Network Structure and Efficiency Gain from Mergers: Evidence from the U.S. Freight Railroads *

Yanyou Chen†

This Version: *February 15, 2021*

Abstract

This paper studies the role of network structure in spatial competition. Markets are interdependent in a network, hence, the network structure (topology of network) affects how much firms benefit from economy of scope. This paper studies the role of networks in affecting the effect of mergers by examining mergers that took place from 1985 to 2010 in the U.S. freight railroad industry. I estimate an optimal transport network model that features firms' pricing, routing, and allocation decisions in multiple origin-destination markets. My model allows origin-destination markets to be interdependent in the cost minimization stage, and the model has rich implications in understanding the change of shipment prices after mergers. On average, shipment price decreased by 9% following a merger, and the price decreased by 11% where railcars no longer had to be switched between two companies as a result of a merger. Results also suggest that if the betweenness centrality of a station is larger, the reduction of cost is larger after the merger. Throughout the period studied, the average degree centrality increased and the average betweenness centrality decreased, at the firm level. This indicates that mergers of railroad firms have generated a lower reduction of cost and a higher increase of mark-up in recent years.

Keywords: Merger, Cost Efficiency, Network Structure, Transport Network, Railroad

*I am indebted to my committee members, Allan Collard-Wexler, Daniel Yi Xu, James W. Roberts, Christopher D. Timmins, and Bryan Bollinger for valuable comments and encouragement throughout this research project. I also thank participants in the Railway Management Certificate Program for their valuable insights on the industry. Any remaining errors are mine alone. Disclaimer: Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Surface Transportation Board. All results have been reviewed to ensure that no confidential information is disclosed.

†Department of Economics, University of Toronto, yanyou.chen@utoronto.ca.

1 Introduction

In the past thirty years, a small number of firms have gained a very large share of the market in the United States, following mergers and consolidations. Autor et al. (2017) document an upward trend over time that industries have become more concentrated on average. There is a large strand of literature documenting and analyzing mergers in multiple industries.[1] However, an open question is that are the efficiency gains of mergers small, large, or absent? In theory, on the one hand a merger gives the combined firm greater market power hence incentive to increase prices. On the other hand, a merger may generate efficiencies, reduce marginal costs, and give the combined firm an incentive to lower prices. Economists are aware of this trade-off since Williamson (1968), yet there is very little direct empirical evidence showing efficiency gain of mergers, and if it offsets the incentive to raise prices. This is largely because it is difficult to measure and quantify if mergers lower the marginal cost of production of the combined firm. The objective of this paper is to study the mechanism of cost efficiency after mergers in U.S. railroads, and evaluate if the efficiency gain offsets the incentive to raise prices. Moreover, I contribute to the merger analysis by proposing a novel way to study efficiency gain at a network industry where origin-destination markets are interdependent.

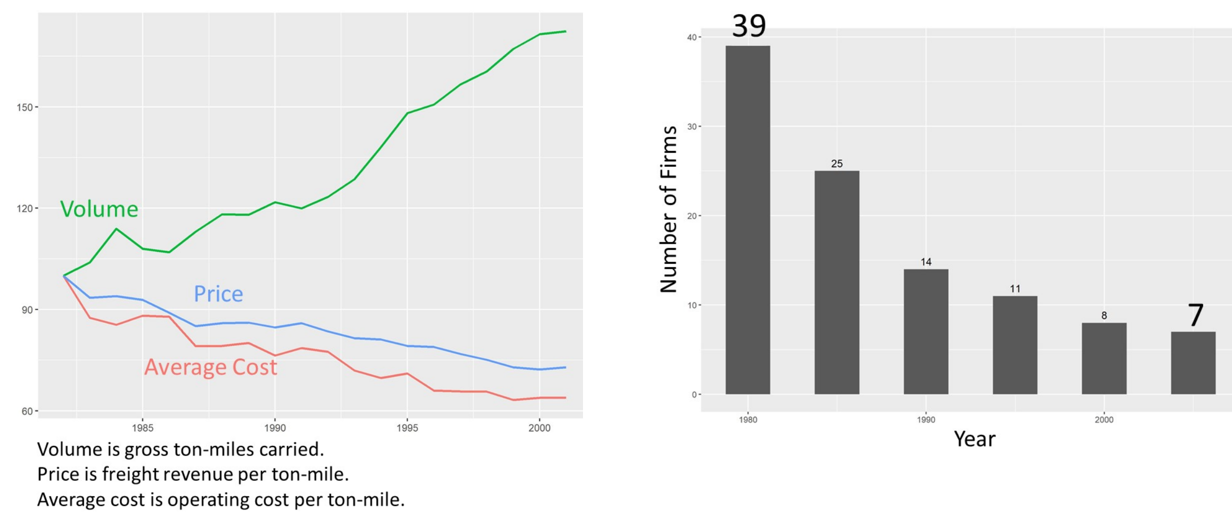
I do this by looking at the U.S. freight railroad industry. First, there was a series of merger from 1980 to 2005. The number of Class I railroads[2] dropped from 39 to 7, and market share of the top four firms increased from 66% to 94%. Although concentration has gone up in this industry, prices have decreased steadily. As illustrated in the left panel in Figure 1, from 1980 to 2005, prices per shipment have decreased by 20%, while the total volume of shipment has doubled. Given there is limited technological change in the studied period, the price reduction indicates that there might be efficiency gains following these railroad mergers. According to the Department of Transportation, in 2017, railroads are the second largest transport mode in providing freight service in the United States, carrying 1,675 billion ton-miles of freight, and accounts for around 30% of total freight transportation. According to the Association of American Railroads (AAR), in 2016 the railroads

[1] In the airline industry, a wave of consolidation happened during a short period of time in the late 1980s (Peters, 2006); in the telecommunication industry, over 6,000 acquisitions occurred between 1996 and 2006 (Leeper, 1999; Jeziorski, 2014); a dozen global hard-disk manufacturers consolidated into only three in the last 20 years from 1996 to 2015 (Igami, 2017; Igami and Uetake, 2016); 190 hospital mergers occurred between 1989 and 1996 (Dafny, 2009; Bazzoli et al., 2002; Dranove and Lindrooth, 2003), and in the dialysis industry, more than 1,200 acquisitions occurred between 1998 and 2010 (Eliason et al., 2018).

[2] Class I railroads are defined as "having annual carrier operating revenues of $250 million or more in 1991 dollars." According to AAR, Class I railroads account for more than 95% of revenues generated in the U.S. freight railroad industry in 2016.

transported about 40% of intercity ton-miles, more than any other mode of transportation.[3] Unlike many other industries, metrics of efficiency can be directly observed in railroad transport. At individual route level, the combined firm can abandon the redundant rail lines that serve the same origin and destination market. Also, the combined firm can eliminate the interchange cost where railcars need to be switched between two railroad companies before the merger. At network level, the combined firm can consolidate traffics and choose shorter efficiency-weighted routes that are not available before the merger.[4]

Figure 1: U.S. Freight Railroad Performance and Number of Class I Railroads



Source: AAR, Analysis of Class I Railroads Annual Issues

I use detailed shipment data on 12 million waybills to quantify the efficiency gain of mergers at both individual route level and network level. Railroad companies conduct point-on-point pricing, and the waybill data contains price information at firm-origin-destination level. The waybill data also contain detailed shipment information on such attributes as commodities carried, total billed weight, participating railroads, and origin, destination locations of each load. I begin my analysis by examining the price effect of mergers. By conducting

[3]One evidence that shows limited technological change in the studied time period is that until the late 1990s, the vast majority of rail lines in the U.S. still relies on the human crew for complying with all safety rules. One accident that highlights this fact is the Big Bayou Canot rail accident, where a tow of heavy barges collided with the rail bridge and later caused the derailing of an Amtrak train on the CSXT Big Bayou Canot bridge with 47 people killed and 103 more injured. The accident happened partly because no technology was implemented to monitor the real-time condition of tracks and train movements, and the Amtrak engineer was not notified of the collision in time. Positive Train Control system that monitors and controls train movements was required by rail safety law in 2008, but full implementation is still underway and is supposed to be done by the end of 2020.

[4]According to Kwoka and White (1997), efficiency gain of mergers include alleviation of capacity constraints, integration of track networks, reduction of mileage along major routes, and improved utilization of locomotives and railcars.

reduced-form analysis at individual route level for each origin-destination market, I study how prices change before and after mergers. After controlling for observable characteristics and a rich set of fixed effects, I find that on average shipment prices have decreased by 9% after mergers. Then I open up the merger cases and examine the price effects for different route types. For route where railroad companies exchange railcars before the merger, the price effect of merger is 11%. For other types of route, the price effect of merger is only 6%. The results suggest that efficiency gain of mergers vary by route types, and elimination of interchange cost where railroad firms exchange railcars is an important source of cost efficiency. One concern regarding these price effects is that such price effects might be driven by competition from other transportation modes, instead of mergers. I address this concern by showing that these price changes occur both for commodities that are mostly transported by rail such as coal, and commodities that are also largely transported by other modes such as electrical machinery. The results suggest that the price effects are not driven by competition from other transportation modes.

I then provide motivating evidence to show that there is efficiency gain of mergers. I do this by examining changes of aggregate operational statistics at firm level, and also examine changes of operations at individual route level for each origin-destination market. At firm level, after controlling for observable characteristics, the reduced-form results show that on average the ratio of switching hours to road service hours has decreased by 11.5%, and length of haul has increased by 5.3% after railroad mergers. The results indicate that there is efficiency gain of mergers because the less time a train spends switching in yards the more efficient the operation is, and it is more fuel efficient for a train to run a longer distance. Results at individual route level also show that there is efficiency gain of mergers. By comparing the changes before and after mergers, I show that the number of interchanges has decreased following railroad mergers. Moreover, firms utilize more unit trains after mergers. A unit train is a train in which all railcars carry the same commodity and are shipped from the same origin to the same destination. The use of unit trains saves time and money. The combined firm can reoptimize routing and consolidate traffics to initiate more unit trains after the merger.

One challenge in quantifying merger effects in the freight railroad industry, however, is that origin-destination markets in a railroad network are interdependent. The shipment cost from one origin to destination depends on the routing and investment decision, which depends on the location and demand of other markets in the railroad network. Therefore, looking only at changes at individual route level is insufficient for understanding mergers in this industry. To overcome this problem, I estimate a structural model that features firms pricing, routing, and investment decisions in multiple origin-destination markets based on the framework of Galichon (2016) and Fajgelbaum and Schaal (2020).

With the empirical evidence in mind, I propose an optimal transport network model that allows firms to choose pricing, routing, and investment for thousands of origin-destination markets. In the model, firms choose prices for the markets they serve, conditional on their perceived cost of operating there and the prices of competitors. Given the prices, the transportation service demanded for each firm in each origin-destination market is determined. Then firms jointly decide how to allocate infrastructure investment and make routing decisions in their network by minimizing the operational cost. Intuitively, firms allocate more resources to routes that carry larger volume of traffic, and choose the shortest efficiency-weighted routing to serve each origin-destination market. I model demand of transportation service using a nested logit model. Shippers choose between different railroad companies and an outside option which is trucking. Firms compete in prices, and in equilibrium, the perceived cost of operating at the pricing stage is consistent with the outcome from firms cost minimization problem. I consider the interdependence of markets in modeling routing and investment. Railroad firms can consolidate infrastructure investment into a small number of corridors, and then route traffic from multiple origins and destinations over these efficient corridors to reduce shipment cost. This allows operational cost to be dependent on topology of the network. For example, if origins and destinations locate near hubs, firms can consolidate traffics into a small number of routes through these hubs. If origins and destination locate in disperse regions, it is difficult to consolidate traffics from different origins and destinations into few efficient routes. I estimate the model using generalized method of moments, and the data moments are obtained from the detailed shipment data.

I then use the estimated parameters to conduct counterfactual simulations. Results of counterfactual simulations show that shipment cost decreases by 32% after mergers, which leads to a model predicted price decreases of 10%. The counterfactual simulations also decompose sources of cost efficiency into elimination of interchange cost, reallocation of resources for investment, and re-optimization of routing. The results suggest that elimination of interchange cost is the predominant source, which accounts for 44% of the cost reduction. Network efficiency that comes from routing accounts for 36% of the cost reduction. In general, the level of cost efficiency from a merger depends on the number of interconnecting services provided by the two merging parties, and it also depends on how traffic can be consolidated between the two merging parties. Last, I use the model to examine the trade-off between efficiency gains and increased market concentration, and study the implied welfare changes in different geographic markets. On the one hand, the degree of overlap between the two networks determine the incremental of market power, because the number of firms providing freight service decreases in the overlapped regions. On the other hand, the topology of the network and distribution of demand affects how the traffics can be consolidated, hence affect the degree of efficiency gain. By combining these two forces, I show that there is a large degree of heterogeneity of welfare changes in different geographic markets, depending upon the location of each market within the network.

Contributions to the Literature. This paper adds to the literature on the effects of horizontal mergers by providing evidence of efficiency gain after mergers and uncovering the sources. First, the price effects of mergers are extensively studied in the literature. Borenstein (1990) and Kim and Singal (1993) analyze the effects of airline mergers, and Kim and Singal (1993) find that airfares increase by 9.44% on airline routes served by the merging firms. Dafny (2009) evaluates the impact of independent hospital mergers between 1989 and 1996 and finds evidence of sharp increases in rivals' prices of about 40% following a merger. In comparison, my paper finds that on average price decreases by 9% after a railroad merger, and price decreases by 11% for interconnecting routes. By contrast, the literature on the cost efficiency of mergers is sparse. One good exception is Ashenfelter et al. (2015), the authors use panel scanner data and geographic variation in the U.S. beer industry, and find that the average predicted increase in concentration leads to price increases of 2% but that this is offset at the mean by a nearly equal and opposite efficiency effect. Some other paper includes McGuckin and Nguyen (1995), which use an unbalanced panel of plants from the Longitudinal Research Database to examine the relation between ownership changes and productivity growth, and they find that plant productivity growth is positively related to a change in ownership.[5] My paper contributes to the literature on the cost efficiency by not only providing reduced-form evidence on improvement of metrics of efficiency and price reduction after mergers, but also uncovering the mechanism of efficiency gain and how it changes with features of the merger through a structural model.

This paper also contributes to the literature that studies the freight railroad industry. The existing literature on the freight railroad industry examines multiple aspects of the effect of deregulation and mergers. Casavant et al. (2012) study the rail rate structure for agricultural commodities and compare it with rates for other commodities. Friebel, McCullough and Angulo (2014) investigate the restructuring of the U.S. freight railroad after deregulation and document both network reductions (the abandonment of redundant rail lines) and labor downsizing after mergers. Prater et al. (2012) examine the sufficiency of

[5]Given the fact that direct metrics of efficiency is hard to define and measure, some other literature calculates cost savings through merger simulations. Jeziorski (2014) uses a dynamic oligopoly model to estimate the cost savings resulting from mergers in the U.S. radio industry, with the estimated resulting savings amounting to $1.2 billion per year. Pesendorfer (2003) examines mergers in the U.S. paper and paperboard industry, comparing equilibrium investment decisions before and after the merger wave and finding that total welfare increases by $583.5 million as a result of the mergers.

rail freight competition and the effects of intramodal competition on rail rates. Chapin and Schmidt (1999) find that merged firms are larger than efficient scale. McCullough (2005) documents that changes in output composition along with line abandonment and a significant degree of industry consolidation lead to longer haul lengths and higher traffic densities. Event studies on special merger cases in the literature also provide explanations on the origination of merger efficiency. Kwoka and White (1997) study the Union Pacific–Southern Pacific merger, and Pittman (2009) examines the Santa Fe and Southern Pacific merger proposal. However, most of the literature studies change of some aggregate cost or price index, or examine merger effect by looking at individual markets in a regression. Virtually no research looks at merger effect by considering the interdependent nature of railroad networks. My paper contributes to filling in this gap by considering the interaction of cost efficiency and incentives to raise prices in railroad network, and understanding welfare implications in different geographic markets.

