	2.04
12	New York Central and Hudson River RR	2.07
12	Lake Shore and Michigan Southern RR	2.16
16	Pennsylvania RR	1.96
17	Florida	3.55
6	Texas	8.37

Source: H.T. Newcomb, Edward G. Ward Jr: Changes in the Rates of Charge for Railway and Other Transportation Services,
U.S. Department of Agriculture, Washington, Government, Printing Office, 1901, Table 3, pp. 21-29.

Notes:

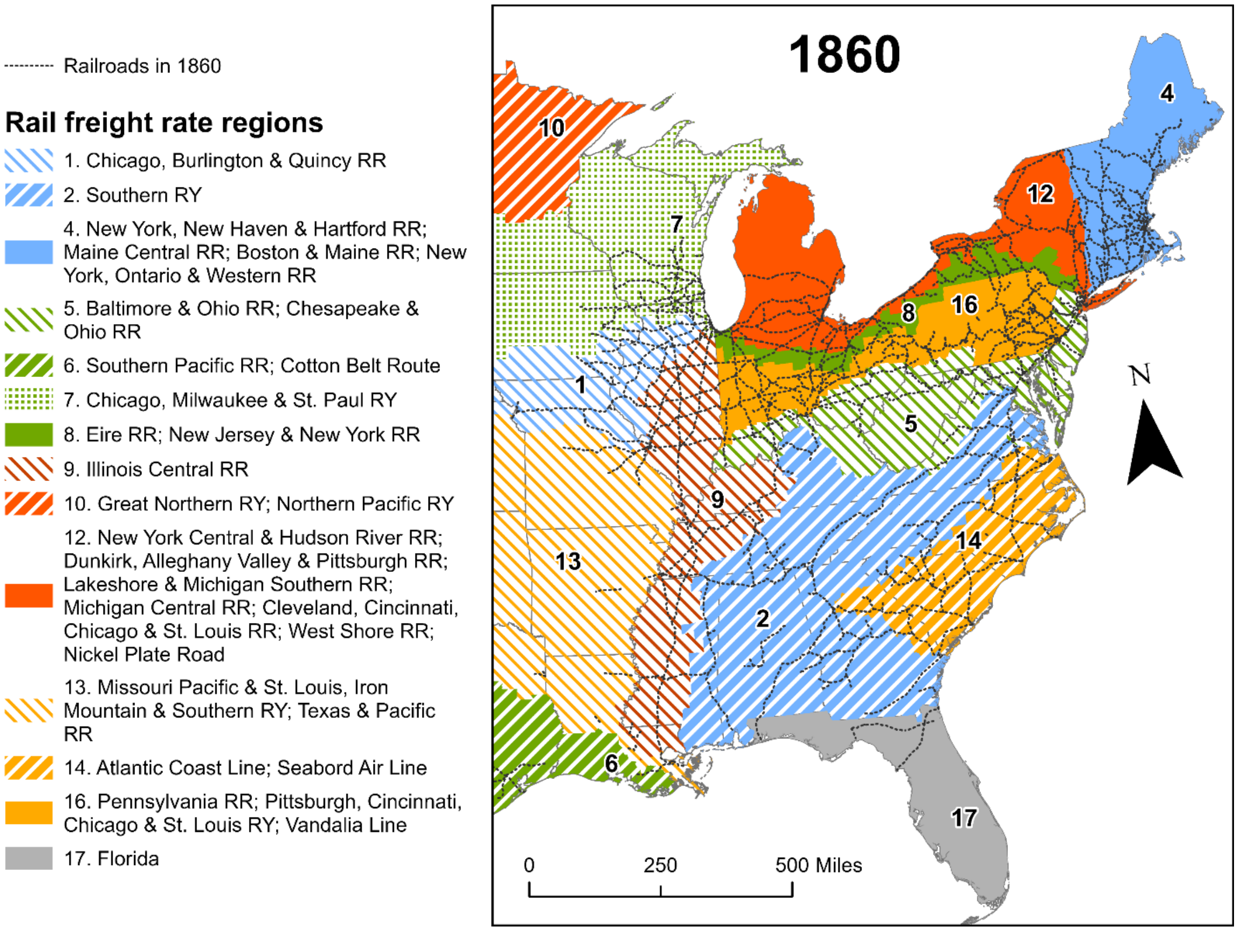
Chicago, Burlington and Quincy RR: East of Missouri - the freight rate is for 1866