In this paper, I jointly model firms pricing, investment, and routing decisions by employing an optimal transport network method, which is similar to the method used in Fajgelbaum and Schaal (2020). I differ from their paper by adding competition and routing choices of each railroad company. My research is also broadly related to other papers that use spatial analysis in understanding distributional impact of economic activities, such as Buchholz (2015), Donaldson (2018), and Brancaccio, Kalouptsidi and Papageorgiou (2017).[6] In both Buchholz (2015) and Brancaccio, Kalouptsidi and Papageorgiou (2017), the network interdependence comes from the fact that once I drop somebody off, I end up in the location where I drop them off for future ride, which is more limited than the interdependencies in routing in railroads.

The remainder of the paper is organized as follows. Section 2 describes the industry background and explains sources of cost efficiency. Section 3 outlines the three datasets used in the paper. Section 4 provides reduced-form evidence on how shipment prices and metrics of efficiency change after railroad mergers. Section 5 constructs the structural model of firm pricing, routing and investment decisions in a rail network. Section 6 shows the estimation results and assesses the validity of the model. Section 7 shows the counterfactual simulations that decompose the sources of cost efficiency, and explains why there is tremendous heterogeneity of welfare changes after mergers in different geographic markets. Section 8 concludes.

[6] Buchholz (2015) analyzes the dynamic spatial equilibrium of taxicabs and shows how common taxi regulations lead to substantial inefficiencies; Donaldson (2018) uses archival data from colonial India to investigate the impact of India's vast railroad network; Fajgelbaum and Schaal (2020) study optimal transport networks in spatial equilibrium; and Brancaccio, Kalouptsidi and Papageorgiou (2017) use detailed data on vessel movements and shipping contracts to study world trade costs and trade flows.

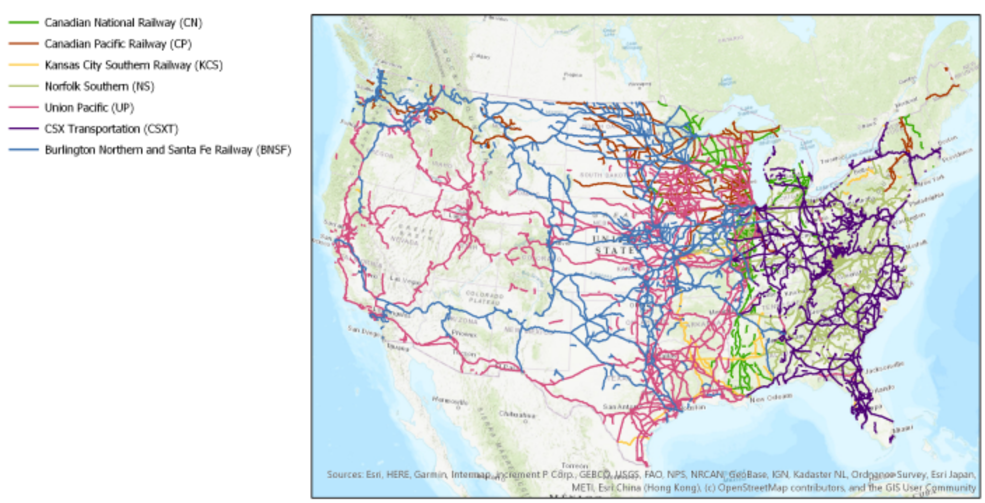
2 Industry Background

2.1 Deregulation and Background

As explained in Section 1, the freight railroad industry plays a vital role in the U.S. economy. The railroad industry, however, has not always enjoyed financial success, and in recent decades it has undergone a remarkable evolution. Following a cycle of decline that began in the 1960s, many freight rail carriers faced liquidation.[7] At the start of 1980s, the U.S. railroad industry accounted for only a small proportion of total ton-miles of freight, around 20%, and carrying less than pipelines.[8]

In response, a series of laws to deregulate the industry followed in the years 1973–1980.[9] Among them, the 1980 Staggers Act formally deregulated the industry, by offering railroad companies much greater pricing and operating freedom. The deregulation sparks a wave of mergers of railroad companies: from 1980 to 2005, the number of Class I railroads decreased from 39 to 7. Figure 2 illustrates the network of current seven Class I railroads: the Burlington Northern and Santa Fe railway (BNSF) competes with the Union Pacific railway (UP) in the west, the CSX Transportation (CSXT) and the Norfolk Southern railway (NS) competes in the east. Two Canadian Class I railroads the Canadian Pacific railway (CP) and the Canadian National railway (CN) connect freight shipment between Canada and U.S., and the Kansas City Southern railway (KCS) locates in the south and connects freight shipment between Mexico and the United States.

Figure 2: U.S. Class I Railroads



[7]In 1970 the nation's largest railroad Penn Central declared bankruptcy along with a dozen other northeastern railroads. See Grimm and Winston (2000) and Gallamore and Meyer (2014) for more discussions.

[8]Figure B.1 and B.2 plots the total ton-miles of freight carried by different mode of transport and how it changed from 1980 to 2011.

[9]Appendix A.2 provides details of regulation changes, and appendix A.1 shows a complete history of railroad mergers.

Since the mergers, the U.S. freight railroad industry has enjoyed a renaissance, becoming not only self-sustaining but one of the most efficient freight railroad systems in the world. Proportion of total ton-miles of U.S. freight carried by rail increased after the deregulation of 1980. According to Department of transportation, in 2017, railroads are the second largest transport mode in providing freight service in the United States. Railroads carry 1,675 billion ton-miles of freight, accounting for 30% of total freight transportation.

A majority of the freight revenues in the U.S. freight railroad industry are generated by so called Class I railroads.[10] According to AAR, in 2012, the seven Class I railroads generate $67.6 billion freight revenues, which accounts for around 95% of the total freight revenues generated by railroad transport.[11] In this paper, I focus on mergers of Class I railroads.

The U.S. railroad industry ships various types of commodities, but the majority of carloads are generated by bulk shipments like coal, chemical, and farm products. According to AAR, in 2012, the top four commodities measured by share of total tonnage are coal (37%), chemicals (10%), non-metallic minerals (9%), and farm products (8%). If we look at share of revenues, each of these four commodities accounts for 17% (coal), 14% (chemicals), 5% (non-metallic minerals), 8% (farm products) of total freight revenues. Coal has been the most important commodity in freight railroad industry ("King Coal"), but the revenues generated by shipping coal are decreasing. Instead, the carloads of intermodal shipping[12] (Misc. mixed shipments such as containers) are increasing nowadays, and intermodal shipment accounts for 13% of total freight revenues in 2012.[13]

2.2 The Story of Train 9-698-21

To better explain sources of cost efficiency and my model in the later sections, here I introduce three concepts of railroading: interchange cost, interconnecting route versus competing route, and allocation decisions regarding track maintenance and locomotives. I use the example of train 9-698-21 to explain these three concepts.[14]

[10] The Surface Transportation Board defines a Class I railroad as "having annual carrier operating revenues of $250 million or more in 1991 dollars", which is equivalent to $464 million in 2017 dollars.

[11] The rest of the revenues are generated by regional and short lines. Check Appendix B.1 for more details on regional and short lines.

[12] Intermodal is the use of two modes of freight, such as truck and rail.

[13] Check Appendix B.1 for more details on carloads originated by different types of commodities.

[14] Train 9-698-21 runs from Birmingham, Alabama to Los Angeles. To explain the concepts, I focus on the section between Los Angeles and Memphis, Tennessee. The original story provides more details on why interchange is costly and why coordination is a problem when two railroads are involved in a shipment.

Figure 3: Story of train 9-698-21

Source: Original story from *Trains* magazine "Twenty-four hours at Supai Summit"

Train 9-698-21 went from Birmingham to Los Angeles, and it begun with Burlington Northern via the Avard gateway in summer 1994. This was before the Burlington Northern railway and the Atchison, Topeka and Santa Fe railway discussed merging. The main customers of train 9-698-21 are UPS and J.B. Hunt, and the train is an express freight train and initiated to "reach downtown L.A. in time for UPS to deliver the next morning". The contract is for Santa Fe to be given haulage rights over BN to Memphis and Birmingham. Haulage rights means Santa Fe sold the service, then paid BN to run the trains east of Avard. However, according to Rollin Bredenberg, BNSF's vice president of transportation at that time, everything has not gone right with 9-698 though:

> *"It was very unreliable under the haulage agreement, pre-merger," reports Bredenberg, "BN's internal measurement of how well they ran trains did not include the performance of the Santa Fe haulage trains, so you can guess what happened." In an interview last year, Krebs (chairman of Santa Fe railway) said he finally had to tell key customers such as Hunt that they were free to go elsewhere until Santa Fe and BN could get their acts together.*

First, interconnecting and competing routes are both shown in this example. In Figure 3, the route from Los Angeles to Memphis is an example of an interconnecting route, because the train needs to ride both Santa Fe and Burlington Northern tracks. If a shipment originates in Claremore, Oklahoma, and is bound for Memphis, the owner has a choice of riding the Union Pacific or the Burlington Northern line, which is an example of competing routes. Second, this example shows the micro-foundation of interchange cost. When train 9-698-21 arrives at Avard gateway in Oklahoma, it needs to exchange crews and rolling stocks (railcars and locomotives) between the Burlington Northern railway and the Santa Fe railway. However, because BN and SF have different priorities over this train, it usually results in delays to finish this process. Moreover, check the condition of railcars and exchange rolling stocks take time and efforts, which further add to the interchange cost.

Last, I want to introduce the allocation decisions regarding track maintenance and locomotives. If more locomotives are available in the Avard gateway, the waiting time of train 9-698-21 will be shortened to finish the interchange. Meanwhile, adequate and constant maintenance of tracks is essential for railroad operation.[15] Regular track maintenance is costly,[16] and railroad companies decide on the frequencies of track evaluation in each region. In the example of train 9-698-21, if the railroad companies invest more in the route from Los Angeles to Memphis by allocating more locomotives and conducting more frequent track maintenance, the route efficiency from Los Angeles to Memphis will increase.

2.3 Why Does Network Structure Matter?

I will use a simple example below to elaborate why network structure (topology of the network) matters when there is economy of scope.

Figure 4 shows two networks A and B. There are two firms I and II in each network. Firm I owns nodes 1, 2, 3, 4. Firm II owns nodes 5, 6, 7, 8. There are six within-firm traffic, $1 \to 2$, $1 \to 3$, $1 \to 4$ within firm I, and $5 \to 6$, $5 \to 7$, $5 \to 8$ within firm II. Meanwhile, there are six across-firm traffic, $1 \to 6$, $1 \to 7$, $1 \to 8$, and $5 \to 2$, $5 \to 3$, $5 \to 4$. Each connected arc has a travel cost of 1, and an across-firm traffic incurs an extra interchange cost η. So, the cost of traveling from $1 \to 2$ is 1, and the cost of traveling from $1 \to 5$ is $2 + \eta$. Network B has exactly the same set-up.

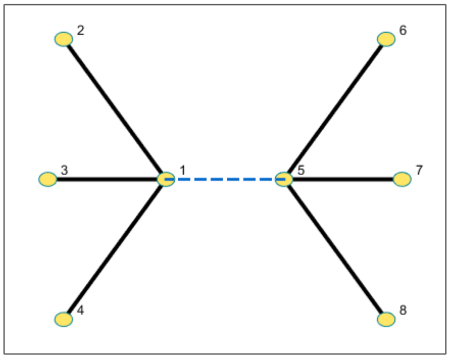
First, if there is no economy of scope, each traffic will have exactly the same travel cost in network A and network B. The travel cost of within-firm traffic in both network A and B are 1, and the travel cost of across-firm traffic in both network A and B are $2+\eta$. This is because when there is no economy of scope, each origin-destination market is independent and the travel cost only depends on the shortest travel distance from the origin to destination.

However, if there is economy of scope, the travel cost will be different in the two networks. I assume that the per-unit shipment cost is proportional to the inverse of total traffic. The numerical example shows that the average cost of within-firm shipment of network A and B is 0.11 and 0.16. The average cost of across-firm shipment of network A and B is 0.24 and 0.35 respectively. So, cost of shipment is larger in network B when there is economy of scope. This is because the ability to consolidate traffic is different in network A and network

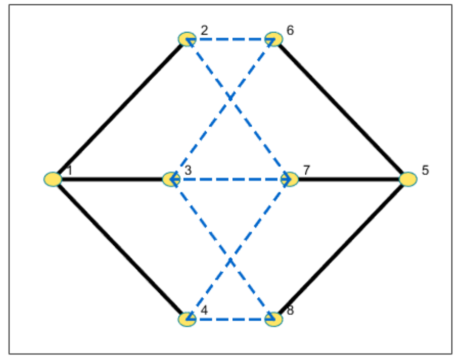
[15] To perform regular maintenance, railroad companies need to operate track evaluation cars to evaluate track geometry, performance based track geometry, rail wear etc and other things. Then the track evaluation cars will report the data for capital and maintenance planning.

[16] According to a Bloomberg report, a New Jersey Transit safety project costs more than $320 million, but the agency still fell behind on rail maintenance, let its ranks of train engineers dwindle and triggered a federal operations audit. https://www.bloomberg.com/news/articles/2018-10-18/how-nj-transit-s-lifesaving-rail-task-dragged-while-cost-doubled

Figure 4: Numerical Example (Why Network Structure Matters)



(a) Network A



(b) Network B

B. In network A, all 6 across-firm traffic go though arc $(1, 5)$. However, in network B only 2 across-firm traffic go through arc $(3, 7)$ — $1 \rightarrow 7$, $5 \rightarrow 3$. In network A, both within-firm traffic $1 \rightarrow 2$ and across-firm traffic $5 \rightarrow 2$ will go through arc $(1, 2)$, but in network B, only within-firm traffic $1 \rightarrow 2$ will go through arc $(1, 2)$. Therefore, on average network A has lower shipment cost because traffic can be better consolidated in network A than in network B.[17] Therefore, network structure (topology of network) matters when there is economy of scope, because network structure affects how much firms benefit from economy of scope.

[Insert Figure 7 here.]

The next question is, is it true that firms can achieve economy of scope by consolidating traffic in the railroad industry in reality? Yes. The Chief Operating Officer, Cindy Sanborn of CSXT states that "An essential feature of the operating plan is to consolidate traffic over a smaller number of efficient, high-volume routes."[18] The idea of achieving greater cost efficiency by consolidating traffic is also supported by the former CEO of the Southern Pacific railway Krebs (2018) and the former CEO of the Canadian National railway Harrison (2005).

In reality, in a much more complicated network, firms pricing decision and routing decision will significantly affect the degree of economy of scope that they can achieve before and after

[17] This does not mean that the shipment cost of every origin-destination pair is smaller in network A than in network B. For example, the shipment cost of all within-firm traffic is 0.11 in network A, but the shipment cost of within-firm traffic in network B is: cost of $1 \rightarrow 2$, $1 \rightarrow 4$, $5 \rightarrow 6$, $5 \rightarrow 8$ is 0.2192, which is larger than 0.11, but cost of $1 \rightarrow 3$ and $5 \rightarrow 7$ is 0.0563, which is smaller than 0.11.

[18] Ex Parte No. 711 (Sub-No.1) Reciprocal Switching, Opening Comments of CSX Transportation INC.

mergers. This is what I will capture in my full model, by considering firms pricing, routing, and allocation decision in their network.

3 Data

I use three main datasets to document the changes in cost efficiency following railroad mergers. The confidential Carload Waybill sample from the Surface Transportation Board (STB), the Class I Railroad Annual Report, and the Commodity Flow Survey. The confidential version of the Carload Waybill Sample provides detailed information on shipment price and corresponding shipment attributes; the Class I Railroad Annual Report R-1 dataset contains information on firm attributes and aggregate operational statistics; and the Commodity Flow Survey has information on shipment volumes of different transportation mode. I also obtain geographic information from Department of Transportation on all U.S. rail lines and their associated railroad companies. Based on the ancestry of each rail line, I trace back to time and reconstruct the rail network of each railroad firm between 1984 and 2010.

The Carload Waybill Sample is a sample of carload waybills for all U.S. rail traffic submitted to the STB by those rail carriers terminating 4,500 or more revenue carloads annually.[19] The data contain detailed shipment information on such attributes as commodities carried, total billed weight, equipment used, participating railroads, and origin, destination, and interchange locations of each load. The locations are recorded in BEA Economic Areas, and there are 173 BEA areas in total (Figure 5). The confidential Carload Waybill Sample that I have access to ranges from 1984 to 2010. Table 1 shows the summary statistics of key variables of the waybill data.

Table 1: Summary Statistics of Variables

	Mean	Std. Dev.	25th Percentile	Median	75th Percentile
Price per railcar (in $)	1,034	1,399	384	703	1,266
Shipment weight per railcar (in ton)	54	46	16	26	102
Travel distance (in miles)	1,045	773	404	854	1,647
Number of waybills (Carrier-Origin-Destination-Date)				12,113,581	

Source: STB, Carload Waybill Sample

[19]STB is an independent adjudicatory and economic-regulatory agency charged by Congress with resolving railroad rate and service disputes and reviewing proposed railroad mergers. STB is created on January 1, 1996 by the ICC Termination Act of 1995, the Board is the successor to the former Interstate Commerce Commission. More details about Carload Waybill Sample can be found at https://www.stb.gov/stb/industry/econ_waybill.html.

Shipment price per carload ranges from \$384 to \$1,266 at 25th to 75th percentile, with an average price at \$1,034. Shipment weight per carload ranges from 16 tons to 102 tons at 25th to 75th percentile, with an average weight at 54 tons. Average travel distance of each shipment is at 1,045 miles. The 25th percentile of travel distance is 404 miles (650 km). It shows that railroad shipments are mostly long-distance shipping (> 300 km). In my data the median shipment price per ton-mile is 2.65 cents. This number is comparable to the price per ton-mile reported in the industry. In year 2001, the average price per ton-mile in the U.S. is 2.32 cents according to AAR.

Then I show the level of competition in this industry. Table 2 shows the total number of O-D markets, average number of competitors serving each O-D market, and the percentage of interchange lines. I also show how these numbers change across time.

Table 2: Summary Statistics of Market Competition

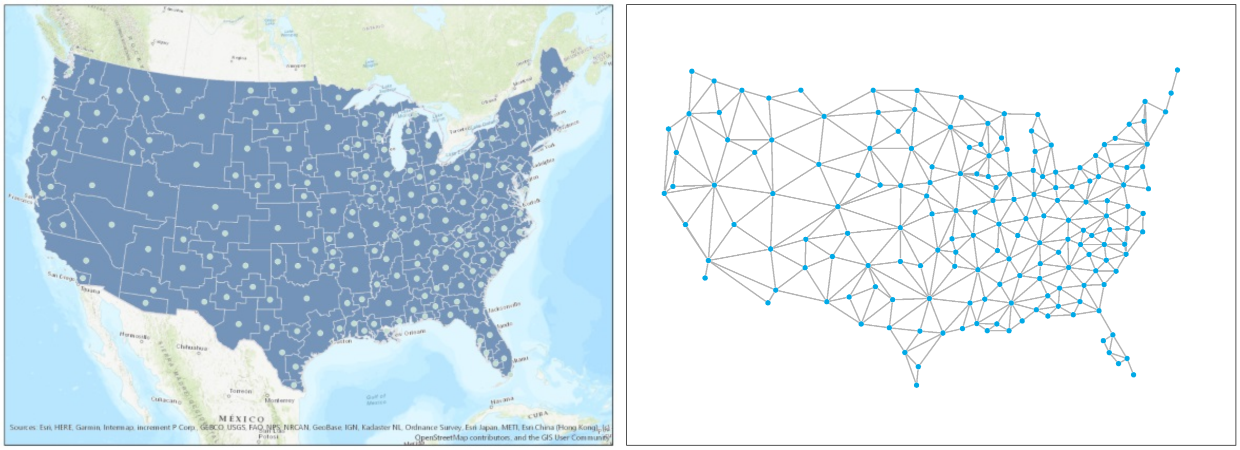
Year	Number of Waybills	Percentage of Interchange Lines	Number of Competitors in an O-D Market			Number of O-D Market (at BEA-to-BEA level)
			mean	25th percentile	75th percentile	
1985	262,703	41%	3	1	3	12,088
1990	323,570	35%	2	1	3	11,835
1995	453,802	26%	2	1	3	11,632
2000	544,738	14%	2	1	2	11,732
2005	611,033	11%	2	1	2	11,611
2010	533,364	10%	2	1	2	10,485

Source: STB, Carload Waybill Sample

First, Table 2 shows that the total number of waybills in the waybill sample (the waybill sample is 2% of total waybills) changed from around 263,000 to 533,000 from year 1985 to 2010. This shows that total volume of railroad shipment doubled from 1985 to 2010. This is consistent with the story I show in figure 1. Meanwhile, the percentage of interchange lines decreased from 41% to 10% while the total traffic volume doubled. It shows that following the wave of mergers from 1985 to 2010, there is a significant decrease of interchanges. The number of O-D markets is relatively stable across the years, with a small decrease from 11,835 to 11,611 from year 1990 to 2005. Therefore, the change of extensive margin after the mergers does not seem to be a big concern. Last, the average number of competitors in each O-D market slightly decreased from 3 to 2 from 1985 to 2010. It indicates that firms conduct oligopolistic competition in the local markets.[20]

[20] To save space, here I only show the statistics every five years from 1985 to 2010. The year-by-year table is shown in Appendix B.3. The year-by-year table tells the same story as in Table 2 here.

Figure 5: U.S. Rail Network



(a) BEA Regions and Their Centroids

(b) Virtual Network

Note: Panel (a) of figure 5 shows the 170 BEA regions in the US and their centroids. Based on the locations of the BEA regions and their adjacency, I construct the virtual network in Panel (b).

4 Definition of Network

Following Galichon (2016), I define a network $\mathcal{G}$ as a directed graph $(\mathcal{Z}, \mathcal{A})$, where $\mathcal{Z}$ is a set of nodes and $\mathcal{A}$ is a set of arcs $\mathcal{A} \subseteq \mathcal{Z}^2$ which are pairs (x, y) where $x, y \in \mathcal{Z}$. In the model I define nodes as centroids of BEA regions[21] and arcs as rail lines that connect each BEA economic area. Panel (a) of figure 5 shows the 170 BEA regions in the US and their centroids. Based on the locations of the BEA regions and their adjacency, I construct the virtual network in Panel (b).

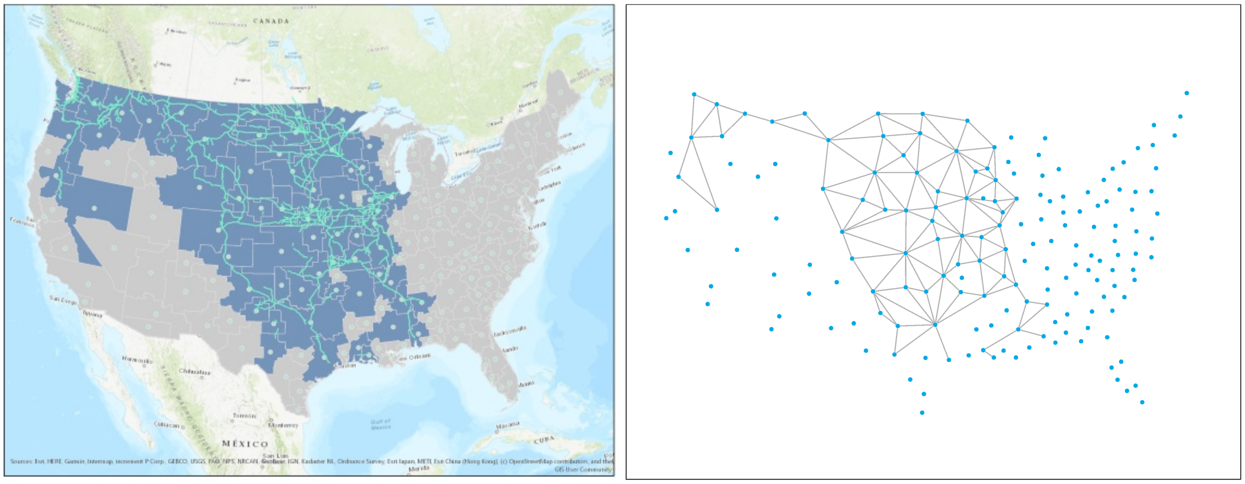
Then I need to obtain network information of each railroad firm. I denote the network of each firm j as $\mathcal{G}_j$, with arcs $\mathcal{A}_j$ and nodes $\mathcal{Z}_j$. Panel (a) of Figure 6 shows the actual rail network of BN in 1994, and panel (b) shows the virtual network of BN that is constructed from the actual information. For example, if two adjacent regions, region a and region b, are connected by rail lines owned by BN, then nodes $a \in \mathcal{Z}_j, b \in \mathcal{Z}_j$, and arc $(a, b) \in \mathcal{A}_j$.

Following the same process, I construct the virtual network of every railroad firm in every year from 1986 to 2005. I am able to do that because I obtain detailed geographic information of each rail line (coordinates) and information about the ancestry of rail lines from the Federal Transit Administration. The geographic coordinates and historical information allow me to reconstruct the rail network back in time.

Network Measures. I use three types of network measures in my analysis. Two are

[21] BEA Economic Areas are used by the STB to conduct economic analysis.

Figure 6: The Rail Network of BN



(a) Actual Rail Network of BN in 1994 (b) Virtual Network of BN

Note: Panel (a) of Figure 6 shows the actual rail lines owned by BN in 1994. Panel (b) shows the virtual network of BN that is constructed from the actual information. For example, if two adjacent regions, region a and region b, are connected by rail lines owned by BN, then nodes $a \in \mathcal{Z}_j, b \in \mathcal{Z}_j$, and arc $(a, b) \in \mathcal{A}_j$.

used to measure the connectivity of the network of each railroad firm, and two are used to capture the centrality (the hubness) of an node in the network.[22]

I start from the notion of *Degree Centrality*. The degree of a node i is the number of neighbors of node i, and it equals the total number of other nodes (BEA regions) with a connection to node i in network $\mathcal{G}_j$ of firm j. Degree Centrality of node i in network $\mathcal{G}_j$ is written as:

$$\begin{aligned} d_i(\mathcal{G}_j) &= \#\{k : a_{ik} = 1\} \\ &= \sum_{(i,k)\in\mathcal{A}_j} 1 \end{aligned} \tag{1}$$

Degree Centrality of node i is larger when the number of links connected to i increases.

The second measure I utilize is *Betweenness Centrality*, which captures how frequently the node is found on the shortest path from an origin to a destination. It is calculated as follows: [23]

[22] I define the network measures similarly as in Ciliberto et al. (2019) and Chen et al. (2020). See Jackson (2010) for a more detailed discussion.

[23] As a robustness check, I add weights in calculating the centrality measures. For example, the betweenness

$$B_i(\mathcal{G}) \coloneqq B_i(\mathcal{G}_j) = \sum_{o,d \in \mathcal{Z}_j} \frac{\mathbb{1}\{i \in l(o,d)\}}{(Z_j - 1)(Z_j - 2)} \tag{2}$$

where Z_j is the total number of nodes in network $\mathcal{G}_j$.

The last measure is the Weighted Average Neighbor Betweenness. This measure captures the extent to which node i is connected to other highly traveled nodes in the network of firm j:

$$\bar{B}_i^w(\mathcal{G}_j) = \frac{\sum_{k=1}^{Z_j} a_{ik} B_k}{\sum_{k=1}^{Z_j} a_{ik}} \tag{3}$$

where $a_{ik} = 1$ if $(i,k) \in \mathcal{A}_j$.[24]

I illustrate how to understand the three network measures in Appendix C. I show when the value of the network measures changes, how the network structure of railroad changes visually in my data.

5 Reduced-form Evidence

To provide evidence of efficiency gain, I examine how price changes following railroad mergers. The regression model studying change in price is specified as:

$$\log P_{s,odt} = \mu_{od} + \gamma_s + \lambda_t + \delta_1 D_{s,odt} + X'_{s,odt}\beta + \epsilon_{s,odt}$$

centrality can be calculated as follows

$$B_i(\mathcal{G}) = \sum_{o \neq i, d \neq i, o \neq d} \frac{\mathbb{1}\{i \in l(o,d)\}\omega_{o,d}}{\sum_{m \neq i, n \neq i, m \neq n} \omega_{m,n}}$$

where $l(o,d)$ is the shortest path from o to d, and $\omega_{o,d}$ is the weight of route O-D. I define $\omega_{o,d}$ in three ways:

a $\omega_{o,d} = 1$ if $o \in \mathcal{Z}_j, d \in \mathcal{Z}_j$. This is my baseline specification, which measures the Betweenness Centrality *within* each firm j's network $\mathcal{G}_j$. Under this case the Betweenness Centrality measure is the same as in Ciliberto et al. (2019).

b $\omega_{o,d} = 1$ if $o \in \{\mathcal{Z}_1 \cup \mathcal{Z}_2\}, d \in \{\mathcal{Z}_1 \cup \mathcal{Z}_2\}$, where $\mathcal{Z}_1$ is the node of merging firm 1 and $\mathcal{Z}_2$ is the node of merging firm 2. This specification measures the Betweenness Centrality within the network of two merging firms.

c $\omega_{o,d} = q_{o,d}$. The weight matrix is the total quantity of shipment. I use actual waybill data to define O-D markets here. Interchange lines are also included. For example, for a route from $a \to b \to c$, carried by firm 1 then firm 2. The (o,d) of firm 1 is defined as (a,b).

[24] I can use a more general form where $\bar{B}_i^w(\mathcal{G}_j) = \frac{\sum_{k=1}^{Z_j} \omega_{ik} a_{ik} B_k}{\sum_{k=1}^{Z_j} \omega_{ik} a_{ik}}$. For now I am having $\omega_{ik} = 1, \forall (i,k)$.

where $D_{s,odt}$ is an indicator of if a merger has happened to firms that provide service s from o to d before or equal to time t, and $X_{s,odt}$ are shipment attributes. I also control for firm, route, and year fixed effect in the regression. Service s is either single-line service or joing-line service. Single-line service is carried by one railroad firm from the origin all the way to the destination. Joint-line service is carried by at least two railroad firms with interchange involved.

Table 3 shows the estimation result of price effect of mergers. The result suggests that on average a railroad merger reduces the shipment price by 9.4%. By opening up the mergers and examining each individual merger, I find that the price effect is largely consistent across individual mergers.[25] Then, to further decompose the effect of railroad mergers on price changes by different route types, I interact the merge dummy with three route types: interconnecting, competing, and non-interconnecting, noncompeting. As explained in section 2, an interconnecting route is a route in which two firms conduct interchange and finish the shipment jointly. Results in column 2 of Table 3 show that the interconnecting routes have the largest price reduction among all route types, with price decreasing by 11% after mergers. In comparison, the other route types have a price reduction of about 6.5% following mergers.

Table 3: Effect of Mergers on Price Change (by Route Types)

	(1) Log Price	(2) Log Price
Indicator of Merger	-0.0935***	
	(0.0142)	
Indicator of Merger × Indicator of Interconnecting Route		-0.107***
		(0.0178)
Indicator of Merger × Indicator of Competing Route		-0.0690***
		(0.0180)
Indicator of Merger × Non-interconnecting, Noncompeting Route		-0.0641***
		(0.0171)
N	12,110,107	12,110,107
Firm FE	Yes	Yes
Year FE	Yes	Yes
O-D Route FE	Yes	Yes

Standard errors in parentheses. Clustered at route level.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Surface Transportation Board, Carload Waybill Sample

[25] Table D.2 shows the robustness check results by looking at change of prices following each railroad merger.

However, there might be concerns that these price effects are driven by competition from other transportation modes such as trucking, not from the effect of railroad mergers. Because comparable origin-destination-level shipment data for trucking is lacking, I cannot directly run the price regressions by controlling for competition of trucking. Instead, to address this concern, I examine the price effects of mergers for different types of commodities. The argument is that the shipment of different types of commodities faces different levels of competition from other transportation modes. Therefore, if the price effects are driven by changes in other transportation modes, the price effects should be greater for commodities facing higher competition from other modes of transport. Table 4 summarizes the two commodities used in this analysis: the Commodity Flow Survey (CFS) of 2012 shows that coal is mainly shipped by railroads. Only 1.5% of coal is shipped by trucking, and 94.8% is shipped by rail. In comparison, food or kindred products are largely shipped by trucking: 76.2% of food or kindred products is shipped by trucking, and only 23.5% is shipped by rail.[26]

Table 4: Summary Statistics of Coal and Food Shipment

	Coal (STCC 11)	Food or kindred products (STCC 20)
Total Ton-Miles in 2012 (Truck)	1.5%	76.2%
Total Ton-Miles in 2012 (Rail)	94.8%	23.5%

Source: U.S. Census Bureau, Commodity Flow Survey

Table 5 shows the estimation results of price effect of mergers for coal and food or kindred products respectively. The results show that railroad mergers have a significantly negative price effect for both coal and food or kindred products. Moreover, the price effect of mergers is greater for coal. This is contrary to the prediction of the hypothesis that the price effect is driven by changes in other transport modes.