Southern Railway: 1860 freight rate is proxied by the freight rate of Louisville and Nashville RR in 1866

Texas: Galveston and Houston and Henderson RR freight rate calculated using backward projection from 1870 using the changes in freight rate between 1860 and 1870 of Georgia RR in 1866 / Southern Carolina and Georgia RR in 1870

Florida: proxied by South Carolina and Georgia RR in 1866. ‘Map reference number’ refers to the region on Map A1.

Map A1: Rail lines in 1860



Source: Atack (2016), Vance (1995)

Part 2: Additional Tables

Table A4: Number of Ships and Tonnage in Operation on Western Rivers 1811-1860.

Year	Ships in Operation (number)	Tonnage (tons)
1811	1	371
1820	69	14,208
1830	151	24,574
1840	494	82,626
1850	638	134,566
1860	817	195,022
1868	874	212,203

Source: Haites, Mak, Walton (1975), Table B-1, page 130.

Table A5: Freight Rates on Cargo and Fares for Deck and Cabin Passengers from Louisville to New Orleans 1810-1860 (in current $US).

	Freight rates ($US per ton)		Passenger fares ($US per person)			
	Upstream	Downstream	Upstream		Downstream	
			Cabin	Deck	Cabin	Deck
Before 1820	100	20	125	25	75	18
1820-29	20	12.5	50	10	25	6
1830-39	10	10	25	6	25	6
1840-49	5	6	20	4	20	4
1850-59	5	6.5	15	3	15	3

Source: Haites, Mak, Walton (1975), Table 7, page 32.
Note: the original freight rates were in current $US per 100 pounds and were recalculate to tons.
The freight rates from 1840 are for flour, pork, whiskey, and pound freight. The rates were calculated as the weighted average of monthly rates with the number of steamboat departures as weights. Details of the sources are discussed in Haites, Mak, Walton (1975), pp. 150-157.

Table A6: Freight Rates between Louisville and Selected Destinations, 1840-1850.

Route	River	Category	Cents per ton mile
Louisville to:			
Bowling green	Barren	Tributary	1.94
Terre Haute	Wabash	Tributary	1.35
Florence	Tennessee	Tributary	1.73
Nashville	Cumberland	Tributary	1.38
New Orleans	Mississippi	Trunk	0.44

Source: Haites, Mak, Walton (1975), Table 13, page 55.

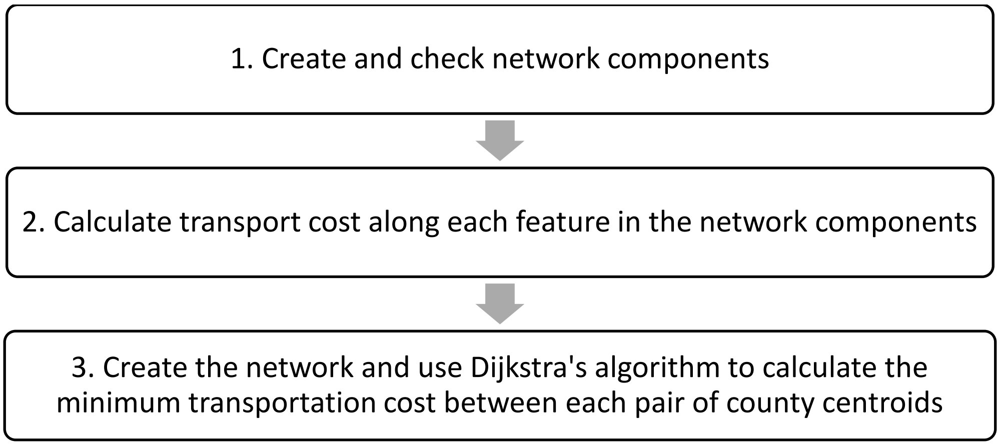
Table A7: Decomposition of the changes in the average county-pair transport costs ($US 2000) between 1820 and 1860.