The estimated price effect is comparable to other analysis of freight railroad mergers. In the Surface Transportation Board analysis of the Union Pacific and Southern Pacific merger, the shipment price of coal was found to decrease by 11%, and that of other commodities by 6% after the merger. The price effect of mergers found in the U.S. freight railroad industry is larger than the price effect of mergers found in some other industries. For example, Ashenfelter et al. (2015) find that the estimated price reduction caused by merger efficiency is 2%. In comparison, I find a 9.4% overall shipment price reduction, and a 17.9% average price reduction for coal shipment, which suggests that cost efficiency following mergers is very important in this industry.

[26] All weights are calculated by total ton-miles of shipment.

Table 5: Effect of Merger on Price Change (by Commodities)

	(1)	(2)
	Log Price (Coal)	Log Price (Food or Kindred Products)
Indicator of Merger	-0.179***	-0.052***
	(0.028)	(0.014)
Log Billed Weight	-0.030	-0.212***
	(0.020)	(0.010)
Ownership of Railcar (Private)	-0.096***	-0.132***
	(0.027)	(0.008)
Ownership of Railcar (Trailer Train)	-0.021	-0.144***
	(0.071)	(0.016)
N	1,002,552	882,066
Firm FE	Yes	Yes
Year FE	Yes	Yes
O-D Route FE	Yes	Yes

Standard errors in parentheses. Clustered at route level.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Surface Transportation Board, Carload Waybill Sample

As a robustness check of the price effect of mergers, I run the price regression for each type of commodity in Appendix D. The results show that price reduction following railroad mergers is consistent across different types of commodities. If we look particularly at commodities that are largely shipped by rail, such as coal, chemicals, and construction materials (clay, concrete, etc.), there is a large and significant price reduction following railroad mergers.

The Need for a Model. In this section, I show that prices decrease after mergers, and the price effect is greater for inter-connecting routes. Examining individual routes only is insufficient for understanding mergers in this industry, however, because origin-destination railroad markets are interdependent. Furthermore, in this networked industry it is hard to find a clean control group for performing a difference-in-differences analysis. Efficiency changes in one part of the network will affect the routing and optimization problem of the whole network, thereby affecting the efficiency in another part of the network. To examine the mechanism of cost efficiency following railroad mergers, I therefore build a structural model to capture firms' endogenous decisions of pricing, routing, and allocation of resources.

6 Model

In the model, railroad firms play a two-stage game. First, firms compete in local markets and choose prices simultaneously, conditional on their perceived operational cost of themselves and their rivals. Then given the shipment demand in each local market, in the second stage firms choose routing and allocation decision to minimize operational cost. In equilibrium, the perceived cost of operating at the pricing stage is consistent with the outcome from firms cost minimization problem.

I define a market as an origin-destination pair. Firms provide either single-line service or joint-line service to serve a market. Single-line service is carried by one railroad firm from the origin all the way to the destination. Joint-line service is carried by at least two railroad firms with interchange involved. Services serving each market is exogenous and obtained from the data. For interconnecting routes, I assume that the originating firm of the joint-line service determines the price of the joint-line service.[27] The network structure of each railroad firm is also exogenous and obtained from the data.

On the demand side, I assume customers make discrete choices based on shipment price and shipment characteristics. Customers also face an outside option, which is shipping by other transportation mode rather than railroads.

6.1 Demand

I assume a logit demand of railroad shipment in this paper. The assumption of logit demand is standard in estimating transportation demands, such as in Peters (2006) and Grieco et al. (2020). In each origin-destination market, consumer i chooses service s to ship from origin o to destination d. Service s is either single-line service or joint-line service. For example,

[27]I assume that there is no issue with double marginalization in the pricing of the interconnecting route based on findings from Alexandrov et al. (2018). Assuming that the origin railroad has all the bargaining power is one way to interpret this assumption. Another way is to think that the two railroad firms in the joint-line service jointly determine the price as one entity, and then divide the revenue between them in a meaningful way. The only assumption that I need is that there is no double-marginalization of pricing. In reality how the price and revenue-sharing of joint-service is much more complicated. For those who are interested, Appendix B.2 documents interviews with managers in this industry. Introducing other mechanism of revenue sharing like Nash bargaining to my model is possible, but it does not provide extra insight into how network measures affect effects of mergers, which is the key question that I am studying in this paper. That is why I am assuming away from those complications. I also assume that firms do not consider the cannibalization between their own single-line service and joint-line service that they participate in for a given o-d market. The last assumption has little impact on the empirical results because in the data a firm rarely provides both single-line service and joint-line service in the same origin-destination market.

shipment carried only by the Union Pacific railway from origin to destination is a single-line service. Shipment carried first by the Burlington Northern railway from origin to interchange station, and then carried by the Santa Fe railway from the interchange station to destination is a joint-line service.

The utility function of consumer i choosing service s is:

$$
\begin{aligned}
u_{is,od} = \alpha \cdot p_{s,od} + \beta_1 \cdot \log TravelDistance_{s,od} \\
+ \beta_2 \cdot \log TotalTrackMiles_{s,od} + \xi_{s,od} + \varepsilon_{is,od}
\end{aligned}
$$

Customers care about shipment price $p_{s,od}$ and travel time. Because I do not observe travel time in my data, I use two variables *travel distance* and *total track miles* to approximate travel time. *Travel distance* measures the average travel distance of service s to go from o to d. Because the network of each firm providing the service is different, the routing of each service is different hence the travel distance of each service varies. I expect β_1 to be negative, because the longer the travel distance is conditional on an *o-d* market, the longer the travel time is likely to be. *Total track miles* measures the total amount of physical tracks that the firm providing service s has in the origin and destination area. For example, say the *o-d* market is from LA to Houston, and the service is a single-line service provided by the Union Pacific railway. Then $TotalTrackMiles_{s,od}$ measures the total amount of physical tracks that the Union Pacific railway has in LA and Houston area. The idea is that the more physical tracks there are, the easier it is to move things from the customer to the railroad, hence the travel time will be shorter.

To control for unobservable market-specific demand factors, and efficiency that are constant within significant subsets of the data, I include fixed effects of O-D markets and firm fixed effects. I also put standard i.i.d. extreme value type assumption on individual shocks $\varepsilon_{is,od}$.

6.2 Firm's Problem

Each firm j owns a network $\mathcal{N}_j$ with corresponding nodes $\mathcal{Z}_j$ and arcs $\mathcal{A}_j$. Firms play a two-stage game. First, firms compete in local markets and choose prices simultaneously, conditional on their perceived operational cost of themselves and their rivals. Then given the shipment demand in each local market, in the second stage firms choose routing and allocation decision to minimize operational cost. In equilibrium, the perceived cost of operating at the pricing stage is consistent with the outcome from firms cost minimization problem.

There are multiple services $\{s\}_{od}$ serving each o-d market. Service s can be either single-line service or joint-line service. $s := [j_1, j_2]$ where j_1 is the origin railroad and j_2 is the termination railroad.

1. If $j_1 = j_2 = j$, s is a single-line service and carried by railroad firm j from the origin all the way to the destination.

2. If $j_1 \neq j_2$, s is a joint-line service. Shipment from o to d will be carried by firm j_1 from o to the interchange station m, and then by firm j_2 from the interchange station m to destination d. Location of interchange station m is specific to each service-origin-destination pair, and exogenously obtained from data.[28]

Type of services $\{s\}_{od}$ in each O-D market is also exogenous and obtained from data.[29]

6.2.1 Local Market Competition

In the first stage, firms compete in prices in local markets, conditional on their perceived cost of themselves and their rivals. A market is an origin-destination pair. Firm j chooses price of each single-line service that it provides, and also chooses price of each joint-line service that firm j is the origin railroad.

The optimization problem of pricing is denoted as:

$$\pi_{s,od} := \max_{p_{s,od}} \left[p_{s,od} - \widetilde{C}_{s,od} \right] \cdot Q_{s,od}(p_{s,od}, p_{-s,od})$$

where $\widetilde{C}_{s,od}$ is the perceived cost of providing service s to ship from origin o to destination d.[30]

The per-unit transportation cost of a service $C_{s,od}$ depends on the routing decision $\mathcal{R}_{s,od}$ and allocation decision $\boldsymbol{I}_j$. Each element $I_{j,ab}$ of $\boldsymbol{I}_j$ measures the amount of resources that firm j allocates to arc (a, b).

[28] In the example of train 9-698 in section 2.2, the Burlington Northern railway and the Santa Fe railway provide a joint-line service from LA to Memphis. The shipment is carried by the Santa Fe railway from LA to Avard, OK, conduct interchange with the Burlington Northern railway, and then carried by the Burlington Northern railway from Avard, OK to Memphis. In this case, Avard, OK is the interchange station. In practice, the Waybill data shows that the location of interchange station for each service-origin-destination pair barely changes

[29] In the U.S., the difficulty of expropriating trackage rights has got to a point where there is virtually no new tracks that have been built in the last fifteen years. Entry into new markets where the firms have no physical tracks is very difficult.

[30] I assume that each service is priced independently even if a firm may participate in both a single-line service and a joint-line service in the same market. For example, imagine that there are three services serving market from LA to Memphis: a single-line service provided by the Burlington Northern railway, a single-line service provided by the Union Pacific railway, and a joint-line service provided by the Burlington Northern railway and the Santa Fe railway. I assume that the Burlington Northern railway price the single-line service and the joint-line service independently. In reality, the case that a railroad offers both a single-line service and a joint-line service in the same O-D market is very rare, so I made this assumption to simplify computation. Relaxing this assumption barely has any effect on the results.

For single-line service $s \coloneqq [j_1, j_2] = [j, j]$, $C_{s,od}$ is specified as:

$$C_{s,od} \coloneqq C_{s,od}(\boldsymbol{I}_j, \mathcal{R}_{j,od}) = \sum_{(a,b)\in\mathcal{R}_{j,od}} c_{j,ab}(\boldsymbol{I}_j) \tag{4}$$

where $c_{j,ab}$ is the arc-level cost of firm j. Routing $\mathcal{R}_{j,od} \in \mathcal{A}_j$ is a subset of connected arcs that routes firm j from origin o to destination d. Intuitively, equation 4 says that the per-unit transportation cost $C_{s,od}$ is the summation of arc-level costs that firm j routes from o to d.

For joint-line service $s \coloneqq [j_1, j_2]$, $j_1 \neq j_2$, $C_{s,od}$ is specified as:

$$\begin{aligned} C_{s,od} &\coloneqq C_{s,od}(\boldsymbol{I}_{j_1}, \boldsymbol{I}_{j_2}, \mathcal{R}_{j_1,om}, \mathcal{R}_{j_2,md}) && (5) \\ &= \sum_{(a,b)\in\mathcal{R}_{j_1,om}} c_{j_1,ab}(\boldsymbol{I}_{j_1}) + \sum_{(a',b')\in\mathcal{R}_{j_2,md}} c_{j_2,a'b'}(\boldsymbol{I}_{j_2}) + \eta && (6) \end{aligned}$$

where η is the interchange cost, and m is the interchange station. Equation 5 says that the per-unit transportation cost $C_{s,od}$ of a joint-line service is the summation of cost of firm j_1 from o to m, and the cost of firm j_2 from m to d, plus the interchange cost η.[31]

I follow Galichon (2016) and Fajgelbaum and Schaal (2020) in defining the per-unit cost of transportation at arc level. The arc-level transportation cost of firm j at arc (a, b) is parameterized as:[32]

$$c_{j,ab} = \frac{\delta_0 Dist_{j,ab}}{I^{\gamma}_{j,ab}}$$

[31] In my model I consider joint-line service with only one interchange, because in the data more than 90% of joint-line service involves only one interchange. This assumption can easily be relaxed. For example, if there is more than one interchange, say service s from o to d is $s = [j_1, j_2, \cdots, j_n]$, the general form of transportation cost of service s from origin o to destination d is written as:

$$\begin{aligned} C_{s,od} &= \sum_{j\in J(s)} \sum_{(a,b)\in\mathcal{R}_{j,od_j(s)}} c_{j,ab} + \#_{interchanges} \cdot \eta \\ &= \sum_{j\in J(s)} \sum_{(a,b)\in\mathcal{R}_{j,od_j(s)}} \left[\frac{\delta_0 Dist_{j,ab}}{I^{\gamma}_{j,ab}}\right] + \#_{interchanges} \cdot \eta \end{aligned}$$

where $J(s)$ is the set of firms that provides service s, and $\#_{interchanges}$ is the total number of interchanges incurred in providing service s.

[32] It is easy to enrich the model by adding geographic characteristics to the cost function such as in Fajgelbaum and Schaal (2020); δ_0 in my model can be further parameterized as:

$$\begin{aligned} &\delta_0 = \delta_1 \left(1 + |\Delta\text{Elevation}|_{ab}\right)^{\delta_2} \\ \text{or}\quad &\delta_0 = \delta_1 \left(1 + |\Delta\text{Elevation}|_{ab}\right)^{\delta_2} \delta_3^{\text{CrossingRiver}_{ab}} \delta_4^{\text{AlongRiver}_{ab}} \end{aligned}$$

The arc-level transportation cost $c_{j,ab}$ depends on the distance between a and b, and the amount of resources firm j allocates to arc (a, b). The efficiency parameter γ is expected to be positive. Therefore, if firm j allocates more resources to arc (a, b), the arc-level transportation cost $c_{j,ab}$ will be smaller. For any arc (a', b') such that $(a', b') \notin \mathcal{A}_j$, the arc-level cost of transportation $c_{j,a'b'}$ is ∞.[33]

The first-order condition with respect to price $p_{s,od}$ is calculated as

$$\frac{\partial \pi_s}{\partial p_{s,od}} := Q_{s,od} + (p_{s,od} - \widetilde{C}_{s,od}) \cdot \frac{\partial Q_{s,od}}{\partial p_{s,od}}$$

A key assumption here is that firms do not consider how pricing affects the optimal routing and allocation decision in the second-stage, i.e., $\frac{\partial \widetilde{C}_{j,od}}{\partial p_{j,od}} = 0$. This assumption is consistent with what people do in this industry. When railroad firms make pricing decisions, they do not strategically take into consideration of how the resulting demand affects the subsequent operational decision.[34]

Then combine with the logit demand, I can get the equilibrium prices $\{\tilde{p}_{s,od}, \tilde{p}_{-s,od}\}$ and demand $Q_{s,od}(\tilde{p}_{s,od}, \tilde{p}_{-s,od})$ for each service s in market o-d, conditional on firms perceived cost of operating in each market.

6.2.2 Operational Decision in the Network

Given the demand $Q_{s,od}(\tilde{p}_{s,od}, \tilde{p}_{-s,od})$, in the second stage firms choose the optimal routing $\{\mathcal{R}_{j,o_j d_j(s)}\}_{o_j \in \mathcal{Z}_j, d_j \in \mathcal{Z}_j}$ and allocation decisions $\{I_{j,ab}\}_{(a,b) \in \mathcal{A}_j}$ to minimize operational cost. The cost minimization problem is subject to constraint of capital allocation and balanced-flow constraints.[35]

[33] I did not include congestion in my model. I can add congestion in the cost function by modifying c_{ab} as a function of total traffics q_{ab} that go through arc (a, b). The reason I did not include congestion in my model is because the routing problem is no longer a linear programming problem if congestion is included. Adding congestion into the model will not provide new insights into the results, but will add significant computational complexity.

[34] See the interview with business development manager of the Canadian National railway in Appendix B.2. I can relax this assumption by calculating the envelope condition of how the optimal routing (a linear programming problem) and optimal allocation reacts to price, but this will significantly increase the computation time especially in estimation.