	Transport Costs Δ 1820-1860					Decomposition			
	From a county in the US region to the rest of the US	Year 1820	Year 1860	1820 transp infrastructure with 1860 freight rates	Δ 1820-1860	Transport system effect	Freight rate effect	Transport system effect %	Freight rate effect %
(1)	United States: consists of the regions listed below	19.8	5.4	13.1	-14.4	-7.7	-6.7	53.5%	46.5%
(2)	New England	18.5	4.4	11.4	-14.1	-7.0	-7.1	49.7%	50.3%
(3)	Middle Atlantic	17.6	3.7	11.3	-14.0	-7.6	-6.4	54.4%	45.6%
(4)	East North Central	17.6	3.7	11.5	-13.9	-7.8	-6.1	56.1%	43.9%
(5)	South Atlantic	18.3	4.8	11.7	-13.6	-6.9	-6.6	51.0%	49.0%
(6)	East South Central	16.8	4.2	11.1	-12.7	-7.0	-5.7	55.1%	44.9%

Source: see the text for details on the data sources

Part 3: Network construction and analysis – methodology

Costs of transportation between counties were determined by using ArcGIS Pro (ESRI, 2023) to construct a network dataset comprising the road, rail, river, canal, and sea/lake routes across the country, and then employing Dijkstra's algorithm to calculate the cost associated with the cheapest route across this network for each pair of counties. The network construction and analysis followed the methodology described by Donaldson & Hornbeck (2016), with some modifications to allow us to incorporate temporal and spatial variation in freight rates.



1. Creating and checking network components

1.1. GIS network components

The network was constructed following the design specified by Donaldson & Hornbeck (2016), with river, canal, and railroad components replaced with those provided by Atack (2021).

Table S1. Description of each GIS feature layer used as a component in the transport network created for each year / scenario, and the formulae for calculating transportation costs for each component. This table is a modified version of that provided in the supplementary material from Donaldson & Hornbeck (2016).

ID	Component Name	Component Definition	Construction description	Cost ($)
0	Navigable rivers	Fogel's definition of navigable rivers; time invariant component of network	Sourced from Atack (2021). Maps for each decade were created by deleting river sections that were not navigable during the year in question.	[Freight rate]*[Length]/1609.344[39]
1	Constructed canals	Fogel's definition of navigable canals; time invariant component of network	Sourced from Atack (2021). Maps for each decade were created by deleting canals that were not in use during the year in question.	[Freight rate]*[Length]/1609.344
4	Sea/Lake Routes	Multiple point-to-point connections throughout the Great Lakes and the Oceans; time invariant component of network	Sourced from Donaldson & Hornbeck (2016), who created this component manually to effectively saturate area.	[Freight rate]*[Length]/1609.344
5	Railroad harbor	Points where transshipment to Sea/Lake Routes is considered possible, created whenever the 1911 railroad network approaches the coastline; time invariant component of network	Sourced from Donaldson & Hornbeck (2016), who created this component manually as a short line from Sea/Lake Routes to Railroads.	[Transshipment cost]
6	Railroads	Railroad lines as depicted on maps; time variant component of the network	Sourced from Atack (2021). Maps for previous decades were created by deleting railroad lines that do not appear during the year in question.	[Freight rate]*[Length] /1609.344
7	Sarnia-Buffalo railroad line	Created based on 1860 Stratford's map from 1860	Sourced from Donaldson & Hornbeck (2015), who created this component manually.	[Freight rate]*[Length] /1609.344
8	Wagon Routes (Centroid-to-centroid)	Wagon routes connecting any two centroids within 300km of each other; time invariant component of the network	Created automatically in GIS.	[Freight rate]*[Length] /1609.344
9	Sea route Between Coasts	Direct route connecting the West Coast (near San Diego) to the East coast (in the Gulf of Mexico); time invariant component of network	Sourced from Donaldson & Hornbeck (2015), who created this component manually.	1.48
130	In-county centroid-to-railroad connection	Represents the average wagon route from any point in the county to railroad lines that pass through the county; a transshipment cost is then	Created manually to connect a centroid to railroads within the county	[Transshipment cost] + ([Freight rate]*[Mean_Length] /1609.344)