[35] Here is what the notation of routing decision $\{\mathcal{R}_{j,o_j d_j(s)}\}_{o_j \in \mathcal{Z}_j, d_j \in \mathcal{Z}_j}$ means:

- If s is a single-line service of firm j from o to d, then $\mathcal{R}_{j,o_j d_j(s)} = \mathcal{R}_{j,od}$.
- If s is a joint-line service and firm j is the origin railroad that carries the shipment from origin o to the interchange station m, then $\mathcal{R}_{j,o_j d_j(s)} = \mathcal{R}_{j,om}$.
- If s is a joint-line service and firm j is the termination railroad that carries the shipment from the interchange station m to destination d, then $\mathcal{R}_{j,o_j d_j(s)} = \mathcal{R}_{j,md}$.

Denote $J(s)$ as the set of firms that provide service s, the cost minimization problem of firm j is written as:[36]

$$\min_{\{\mathcal{R}_{j,o_jd_j(s)}\},\{I_{j,ab}\}_{(a,b)\in\mathcal{A}_j}} \sum_{s:j\in J(s)} C_{s,o_jd_j(s)}(\boldsymbol{I}_j, \mathcal{R}_{j,o_jd_j(s)}) \cdot Q_{s,od}(\tilde{p}_{s,od}, \tilde{p}_{-s,od}) \tag{7}$$

such that

Constraint of capital allocation:

$$\sum_{(a,b)\in\mathcal{A}_j} I_{j,ab} \leq K_j$$

Balanced-flow constraint: for any service s in any market od, $\forall m' \in \mathcal{Z}_j$,

$$D_{j,m'} + \sum_{a\in\mathcal{Z}_j(m')} Q_{s,od} \cdot \mathbb{1}\{(a,m') \in \mathcal{R}_{j,o_jd_j(s)}\} \leq \sum_{b\in\mathcal{Z}_j(m')} Q_{s,od} \cdot \mathbb{1}\{(m',b) \in \mathcal{R}_{j,o_jd_j(s)}\}$$

The intuition of the capital allocation constraint is that each firm j has a fixed amount of capital K_j that firm j can allocate to arcs $\mathcal{A}_j$ in its network. The intuition of the balanced-flow constraint is that for each node m', the total inflow of traffic plus the net demand equals to the total outflow of traffic.[37]

Let us first look at the routing problem. Firm j's routing problem from origin o_j to destination d_j is written as:

$$\min_{\mathcal{R}_{j,o_jd_j(s)}} \sum_{(a,b)\in\mathcal{R}_{j,o_jd_j(s)}} c_{j,ab} \cdot Q_{s,od} \tag{8}$$

[36] If s is a single-line service provided by firm j, then $J(s) = \{j\}$. If s is a joint-line service provided by firm j_1 and j_2, then $J(s) = \{j_1, j_2\}$.

[37] Here are the details of the balanced-flow constraint:

- $\mathbb{1}(\cdot)$ is an indicator function and $\mathbb{1}\{(a,m') \in \mathcal{R}_{j,od_j(s)}\} = 1$ if arc (a,m') is in the routing from o_j to d_j of firm j.
- $a \in \mathcal{Z}_j(m)$ means that a is in the neighborhood of node m. The total inflow of traffic into node m' is the summation of traffic from all the arcs (a,m') of firm j such that $a \in \mathcal{Z}_j(m)$.
- $D_{j,m'}$ is the net demand at node m'. $D_{j,m'} = \begin{cases} Q_{s,od} & \text{if } m' = o \\ -Q_{s,od} & \text{if } m' = d \\ 0 & \text{otherwise} \end{cases}$

s.t. $\forall m' \in \mathcal{Z}_j$, the balanced-flow constraint is satisfied:

$$D_{j,m'} + \sum_{a \in \mathcal{Z}_j(m')} Q_{s,od} \cdot \mathbb{1}\{(a,m') \in \mathcal{R}_{j,o_j d_j(s)}\} \leq \sum_{b \in \mathcal{Z}_j(m')} Q_{s,od} \cdot \mathbb{1}\{(m',b) \in \mathcal{R}_{j,o_j d_j(s)}\}$$

Solving equation 8 is equivalent to solving the minimization problem:

$$\min_{\mathcal{R}_{j,o_j d_j(s)}} \sum_{(a,b) \in \mathcal{R}_{j,o_j d_j(s)}} c_{j,ab}$$

$$\Rightarrow \min_{\mathcal{R}_{j,o_j d_j(s)}} \sum_{(a,b) \in \mathcal{R}_{j,o_j d_j(s)}} \frac{\delta_0 Dist_{j,ab}}{I^{\gamma}_{j,ab}}$$

s.t. $\forall m' \in \mathcal{Z}_j$,

$$\mathbb{1}\{m' = o\} - \mathbb{1}\{m' = d\} + \sum_{a \in \mathcal{Z}_j(m')} \mathbb{1}\{(a,m') \in \mathcal{R}_{j,o_j d_j(s)}\} \leq \sum_{b \in \mathcal{Z}_j(m')} \mathbb{1}\{(m',b) \in \mathcal{R}_{j,o_j d_j(s)}\}$$

The intuition of the routing problem is that firm j chooses the shortest efficiency-weighted route to travel from origin o_j to destination d_j in its own railroad network. Firm j obtains the optimal routing for each market from o_j to d_j by solving a linear programming problem.[38]

The optimal allocation decision is obtained by solving the cost minimization problem in equation 7. By taking the derivatives with respect to $I_{j,ab}$ for every arc $(a,b) \in \mathcal{A}_j$, the optimal allocation decision is solved as:

$$\begin{aligned} I_{j,ab} &= \left[\frac{\gamma}{\lambda_j} \sum_{s: j \in J(s)} \left(\delta_0 Dist_{j,ab} \cdot Q_{s,od} \cdot \mathbb{1}\{(a,b) \in \mathcal{R}_{j,o_j d_j(s)}\} \right) \right]^{\frac{1}{1+\gamma}} \\ &= \left[\frac{\gamma}{\lambda_j} \cdot \delta_0 Dist_{j,ab} \cdot q_{j,ab} \right]^{\frac{1}{1+\gamma}} \end{aligned} \tag{9}$$

where $q_{j,ab}$ is the total amount of shipment that run through arc (a,b), and λ_j is the La-

[38] To represent a linear-programming problem, the routing problem from origin o to destination d can be written in vectors as:

$$\min_{\boldsymbol{q}} \boldsymbol{c} \times \boldsymbol{q}$$

$$\text{such that } \nabla \boldsymbol{q} = \boldsymbol{Q_s}, \text{ where } \boldsymbol{Q_s} = \begin{bmatrix} \cdots \\ -Q_{s,od} \\ 0 \\ 0 \\ Q_{s,od} \\ 0 \\ 0 \\ \cdots \end{bmatrix}$$

grangian multiplier of the capital constraint of firm j.[39] Therefore, for any non-zero $I_{j,ab}$ and $I_{j,a'b'}$, from equation 9 we know that:

$$\frac{I_{j,ab}}{I_{j,a'b'}} = \left[\frac{Dist_{j,ab} \cdot q_{j,ab}}{Dist_{j,a'b'} \cdot q_{j,a'b'}}\right]^{\frac{1}{1+\gamma}} \tag{10}$$

Equation 10 shows that the optimal allocation decision is to allocate more resources to arcs that carry larger volume of traffic. Combine results in equation 10 with the capital constraint of firm j where $\sum_{(a,b)\in\mathcal{A}_j} I_{j,ab} = K_j$, I can obtain the solution of the optimal allocation problem.

To sum up, given the demand $Q_{s,od}(\tilde{p}_{s,od}, \tilde{p}_{-s,od})$ of each service in each origin-destination market, in the second stage firms make routing and allocation decision by solving the cost minimization problem in equation 7. The optimal routing of firm j is to choose the shortest efficiency-weighted route to travel from origin o_j to destination d_j, and the optimal allocation is to allocate more resources to arcs that carry larger volume of traffic.

In equilibrium, firms choose prices conditional on their perceived cost of themselves and their rivals. Given the prices, the demand for each service in each market is determined. Then, given the demand, firms make routing and allocation decision to minimize operating cost. Last, the outcome from firms cost minimization problem is consistent with the perceived cost of operating at the pricing stage.

7 Estimation

7.1 Demand Estimation

The utility function of consumer i choosing service s is specified as:

$$\begin{aligned} u_{is,od} = \alpha \cdot p_{s,od} &+ \beta_1 \cdot \log TravelDistance_{s,od} \\ &+ \beta_2 \cdot \log TotalTrackMiles_{s,od} + \xi_{s,od} + \varepsilon_{is,od} \end{aligned}$$

[39] The indicator function $\mathbb{1}\{(a,b) \in \mathcal{R}_{j,o_j d_j(s)}\} = 1$ if arc (a,b) is in the routing of firm j from o_j to d_j. Because the amount of allocation cannot be negative, the optimal level of allocation would be $I_{j,ab} = \max\{I_{j,ab}, 0\}$.

The demand estimation follows the standard procedure as in Berry (1994).[40]

$$\ln(S_{s,od}) - \ln(S_{0,od}) = \alpha \cdot p_{s,od} + \beta_1 \cdot \log TravelDistance_{s,od} \\ + \beta_2 \cdot \log TotalTrackMiles_{s,od} + \xi_{s,od}$$

$s_{s,od}$ is the market share of service s in serving market from o to d. $s_{0,od}$ is the market share of the outside option, which is the share of all other transportation mode serving market from o to d. Usually economists do not observe the share of outside option when estimating demand. However, in my case I do observe the share of outside option. The data on shipment by other transportation mode is obtained from the Commodity Flow Survey (CFS).[41]

Table 6 shows the estimated results using data of 2012:

Table 6: Results of Demand Estimation

	(1) OLS	(2) Fixed Effects	(3) Fixed Effects
Price	-0.077**	-0.302***	-0.282***
	(0.036)	(0.038)	(0.038)
Log Travel Distance	1.157***	-0.339***	-0.372***
	(0.058)	(0.107)	(0.107)
Log Amount of Track Miles	0.547***	0.330***	0.349***
	(0.045)	(0.050)	(0.053)
Constant	-14.757***	-2.889***	-3.142***
	(0.440)	(0.740)	(0.744)
Observations	4,481	4,481	4,481
R-squared	0.120	0.038	0.058
Number of marketID		1,420	1,420
Market FE		Yes	Yes
Firm FE			Yes

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Column (1) shows the OLS results. The coefficient of *price* is significantly negative, but the coefficient of *travel distance* is positive, which is counter-intuitive. The coefficient of *travel distance* is likely to be upward-biased in the OLS regression. This is because customers are more likely to use railroads in long-distance shipping since railroads are more efficient

[40] It is more efficient to estimate demand parameters and cost parameters together using GMM. Given the level of complexity of the firm's problem in my model, however, I choose to estimate demand parameters and cost parameters separately.

[41] CFS is conducted every five years by the U.S. Census Bureau. The data I have is in 1993, 1997, 2002, 2007, and 2012.

than other transportation mode like trucking in long-haul shipping. To correct for this kind of bias, and to control for unobservable market-specific demand factors, I control for O-D market fixed effects in columns (2). The coefficient of *travel distance* turns to be significantly negative after controlling for market fixed effects. This means that conditional on each O-D market, customers are less likely to choose service with longer travel distance. Price elasticity is also larger after controlling for market fixed effects. Last, there are concerns that there will be efficiency and other factors that are constant within railroad firms. To address this concern, I further control for firm fixed effects in column (3). The results in column (3) are very similar to results in column (2). In brief, results in Table 6 shows that conditional on each O-D market, customers are more likely to choose service with shorter travel distance and service with more physical tracks in the local market.

Given the estimated price coefficient to be -0.3, the average price elasticity of demand is -0.739 in my data. The estimated average price elasticity is comparable to the estimates in the transportation literature. For example, Oum (1989) estimated an own-price elasticity of rail demand for Canadian inter-regional freight of -0.598, Wilson et al. (1988) estimated a price elasticity of demand for US road freight grain transport of –0.73, and Beuthe et al. (2001) estimated the demand elasticities for long-distance trips ($>$ 300 km) of -0.63.[42]

Because I do not have a time subscript t in my demand parameters and in my model, one concern is that the demand parameters significantly changed throughout the studied period. To address this concern, I replicate my demand estimation using data in 1997, 2002, and 2007 (when the CFS data is available). Results in Appendix E shows that the demand parameters are relatively stable across the studied time period.

Given the estimated demand parameters, then I can solve the firms optimization problem and estimate the cost parameters.

7.2 Estimation of Cost Parameters

In section 6, the arc-level transportation cost of firm j at arc (a, b) is parameterized as

$$c_{j,ab} = \frac{\delta_0 Dist_{j,ab}}{I_{j,ab}^{\gamma}}$$

[42] Graham and Glaister (2004) conducted a nice summary of road traffic-related elasticity estimates in the transportation literature. Compared to the literature, my estimates of price elasticity of demand of -0.727 is within a reasonable range.

the arc-level transportation cost $c_{j,ab}$ depends on the distance between a and b, and depends on the amount of investment $I_{j,ab}$. The per-unit transportation cost of service s is:

$$\begin{aligned} C_{s,od} &= \sum_{j \in J(s)} \sum_{(a,b) \in \mathcal{R}_{j,o_j d_j(s)}} c_{j,ab} + \#_{interchanges} \cdot \eta \\ &= \sum_{j \in J(s)} \sum_{(a,b) \in \mathcal{R}_{j,o_j d_j(s)}} \left[\frac{\delta_0 Dist_{j,ab}}{I^{\gamma}_{j,ab}} \right] + \#_{interchanges} \cdot \eta \end{aligned}$$

Therefore, there are three cost parameters to be estimated, δ_0, γ and η. δ_0 captures the average shipment cost per efficient mile, γ captures the effectiveness of allocated resources, and η captures interchange cost. In Appendix F, I conduct some comparative statistics to show that when the value of parameters changes, how firms routing and allocation decision will be different. I show that when the value of γ is larger, the benefit of economy of scope is larger, hence firms have more incentive to consolidate traffic.

I use simulated GMM to estimate δ_0, γ, and η. I target five data moments in estimating the parameters:

Table 7: Targeted Data Moments

	Data Moments	Identification
Average shipping price per loaded car per mile	\$1.01	pin down δ_0
Change of price per loaded car per actual mile on travel distance (per 100 miles)	\$-0.17	pin down γ, δ_0
Average difference of price per loaded car between interconnecting route and non-interconnecting route	\$264.43	pin down η
Moments related to network measures		
Effect of degree centrality on price per loaded car per mile	\$-0.0002	pin down γ
Effect of betweenness centrality on price per loaded car per mile	\$-0.30	pin down γ

The identification argument is as follows: the first and second moment measures the intercept and the gradient of the effect that travel distance has on average shipping price. Because δ_0 captures the average shipment cost per efficient mile, the larger the value of δ_0, the larger the value of average shipping expense per mile hence the price. Therefore, the first two moments help pin down the value of δ_0. γ captures the effectiveness of allocated resources, and the value of γ largely affects the value of the last two moments regrading network measures. If there is no economy of scope ($\gamma = 0$), then only the travel distance between origin and destination matters for cost of shipping. Therefore, there is no benefit of consolidating traffic at all. Hence, the connectedness of the station and network structure will not matter regarding shipping expense. If there is economy of scope ($\gamma \neq 0$), as explained in the example of section 2.3, the network structure affects how traffic can be consolidated hence the level of economy of scope. For example, the higher the betweenness centrality of the station, the more the station benefits from economy of scope. Therefore, the effect of degree centrality and betweenness centrality on price helps identify parameter γ. Conditional on the value of δ_0 and γ, the parameter of interchange cost η is identified by the average difference of price between an interconnecting route and a non-interconnecting route. When the value of interchange cost η increases, the price difference between an interconnecting route and a non-interconnecting route increases.

Regarding the value of the data moments, average shipping price per loaded car per mile is \$1.01. According to AAR, in 2001 the average shipping price per mile for a loaded car is \$1.69, which is comparable to the number I observe in the data. In my data, if the total travel distance increases by 100 miles, the average shipping price per mile decreases by \$0.17, which is equivalent to a 16.8% decrease of average shipping price. The effect of betweenness centrality on price is calculated to be -\$0.3. To interpret this number, if we compare a station with the highest value of betweenness centrality to a station with the lowest value, the price difference is -\$0.2, which is about 20% difference in price.[43] The average difference of shipping price between an interconnecting route and a non-interconnecting route in the data is \$264.43 per loaded car. As a benchmark, in the data the average shipment price is \$1,034 per loaded car. It means that the interchange cost is equivalent to about 26% of average shipment price.