[39] For reasons of conformity with other parts all lengths of network components were measured in meters (e.g., output of the Near tool was given in meters); we divide all lengths by 1609.344 to convert them to miles, so that we can generate cost values in $ per ton-mile.

		incurred; time variant component of network		
140	Out-of-county centroid-to-railroad connection	Represents the average wagon route from any point in the county to relevant railroad lines outside the county border in various directions; a transshipment cost is then incurred; time invariant component of network	Created manually to connect a centroid to potentially-relevant railroads outside the county	[Transshipment cost] + ([Freight rate]*[Length] /1609.344)
15	Out-of-county centroid-to-harbor connection	Represents the average wagon route from any point in the county to Sea/Lake harbors (ID 5) outside the county in various directions; a transshipment cost is then incurred; time invariant component of network	Created manually to connect a centroid to potentially-relevant harbors outside the county	[Transshipment cost] + ([Freight rate]*[Length] /1609.344)
60	River harbor	Points where transshipment to Sea/Lake Routes is considered possible; created wherever rivers, canals, or proposed canals meet the coastline; time invariant component of network	Created manually as a short line from a Sea/Lake route	[Transshipment cost]
61	Canal harbor			
601	River-to-Canal transshipment station	Points where rivers and canals meet and a transshipment is possible; time invariant component of network	Created manually as a short line to connect the two modes of transportation	[Transshipment cost]
700	River- /Canal- to-Railroad Transshipment point	Points where inland waterways (rivers, canals, proposed canals) meet railroads and transshipment is possible; time invariant component of network		
80	In-county centroid-to-river connection	Represents the average wagon route from any point in the county to waterway lines that pass through the county (river, canal, or proposed canal); a transshipment cost is then incurred; time invariant component of the network	Created manually to connect a centroid to waterways within the county	[Transshipment cost] + ([Freight rate]*[Mean_Length] /1609.344)
81	In-county centroid-to-canal connection			
90	Out-of-county centroid-to-river connection	Represents the average wagon route from any point in the county relevant waterway lines outside the county border in various directions (river, canal, or proposed canal); a transshipment cost is then incurred; time invariant component of the network	Created manually to connect a centroid to potentially-relevant waterways outside the county	[Transshipment cost] + ([Freight rate]*[Length] /1609.344)
91	Out-of-county centroid-to-canal connection			
N/A	County Borders	1990 County borders	Downloaded from nhgis.org	N/A
N/A	County Centroids	The geographical center (centroid) of each 1990 county; the county centroid is considered its "representative point"	Created using ArcGIS "*Feature to point"* tool	N/A

1.2. Incorporating variation freight rates

Our analysis was designed to reflect variation in freight rates across both time and space. For temporal variation, this meant assigning different values to the freight rate attribute for each feature of a network component each year. For spatial variation, this meant dividing up the features in each network component into specific regions (e.g., sections of river, or canals / railroads operated by different companies), and assigning freight rates based on the region that the feature was located in. Further details on how we included spatial variation in freight rates are given below.

1.2.1. River freight rates

For years prior to 1860, we used freight rates that varied between rivers (but from 1860 onwards, river freight rates only varied over time, and not in space). Some navigable rivers were assigned known freight rates using historical sources, but were data on freight rates were unavailable, we estimated rates by assuming that they would be proportional to average discharge at river mouth.

River discharge was plotted against freight rate for those rivers where historic data was available to get an idea of the relationship between discharge (in m^3/s), and therefore identify an approximate formula for estimating river freight rates in each year (Table S2).

Table S2. Relationship between river freight rate and average annual discharge (in m^3/s) at mouth for 1820 – 1850. (For subsequent years, freight rates were the same across all navigable rivers).

Year	River Freight Rate (in 2000 $US)
1820	$\frac{1}{discharge}$ * 4; maximum value $0.016
1830	$\frac{1}{discharge}$ * 3.5; maximum value $0.014
1840	$\frac{1}{discharge}$ * 2.25; maximum value $0.009
1850	$\frac{1}{discharge}$ * 2.05; maximum value $0.008

River discharge estimates were sourced primarily from the US Geological Survey's National Water Information System Web Interface and factsheets (US Geological Survey, 1992, 2023). For smaller tributaries, with inaccessible or incomplete data, the freight rate was assumed to be the maximum for the decade (Table S2).