I estimate the cost parameters by minimizing the weighted distance between data and simulated moments. Table 8 shows the estimated results:

Table 8: Estimation Result of Cost Parameters

	Point Estimate
δ_0	1.8
γ	0.31
η	11

[43] The value of betweenness centrality is in the range of 0 to 1. The lowest value of betweenness centrality in the data is 0.02 and the highest value is 0.7. Therefore, the difference is calculated as $-\$0.3*(0.7-0.02) = -\$0.204$. Compared to the average shipping price \$1.01, the change is about 20%.

The estimated value of γ is less than 1. It indicates that the marginal return of allocating resource is decreasing. Therefore, railroad firms are more likely to allocate resources to multiple arcs, rather than stacking them in only few arcs. Table 9 compares the data moments with simulated moments under the estimated cost parameters.

Table 9: Compare Simulated Moments with Data Moments

	Data Moments	Simulated Moments
Average shipping price per loaded car per mile	\$1.01	\$0.96
Change of price per loaded car per actual mile on travel distance (per 100 miles)	\$-0.17	\$-0.03
Average difference of price per loaded car between interconnecting route and non-interconnecting route	\$264.43	\$257.40
Moments related to network measures		
Effect of degree centrality on price per loaded car per mile	\$-0.0002	\$-0.0006
Effect of betweenness centrality on price per loaded car per mile	\$-0.30	\$-0.32

The model does well in matching the average shipping price, but not the gradient of price change on travel distance. This might be because conditional on routing, my model has a restricted relation between travel distance and transportation cost. Therefore, there is not much flexibility in matching the curvature of the effect of distance on price. The simulated moment of average difference of shipping price between interconnecting route and non-interconnecting route matches the data moment well. Last, regarding moments related to network measures, the simulated moments are close to the data moments.[44]

8 Counterfactuals and Results

My model allows for origin-destination markets to be interdependent in the cost minimization stage, and the model has rich implications in understanding the change of shipment prices after mergers. First, I conduct merger simulations of railroad firms. I follow what happened in the history and simulate every single merger between Class I railroads in the U.S. that

[44]The standard errors of estimated parameters and results of sensitivity analysis are coming in the future version. I follow Andrews et al. (2017) in conducting the sensitivity analysis.

happened between 1985 to 2010.[45] For example, in simulating the merger between the Burlington Northern railway (BN) and the Santa Fe railway (SF), I fix the mass of demand of each origin-destination market, fix the network of railroad firms other than BN and SF, merge the network of BN and the network of SF, combine the capital owned by BN and SF (so the total capital owned by BNSF is the summation of capital of BN and SF), and assume the merged firm BNSF makes pricing, routing, and allocation decision in the newly merged network. Next, I solve for the new equilibrium prices after the merger. By comparing the change of costs and prices before and after the merger in each origin-destination market, I calculate what the effect of merger is. I repeat this process for all the mergers that happened in my studied period. Then given the counterfactual results, I summarize the role of network structure by looking at different centrality measures of a network, and how such measures are related to change of costs and change of markups after a merger.

8.1 Effect of Merger

The counterfactual analyses how prices, costs, and markups change after each merger. For example, below I show you the results of the merger of the Burlington Northern railway (BN) and the Santa Fe railway (ATSF).

[Insert Figure 8 here.]

Panel (a) of Figure 8 shows the network of the two merging firms. The map is color coded with purple showing the network owned only by BN, green showing the network owned only by SF, and yellow showing the overlapped region of the two networks. Panel (b) shows the estimated change of costs after the merger. The changes are aggregated to origins, and the size of the nodes represents the magnitude of changes. Color green means decrease, and color red means increase. Results in panel (b) show that change of costs has a large heterogeneity in different regions. Nodes on the borderline have a larger reduction of costs after the merger. This is because borderline nodes need to travel longer distance to other important markets, such as from LA to Chicago compared to from St. Louis to Chicago. Therefore, borderline nodes benefit more from elimination of interchange cost, and they benefit more from better options of routing after the merger. Panel (c) shows the change of markup, still red means markup increased after the merger. Regarding change of markup, nodes in the overlapped

[45] There were acquisitions between Class I railroads and smaller (Class II, Class III and shortlines) railroads in the studied time period. Because data regarding small railroads such as Class II, Class III and shortlines are not available, such mergers are not considered in my analysis. The complete merger history is shown in Figure A.1.

region have larger increase of mark-up. This is intuitive because the market became more concentrated in the overlapped region after the merger.

Results in figure 8 show that there is large degree of heterogeneity regarding change of costs and change of markups, depending on the location of markets and the topology of the network. Counterfactual results of other mergers also show that different network structures significantly impact the effect of mergers. In the next section, I summarize the role of network structure by looking at different centrality measures of a network, and how such measures are related to change of costs and change of markups after a merger.

8.2 Role of Network Structure

8.2.1 How Network Measures Affect Effects of Merger Within a Merger

By conducting the counterfactuals, I calculate the equilibrium prices and costs before and after each merger. I then regress those changes on centrality measures of the node. The regression results show how the value of centrality measures impact effects of merger.

The main regression is specified as follows:

$$y_{j,od} = \alpha + \beta_1 D_j + \beta_2 d_o(\mathcal{G}_j) + \beta_3 D_{j,od} \times d_o(\mathcal{G}_j) + \epsilon_{j,od} \tag{11}$$

where $y_{j,od}$ is either the shipment price, the shipment cost, or the markup for firm j at market from o to d. $D_{j,od}$ is an indicator of merger of firm j, and $d_o(\mathcal{G}_j)$ is the centrality measure of origin o. The centrality measure is defined as degree centrality, betweenness centrality, and weighted average of neighboring betweenness centrality (WANB) respectively.[46]

[Insert Table 10 here.]

Table 10 shows the results of regressing effects of merger on degree centrality. Degree centrality measures the total number of arcs connected to each node. The higher the value of degree centrality, the more connections a node has to the neighboring nodes (see figure C.1 for an illustration). Panel A shows the regression results on degree centrality of origin. Column (2) of panel A shows that the higher the value of degree centrality, the smaller the reduction of cost is after merger. This might be because that the more connected the origin is, the more routing options there are before the merger. Therefore, the particular O-D market benefits less from other options of routing after the merger, and the cost reduction is smaller compared to other O-D markets which have less routing options before the merger.

[46] I also run the regressions on the centrality measures of destination as a robustness check.

On the other hand, column (3) of panel A shows that the increase of markup is also smaller for origins that have larger degree centrality, because the coefficient of the interaction term is negatively significant. This shows that for more connected routes before the merger, the increased market concentration is less of a concern after merger. Regression results of merger effects on degree centrality of destination in panel B is similar to the results on degree centrality of origin in panel A.

[Insert Table 11 here.]

Table 11 shows the results of regressing merger effects on betweenness centrality. Betweenness centrality measures how frequently the node is found on the shortest path from an origin to a destination (the hubness of a node). Panel A of table 11 shows the regression results on the betweenness centrality of origin. Column (2) of panel A shows that the higher the value of betweenness centrality of the origin, the larger the cost reduction is after the merger. This is because the higher the value of betweenness centrality, the node is found on the shortest path of more O-D pairs. Therefore, firms will allocate more resources to those nodes after the merger because of the larger hubness of those nodes, hence the cost reduction will be larger. However, column (3) of panel A shows that the incremental of markup is larger for the O-D market if the betweenness of the origin is larger. Panel A also shows that the magnitude of the coefficient of the interaction term is much larger than the magnitude of the coefficient of the merger indicator. It means that the betweenness centrality of the origin has a large impact in determining merger effects. Panel B shows the regression results on betweenness centrality of destination, and the findings are consistent with our results in Panel A.

[Insert Table 12 here.]

Table 12 shows the results of regressing merger effects on the weighted average of neighboring betweenness centrality (WANB). Panel A of table 12 shows the regression results on the WANB of origin. Column (2) of panel A shows that the higher the value of WANB of the origin, the larger the cost reduction would be after the merger. The argument is similar to what we make above: the higher the value of WANB, the neighboring nodes are found on the shortest path of more O-D pairs. Therefore, firms will allocate more resources to those neighboring nodes after the merger because of the larger hubness of those nodes. Hence, the studied O-D market can route through the neighboring nodes and take advantage of the low transportation cost there. Column (3) of panel A suggests that however, the increase of markup will be larger if the WANB is larger.

Results above summarize the role of network structure in the merger of BN and SF. As a robustness check, I replicate the same analysis for all the mergers that happened in the studied period. The estimated coefficients are shown in figure 11. The results of the robustness check are consistent with our previous findings: if the degree centrality of a node is larger, both the reduction of cost and increase of markup are smaller after the merger. If the betweenness centrality or the WANB of a node is larger, both the reduction of cost and increase of markup are larger after the merger.

[Insert Figure 11 here.]

8.2.2 How Network Measures Affect Effects of Merger Across Mergers

To further study how topology of firms' network affect merger effects, in figure 12 I plot the average cost changes and markup changes with respect to average network measures of each merger:

[Insert Figure 12 here.]

Panel (a) and (b) of Figure 12 show the comparison between average cost changes after merger and average network measures (degree centrality and betweenness centrality) of merging firms. Panel (c) and (d) of Figure 12 show the comparison between average markup changes after merger and average network measures (degree centrality and betweenness centrality) of merging firms. Results of 12 show that across the 12 mergers from 1985-2005, mergers with higher average degree centrality have lower cost reduction, but higher increase of markup on average. Mergers with higher average betweenness centrality have higher cost reduction, and lower increase of markup on average. In brief, the results suggest that if the merged firm has higher average betweenness centrality and lower average degree centrality, the merger will result in higher reduction of cost and lower increase of markup. Given this result, we can evaluate the change of welfare following the mergers in the studied period.

Figure 13 shows the change of the average degree centrality and the change of the average betweenness centrality across time.

[Insert Figure 13 here.]

Panel (a) of figure 13 shows that the average degree centrality increased across time. This is because throughout the years, mergers between railroad firms caused the network of individual firm to be larger. As a result, the number of connected neighbors to each node increased (degree centrality) within a firm's network. Panel (b) shows that the average betweenness centrality decreased across time. This is because when individual firm's network became larger following the mergers, the importance of particular individual node decreased, because there are more alternative routes in a larger network. Therefore, a particular node is less frequently found on the shortest path from an origin to a destination. Hence the average betweenness centrality decreased. Given these changes regarding average network measures across time, and combine with our results above, it shows that mergers between railroad firms are generating lower reduction of cost and higher increase of markup in recent years.

9 Conclusion

I document evidence of improved cost efficiency following the wave of mergers in the U.S. railroad industry from 1980 to 2010. By conducting a reduced-form analysis with detailed route-level shipment data, I find that following the mergers, the ratio of switching hours to road-service hours decreases by 11% and the average haul length increases by 5%. Moreover, on average the number of interchanges decreases and use of unit trains increases following railroad mergers. Shipment prices decrease by 9.4% on average after the mergers, and interconnecting routes have the largest price reduction, 11%, of all the route types.

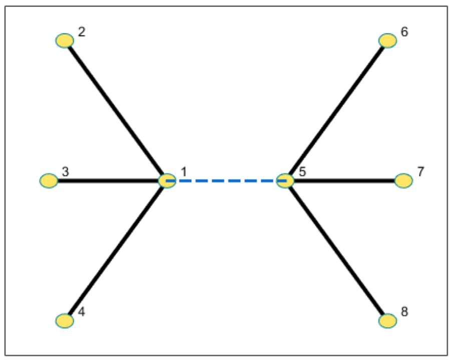
Using the reduced-form evidence and descriptive evidence from interviews and merger documents, I argue that looking solely at the effect of individual routes is not sufficient to understand efficiency gain in this industry. This is because each origin-destination market in the network is interdependent. To capture this important feature and examine how it affects consumer welfare following mergers, I propose an optimal transport network model by endogenizing firm pricing, routing, and allocation decisions. The simulation results show there is a large degree of heterogeneity of welfare changes in different geographic markets following railroad mergers, and it depends on the ownership of networks and the location of each market within the network. Results show that if the degree centrality of a node is larger, both the reduction of cost and increase of markup are smaller after the merger. If the betweenness centrality or the WANB of a node is larger, both the reduction of cost and increase of markup are larger after the merger. Throughout the studied period, the average degree centrality increased, and the average betweenness centrality decreased. Therefore, it indicates that mergers between railroad firms generated lower reduction of cost and higher increase of markup in recent years. The structure model can be further used in conducting various counterfactual results such as policy simulations and evaluation of future mergers.

The result of how cost efficiency and price incentives vary with respect to different topology that a network has, may also help us understand the tradeoff between efficiency gain and market concentration in other network industries, such as telecommunications and electricity, airline, and railroad industries.

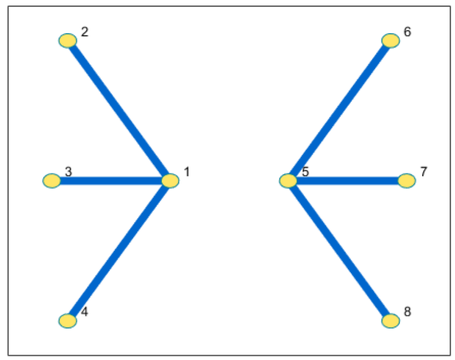
Certain limitations of this paper suggest areas for further exploration. First, the total amount of investment is fixed. One could look further at how the free capital market affects the predicted results by relaxing this assumption. Second, it is difficult to integrate two large networks into one. For example, when the Union Pacific railway acquires the Southern Pacific railway, integration of their networks is difficult and has a large negative impact on freight transportation in the short run. The trade-off between short-term chaos and long-term efficiency gain is not considered in my model. Moreover, when the number of firms decreases, firms may find it easier to collude. In the U.S. freight railroad industry there are more lawsuits alleging collusion by railroad companies after 2004, when the main mergers are finished. These features are not captured in my model but are worth exploring.

Figure 7: Illustrative Example

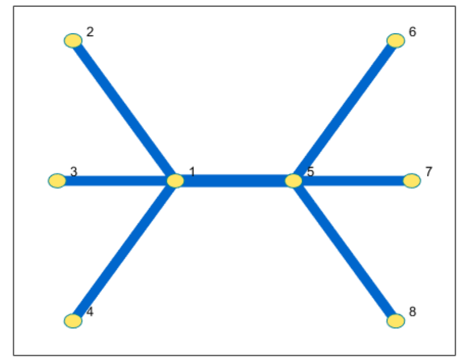
Note: Figure 7 shows the numerical example of two networks A and B. Panel (a) shows the network of A: there are two firms I and II in network A. Firm I owns nodes 1, 2, 3, 4. Firm II owns nodes 5, 6, 7, 8. There are six within-firm traffic, $1 \to 2$, $1 \to 3$, $1 \to 4$ within firm I, and $5 \to 6$, $5 \to 7$, $5 \to 8$ within firm II; and there are six across-firm traffic, $1 \to 6$, $1 \to 7$, $1 \to 8$, and $5 \to 2$, $5 \to 3$, $5 \to 4$. Each connected arc has a travel cost of 1, and an across-firm traffic incurs an extra interchange cost d. So, the cost of traveling from $1 \to 2$ is 1, and the cost of traveling from $1 \to 5$ is $2 + d$. Network B in panel (d) has exactly the same set-up. Panel (b) and (c) shows the traffic volume of network A when the interchange cost is 0.5 and 0.1 respectively. Panel (e) and (f) shows the traffic volume of network B when the interchange cost is 0.5 and 0.1 respectively. Comparing Panel (b) with panel (e) shows that Firm I and II have the exact same network in network A and network B, the only difference is how they are connected to each other. Think in the way that the “fork” is placed towards different direction in network A and network B.



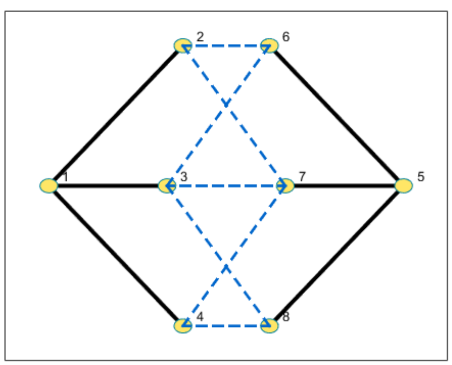
(a) Network A



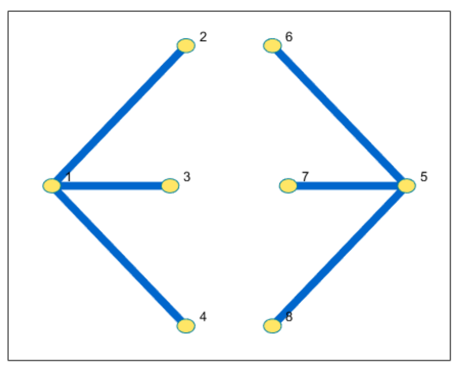
(b) Traffic Volume when interchange cost is 0.5



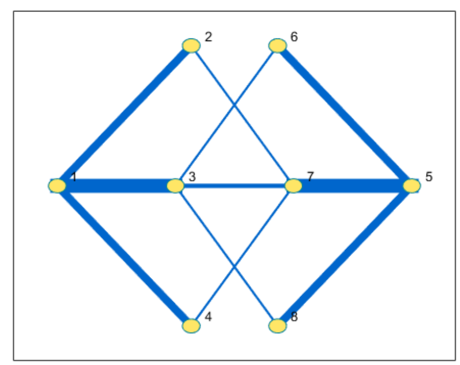
(c) Traffic Volume when interchange cost is 0.1



(d) Network B



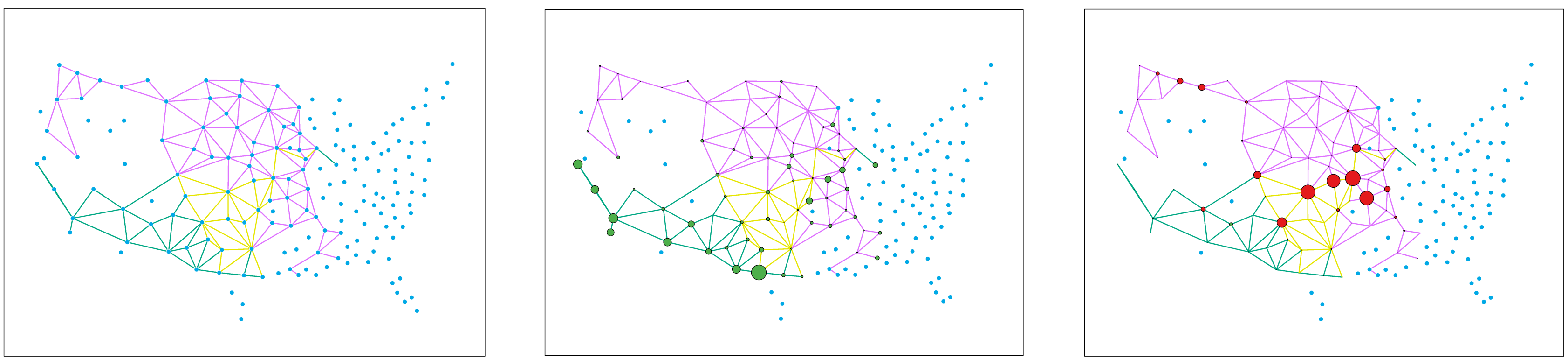
(e) Traffic Volume when interchange cost is 0.5



(f) Traffic Volume when interchange cost is 0.1

Figure 8: Changes of Transportation Cost and Markup

Note: Panel (a) shows the network of the two merging railroad firms the Burlington Northern Railway (BN) and the Santa Fe railway (SF). The map is color coded with purple showing the network owned only by BN, green showing the network owned only by SF, and yellow showing the network owned by both firms. Panel (b) shows the estimated cost changes following the merger of BN and SF. The changes are aggregated at origins, and the size of the nodes represents the magnitude of changes. Color green means reduction, color red means increase. Panel (c) shows the change of markup, still red is increased markup. First regarding cost changes, we can see that the magnitude of cost changes has a large heterogeneity in different regions. It seems that nodes on the borderline has a larger cost reduction. This is because borderline nodes need to travel longer distance to other important markets, like from LA to Chicago, Kansas City, and Minneapolis. So, it means borderline nodes benefit more from elimination of interchange cost, and they are more likely to have better routing options after the merger. Regarding change of mark-up, nodes in the overlapped region seem to have larger increase of mark-up. This result here basically illustrates the tradeoff between efficiency gain and change of market power, but this is only one merger case. If the topology of network changes, what will happen?



(a) Network of BNSF

(b) Changes of Cost after BNSF Merger

(c) Changes of Markup after BNSF Merger

Figure 9: Actual Railroad Networks in 2004

Note: Figure 9 illustrates the network of the seven Class I railroads in 2004: the Burlington Northern and Santa Fe railway (BNSF) competes with the Union Pacific railway (UP) in the west, the CSX Transportation (CSXT) and the Norfolk Southern railway (NS) competes in the east. Two Canadian Class I railroads the Canadian Pacific railway (CP) and the Canadian National railway (CN) connect freight shipment between Canada and U.S., and the Kansas City Southern railway (KCS) locates in the south and connects freight shipment between Mexico and the United States.

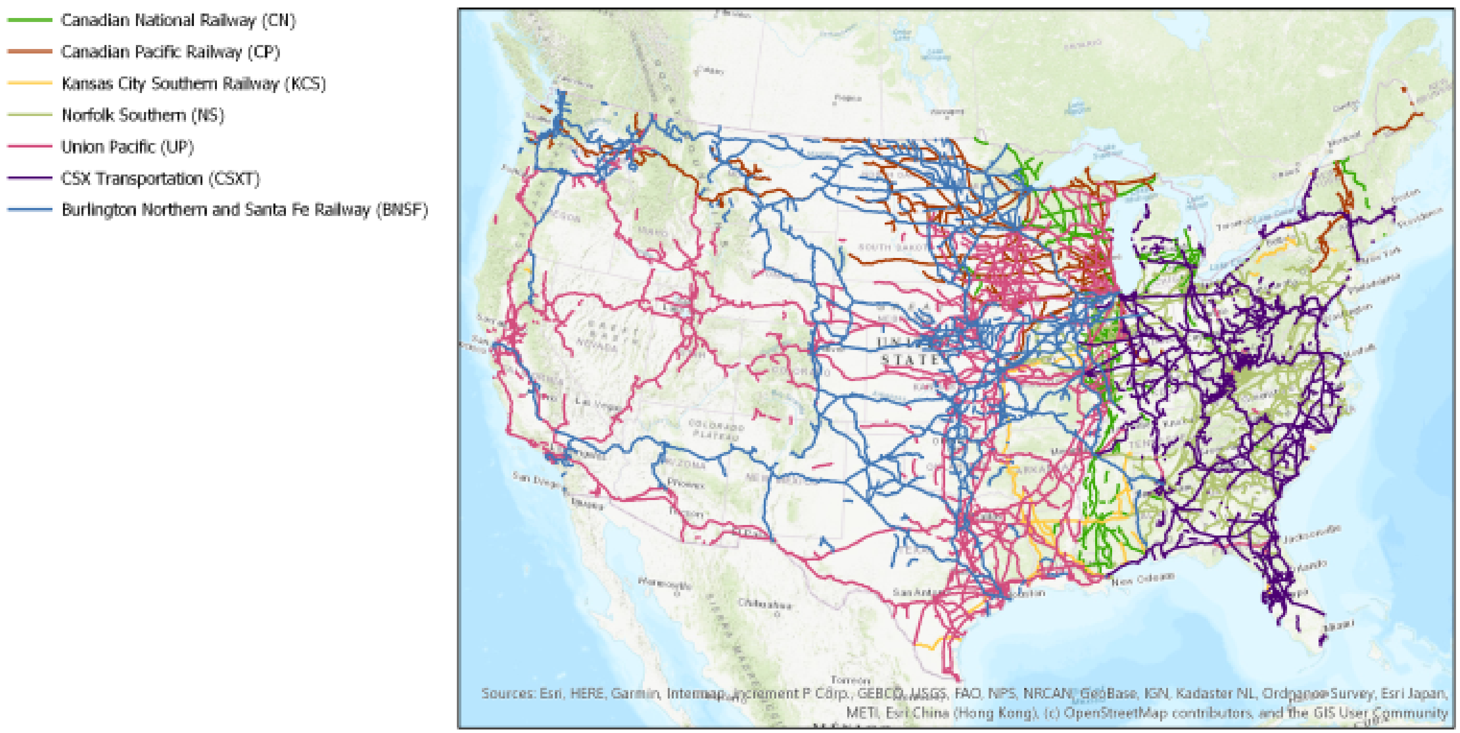
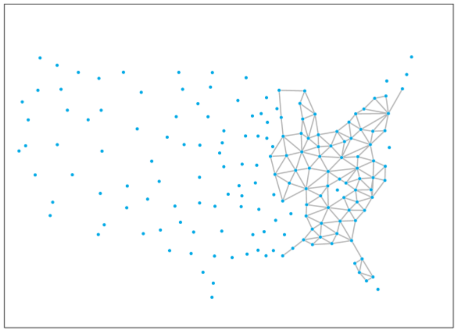
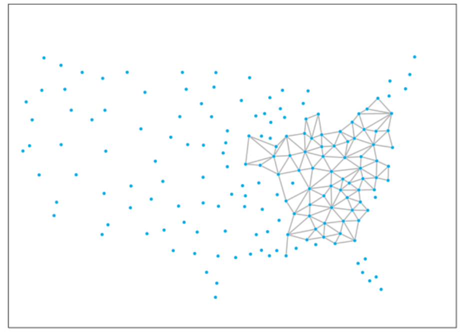


Figure 10: Virtual Networks in 2004

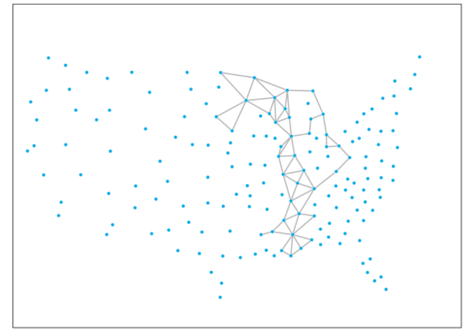
Note: Figure 10 shows the virtual networks we use for analysis. The virtual networks are built from the actual railroad networks shown in figure 9. Notice that this is only the networks of year 2004. In our analysis we build and use the virtual networks of every single firm of every year from 1985 to 2005.



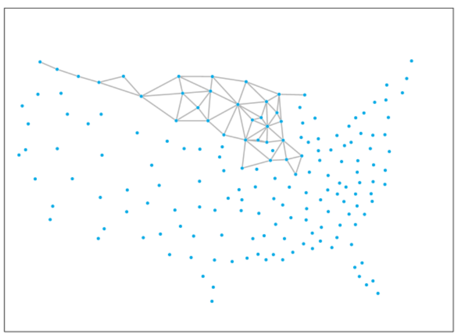
(a) Network of CSXT



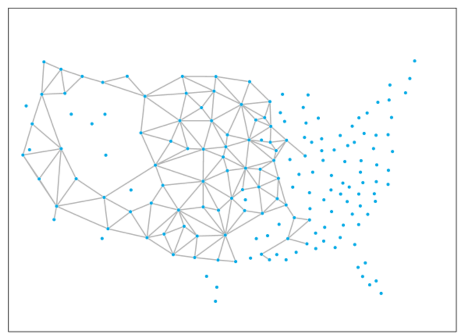
(b) Network of NS



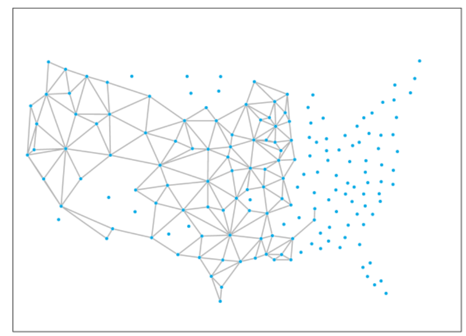
(c) Network of CN



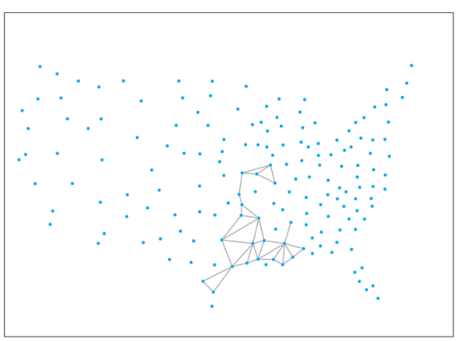
(d) Network of CP



(e) Network of BNSF



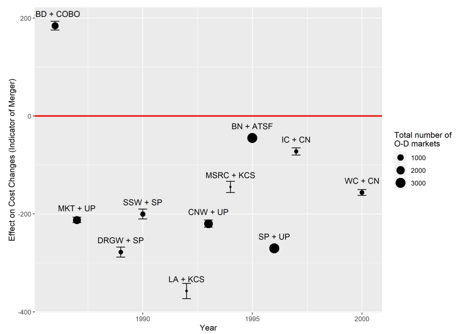
(f) Network of UP



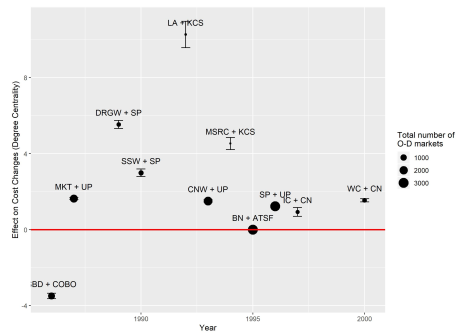
(g) Network of KCS

Figure 11: Estimated Coefficient across Mergers (Regress Merger Effects on Network Measures)

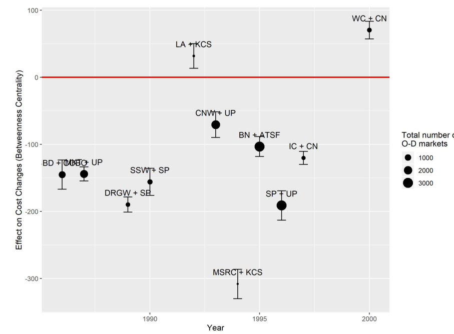
Note: Panel (a) to (c) in Figure 11 shows the estimated coefficients of regressing cost changes on indicator of merger, and the interaction with degree of centrality and interaction with betweenness of centrality respectively. Panel (d) to (e) shows the estimated coefficients of regressing markup changes on indicator of merger, and the interaction with degree of centrality and interaction with betweenness of centrality respectively. Size of the dot represents the total number of O-D markets in the particular merger.