1.2.2. Canal freight rates

We used freight rates that varied across canals. We were able to use historical sources to assign freight rates to some canals for each decade and use this information to create different regions containing networks of canals centered around each of the canals with known freight rates. Canals with unknown freight rates were assigned a freight rate based on which of these regions that they were located within.

1.2.3. Sea / lake freight rates

The sea / lake route features were categorized into three regions (Erie; Huron & Michigan; sea / coastal). Each of these regions was associated with a different freight rate: features were assigned a freight rate based on the region they were located within.

1.2.4. Railway freight rates

For all years of our analysis in which railway were used as a network component (i.e., 1840 onwards), each railway feature had an attribute detailing a regional / company-specific freight rate.

Railways in 1840

Historic sources provided average freight rates for some of the railways in 1840. For those railways for which we did not have data on freight rates, we estimated freight rates by using one of three methods, depending on the location and configuration of the unassigned railroads:

1. For railway sections within longer continuous lines.
   Where there were railways with unknown freight rates situated in between other railways with known freight rates, we calculated the average of the freight rates for the two railways at either end of the unassigned railway. The lengths of the two railways with known freight rates were used to weight this average, and this weighted average was used as the freight rate for the railway.

2. For railway sections within complex networks / clusters of lines.
   Where there were railway with unknown freight rates situated within complex networks or clusters of other railways, with multiple connections and branches, we calculated the average freight rate for all companies operating in the region (for which freight rate data was available) and weighted this average by the lengths of railway operated by each company. This weighted average was then used as the freight rate for all railways within the network / cluster for which no freight rate data was available.

3. For isolated / remote sections of railway.
   Where there were isolated sections of railways with unknown freight rates, that did not obviously belong to any local cluster or network of railways, we used the freight rate from the nearest railway section with a known freight rate.

Railways in 1850

Based on a map of railways in 1848 presented in Fishlow (1965), we created a new GIS layer that divided the country into eight regions and associated a different freight rate with each region. We used the Pairwise

Intersect tool to split railway features where they crossed the boundaries of these regions, and performed a Spatial Join to give each railway feature the freight rate of the region it was located within.

Railways in 1860

We used a nationwide map of railways in 1899 from Vance (1995) to allocate all US counties into regions according to the dominant railroad company / category operating in the area. We created a new GIS layer to represent the different regions, and then used the Pairwise Intersect tool to split railway features where they crossed the boundaries of these regions, and then performed a Spatial Join to give each railway feature the freight rate of the region it was located within.

1.3. Setting up and checking network components

Components 0, 1, 4, and 6 (i.e., rivers, canals, sea/lake routes, and railway) were split wherever they intersected with other components / with themselves, to ensure that the Network Analysis tools (described in Section 3) recognizes that turns are permitted at these locations. This was achieved using the following process:

1. Use the Pairwise Intersect tool to create points at all junctions with other relevant network components / and the component itself.
   E.g., for railroads, this means using the Pairwise Intersect tool to create points at intersections with the following components:
   a. Railways
   b. Railway harbor
   c. Sarnia-buffalo line
   d. In-county centroid-to-railway connection
   e. Out-of-county centroid-to-railway connection
   f. River- /Canal- to-Railrway Transshipment point
2. Merge all the point layers created with the Pairwise Intersect tool, and then use the Split Line at Point tool (with a 5-meter search radius to ensure that all intersection points are used) to split the features at every intersection.

Since components 130, 140, and 700 are time-invariant components of the network, in a certain year, they may include many harmless extra connections that do not connect the county centroids to any other network components. These extra connections correspond to all the potential connections to the transportation network in past or future decades.

Including all the connections across all decades for the network for each year ensures that it is never the case that a connection exists in one decade and, although it should, does not exist in the next/previous decades.

We applied consistency checks on the network components, as described by Donaldson & Hornbeck (2015), to minimize the probability of a network construction error.

2. Transport cost calculations

As shown in Table S1, the cost of transporting freight along a given feature of a network component was calculated by multiplying the network component's freight rate (in $ per ton-mile in 2000 $US) by the length of the feature (in meters) and dividing by 1609.344 to convert distances to miles.

We therefore added a 'Length' field to each component and used the Calculate Geometry tool to assign a length to each feature in meters. We created a 'Cost' field and used the Calculate Field tool to assign values for this field to each feature using the formulae in Table S1.

For components 130, 80, and 81, the transport cost formulae are based on [Mean_Length] rather than [Length]. [Mean_Length] represents the average distance from a random point within a county to the nearest waterway or railway.

This [Mean_Length] is calculated as follows:

1. Run the Create Random Points tool with:
    a. the Constraining Feature Class set as the county boundaries,
    b. and Number of Points = 200.
2. Save a copy of the new random points per county layer (because the following steps will permanently alter the layer).
3. Use the Near tool to calculate the distance from each of the random points to the nearest river/canal/railroad.
4. Use the Dissolve tool on the random points layer, with:
    a. Dissolve Field = unique county IDs from the original county boundaries layer.
    b. Statistics Field = NEAR_DIST (Statistic Type = Mean)
5. Join the dissolved random points layer to the In-county centroid-to-feature layer, so that each in-county centroid-to-feature line has a MEAN_NEAR_DIST value associated with it. Save as a new layer.
6. Create a Cost field in the in-county centroid-to-feature layer and use the Calculate Field tool to assign values for this field to each feature using the transport cost formulae in Table S1.
7. Steps 2 to 7 must be repeated for canals/rivers/ railways for each decade.

3. Network creation and calculation of transportation costs between counties

Once all network components have been created, checked, and a transport cost has been assigned to each component feature, we can use them to build the network for a given year or scenario, and then calculate the minimum transport cost between each pair of counties. This process proceeds as follows:

1. Within a Geodatabase, create a Feature Dataset containing all the required network components for the scenario.
2. Use the Create Network Dataset tool:
    a. Under Target Feature Dataset, select the Feature Dataset from step 1.
    b. Select all the numbered network components as Source Feature Classes.
    c. Run the tool.
3. Open the Properties for the new Network Dataset and go to Travel Attributes -> Costs. Create a new Cost attribute for the network dataset:
    a. Set Units = Other, and Data Type = double
    b. Under Evaluators -> Edges, set the Type for Edges in the (Along) direction as 'Field Script', and leave the Type in the (Against) direction as 'Same as Along'.
    [Unless the network includes rivers where the upstream and downstream transport costs differ, in which case the Type in the (Against) direction should also be set as 'Field Script']
    c. Create the script expression for each edge / network component by selecting the Value option, setting the Language to Python, and setting Result = !Cost! (i.e., the name of the field containing the transport cost values for each feature)
4. Still in the Network Travel Attributes, create a new Travel Mode, where the Impedance is set as equal to the new cost attribute created in step 3.
5. Use the Build Network tool to build the network.
6. Use the Origin-Destination Cost Matrix tool to use Dijkstra's algorithm to calculate travel costs between each pair of county centroids:
    a. For the Import Origins and Import Destinations, use the US county 1990 centroids as the Input Locations.
    b. Ensure that Mode is set as the new travel mode created in step 4, and Output Geometry can be set as 'No Lines', and run the tool.
    c. When the analysis is complete, use the Export Table tool (setting the Input Table as the OD Cost Matrix\Lines layer) to export the results as a .txt file.

References

Atack, J. (2021). *Historical Geographic Information Systems (GIS) database of U.S. Railroads for 1850 - 1911*. https://my.vanderbilt.edu/jeremyatack/data-downloads/

Donaldson, D., & Hornbeck, R. (2016). Railroads and American Economic Growth: A "Market Access" Approach. *The Quarterly Journal of Economics*, *131*(2), 799–858. https://doi.org/10.1093/QJE/QJW002

ESRI. (2023). *ArcGIS Pro (Version 3.2)*. https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview

Fishlow, A. (1965). *American railroads and the transformation of the ante-bellum economy*. Harvard University Press.

US Geological Survey. (1992). *Largest Rivers in the United States*. https://pubs.usgs.gov/of/1987/ofr87-242/

US Geological Survey. (2023). *National Water Information System: Web Interface*. https://waterdata.usgs.gov/ak/nwis/nwis

Vance, J. E. (1995). *The North American railroad: its origin, evolution, and geography*. Johns Hopkins University Press.