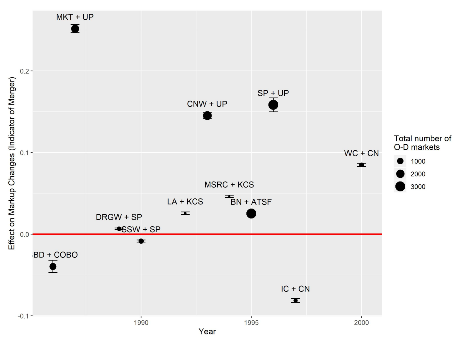
(a) Indicator of Merger on Cost Changes



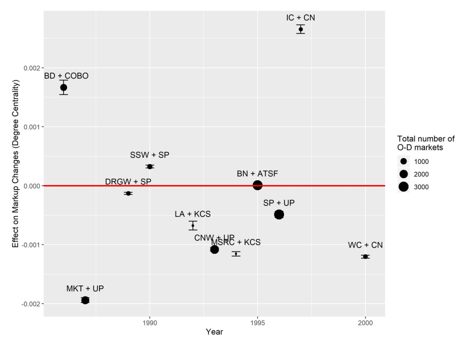
(b) Degree of Centrality on Cost Changes



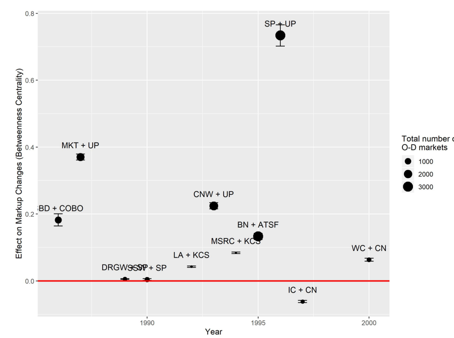
(c) Betweenness of Centrality on Cost Changes



(d) Indicator of Merger on Markup Changes



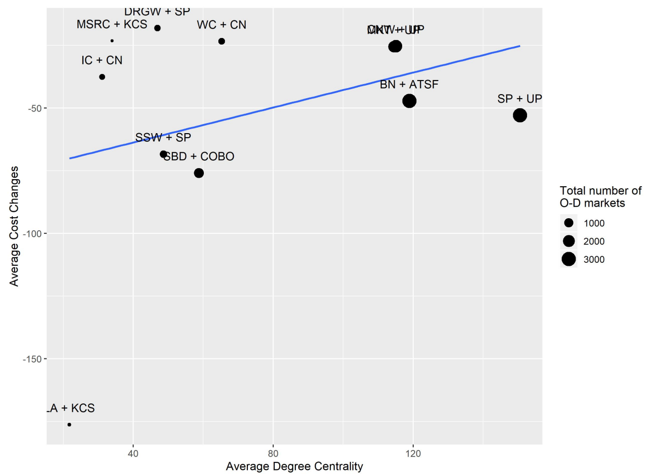
(e) Degree of Centrality on Markup Changes



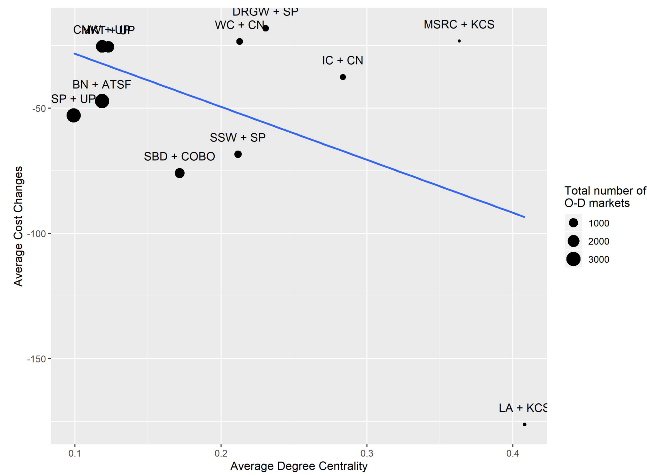
(f) Betweenness of Centrality on Markup Changes

Figure 12: Average Merger Effects v. Average Network Measures

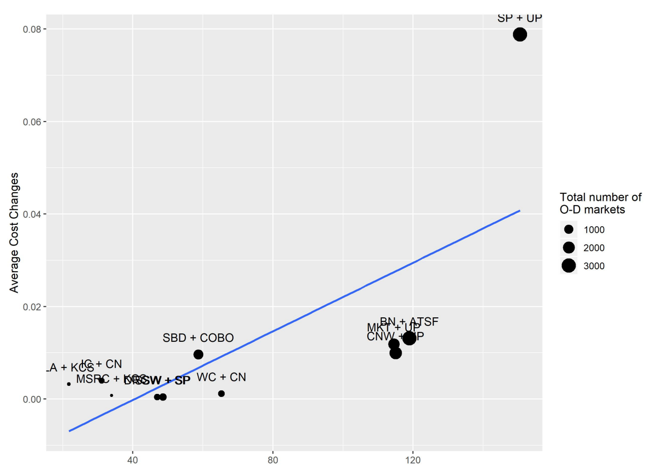
Note: Panel (a) and (b) of Figure 12 show the comparison between average cost changes after merger and average network measures (degree centrality and betweenness centrality) of merging firms. Panel (c) and (d) of Figure 12 show the comparison between average markup changes after merger and average network measures (degree centrality and betweenness centrality) of merging firms. Size of the dot represents the total number of O-D markets in the particular merger.



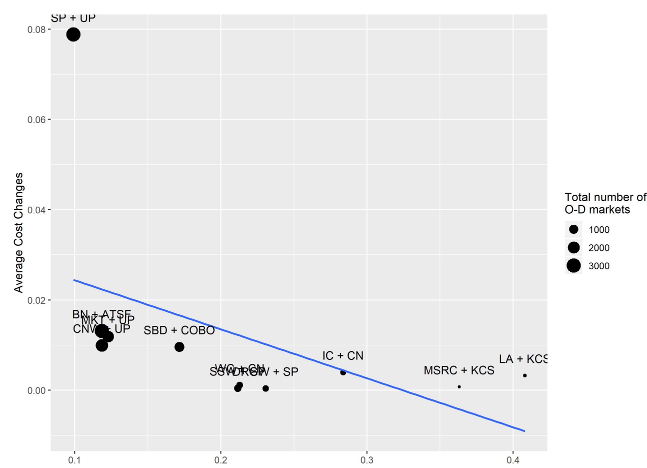
(a) Average Cost Changes v. Average Degree Centrality



(b) Average Cost Changes v. Average Betweenness Centrality



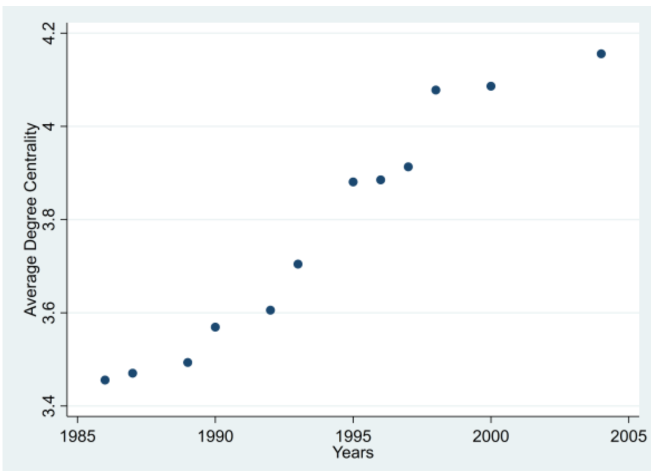
(c) Average Markup Changes v. Average Degree Centrality



(d) Average Markup Changes v. Average Betweenness Centrality

Figure 13: Change of Network Measures Across Years

Note: figure 13 shows the change of the average degree centrality across time in panel (a), and the change of the average betweenness centrality across time in panel (b). The two centrality measures are averaged over all firms within each year. Panel (a) shows that the average degree centrality increased across time. This is because throughout the years, mergers between railroad firms caused the network of individual firm to be larger. As a result, the number of connected neighbors to each node increased (degree centrality) within a firm's network. Panel (b) shows that the average betweenness centrality decreased across time. This is because when individual firm's network became larger following the mergers, the importance of particular individual node decreased, because there are more alternative routes in a larger network. Therefore, a particular node is less frequently found on the shortest path from an origin to a destination. Hence the average betweenness centrality decreased.



(a) Average Degree Centrality

(c) Average Betweenness Centrality

Table 10: How does Degree Centrality Affect Merger Effects

Note: Table 10 shows the results of regressing merger effects on degree centrality. The regression is at firm-market (firm-origin-destination) level. Degree centrality measures the total number of arcs connected to each node. The higher the value of degree centrality, the more connections a node has to the neighboring nodes. Panel A shows the regression results on degree centrality of origin, and Panel B shows the regression results on degree centrality of destination. Column (1) shows the regression results of price changes on degree centrality after merger. Column (2) shows the results of cost changes, and column (3) shows the results of markup changes. Interpretation of the results are in section 8.2.

Panel A: Regress merger effects on degree centrality of origin

	(1) Price of Shipment	(2) Cost of Shipment	(3) Markup
Indicator of Merger	-126.0***	-64.62***	0.0301***
	(2.579)	(1.322)	(0.00109)
Degree Centrality	0.00439	0.00226	-0.00000455
	(0.00666)	(0.00341)	(0.00000280)
Merger × Centrality Measure	0.275***	0.141***	0.0000851***
	(0.0213)	(0.0109)	(0.00000897)
Observations	144774	144774	144774

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel B: Regress merger effects on degree centrality of destination

	(1) Price of Shipment	(2) Cost of Shipment	(3) Markup
Indicator of Merger	-165.6***	-84.91***	0.0290***
	(2.511)	(1.287)	(0.00106)
Degree Centrality	0.00765	0.00393	-0.00000465
	(0.00645)	(0.00331)	(0.00000273)
Merger × Centrality Measure	0.644***	0.329***	0.0000966***
	(0.0209)	(0.0107)	(0.00000883)
Observations	144774	144774	144774

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: How does Betweenness Centrality Affect Merger Effects

Note: Table 11 shows the results of regressing merger effects on betweenness centrality. The regression is at firm-market (firm-origin-destination) level. Betweenness centrality measures how frequently the node is found on the shortest path from an origin to a destination (the hubness of a node). Panel A shows the regression results on betweenness centrality of origin, and Panel B shows the regression results on betweenness centrality of destination. Column (1) shows the regression results of price changes on betweenness centrality after merger. Column (2) shows the results of cost changes, and column (3) shows the results of markup changes. Interpretation of the results are in section 8.2.

Panel A: Regress merger effects on betweenness centrality (origin)

	(1) Price of Shipment	(2) Cost of Shipment	(3) Markup
Indicator of Merger	-76.06***	-39.03***	0.0252***
	(1.530)	(0.784)	(0.000643)
Betweenness Centrality	19.05***	9.767***	-0.000845
	(3.465)	(1.776)	(0.00146)
Merger × Centrality Measure	-149.5***	-76.64***	0.105***
	(8.874)	(4.547)	(0.00373)
Observations	144774	144774	144774

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel B: Regress merger effects on betweenness centrality (dest)

	(1) Price of Shipment	(2) Cost of Shipment	(3) Markup
Indicator of Merger	-63.12***	-32.41***	0.0268***
	(1.505)	(0.771)	(0.000634)
Betweenness Centrality	15.73***	8.062***	0.000109
	(3.363)	(1.723)	(0.00142)
Merger × Centrality Measure	-241.6***	-123.9***	0.0924***
	(8.578)	(4.396)	(0.00361)
Observations	144774	144774	144774

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: How does WANB Affect Merger Effects

Note: Table 12 shows the results of regressing merger effects on the weighted average of neighboring betweenness centrality (WANB). The regression is at firm-market (firm-origin-destination) level. Betweenness centrality measures how frequently the node is found on the shortest path from an origin to a destination (the hubness of a node). Panel A shows the regression results on WANB of origin. Panel B shows the regression results on WANB of destination, and panel C shows the regression results on WANB of interchange station of interchange lines. Column (1) shows the regression results of price changes on WANB after merger. Column (2) shows the results of cost changes, and column (3) shows the results of markup changes. Interpretation of the results are in section 8.2.

Panel A: Regress merger effects on WANB (origin)

	(1) Price of Shipment	(2) Cost of Shipment	(3) Markup
Indicator of Merger	-84.73***	-43.49***	0.0359***
	(2.215)	(1.135)	(0.000933)
WANB	-0.0382	-0.0152	0.000110
	(5.220)	(2.675)	(0.00220)
Merger × Centrality Measure	-83.85***	-42.95***	0.0270***
	(15.06)	(7.718)	(0.00634)
Observations	144774	144774	144774

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Panel B: Regress merger effects on WANB (dest)

	(1) Price of Shipment	(2) Cost of Shipment	(3) Markup
Indicator of Merger	-64.43***	-33.08***	0.0356***
	(2.182)	(1.118)	(0.000920)
WANB	14.94**	7.652**	0.000121
	(5.148)	(2.638)	(0.00217)
Merger × Centrality Measure	-238.1***	-122.0***	0.0293***
	(14.83)	(7.597)	(0.00625)
Observations	144774	144774	144774

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

References

Alexandrov, Alexei, Russell W Pittman, and Olga Ukhaneva, "Pricing of Complements in the US freight railroads: Cournot versus Coase," *US Department of Justice, Antitrust Division, Economic Analysis Group Discussion Paper EAG*, 2018, pp. 18–1.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M Shapiro, "Measuring the sensitivity of parameter estimates to estimation moments," *The Quarterly Journal of Economics*, 2017, *132* (4), 1553–1592.

Ashenfelter, Orley C, Daniel S Hosken, and Matthew C Weinberg, "Efficiencies brewed: pricing and consolidation in the US beer industry," *The RAND Journal of Economics*, 2015, *46* (2), 328–361.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The fall of the labor share and the rise of superstar firms," 2017.

Bazzoli, Gloria J, Anthony LoSasso, Richard Arnould, and Madeleine Shalowitz, "Hospital reorganization and restructuring achieved through merger," *Health Care Management Review*, 2002, *27* (1), 7–20.

Berry, Steven T, "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 1994, pp. 242–262.

Beuthe, Michel, Bart Jourquin, Jean-Francois Geerts, and Christian Koul à Ndjang'Ha, "Freight transportation demand elasticities: a geographic multimodal transportation network analysis," *Transportation Research Part E: Logistics and Transportation Review*, 2001, *37* (4), 253–266.

Borenstein, Severin, "Airline mergers, airport dominance, and market power," *The American Economic Review*, 1990, *80* (2), 400–404.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou, "Geography, Search Frictions and Endogenous Trade Costs," Report, National Bureau of Economic Research 2017.

Buchholz, Nicholas, "Spatial equilibrium, search frictions and efficient regulation in the taxi industry," Report, Technical report, University of Texas at Austin 2015.

Casavant, Ken, Eric Jessup, Marvin E Prater, Bruce Blanton, Pierre Bahizi, Daniel Nibarger, Johnny Hill, and Isaac Weingram, "Rail rate and revenue changes since the Staggers Act," in "Journal of the Transportation Research Forum," Vol. 50 2012.

Chapin, Alison and Stephen Schmidt, "Do mergers improve efficiency? Evidence from deregulated rail freight," *Journal of Transport Economics and Policy*, 1999, pp. 147–162.

Chen, M Keith, Judith A Chevalier, and Elisa F Long, "Nursing home staff networks and COVID-19," Technical Report, National Bureau of Economic Research 2020.

Ciliberto, Federico, Emily E Cook, and Jonathan W Williams, "Network Structure and Consolidation in the US Airline Industry, 1990–2015," *Review of Industrial Organization*, 2019, *54* (1), 3–36.

Dafny, Leemore, "Estimation and Identification of Merger Effects: An Application to Hospital Mergers," *Journal of Law and Economics*, 2009, *52.*

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

Dranove, David and Richard Lindrooth, "Hospital consolidation and costs: another look at the evidence," *Journal of health economics*, 2003, *22* (6), 983–997.

Eliason, Paul J, Benjamin Heebsh, Ryan C McDevitt, and James W Roberts, "How ownership changes affect behavior and outcomes: Evidence from dialysis facility acquisitions," Technical Report, Working paper 2018.

Fajgelbaum, Pablo D and Edouard Schaal, "Optimal transport networks in spatial equilibrium," *Econometrica*, 2020, *88* (4), 1411–1452.

Friebel, Guido, Gerard McCullough, and Laura Padilla Angulo, "Patterns of Restructuring The US Class 1 Railroads from 1984 to 2004," *Journal of Transport Economics and Policy (JTEP)*, 2014, *48* (1), 115–135.

Galichon, Alfred, *Optimal Transport Methods in Economics*, Princeton University Press, 2016.

Gallamore, Robert E and John Robert Meyer, *American Railroads*, Harvard University Press, 2014.

Graham, Daniel J and Stephen Glaister, "Road traffic demand elasticity estimates: a review," *Transport reviews*, 2004, *24* (3), 261–274.

Grieco, Paul L. E., Charles Murry, and Ali Yurukoglu, "The Evolution of Market Power in the US Auto Industry," Technical Report, Working paper 2020.

Grimm, Curtis and Clifford Winston, "Competition in the deregulated railroad industry: Sources, effects, and policy issues," *Deregulation of network industries: What's next*, 2000, pp. 41–71.

Harrison, E. Hunter, *How We Work and Why*, Canadian National Railway Company, 2005.

Igami, Mitsuru, "Estimating the innovator's dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981–1998," *Journal of Political Economy*, 2017, *125* (3), 798–847.

_ **and Kosuke Uetake**, "Mergers, Innovation, and Entry-Exit Dynamics: Consolidation of the Hard Disk Drive Industry, 1996-2015," 2016.

Jackson, Matthew O, *Social and economic networks*, Princeton university press, 2010.

Jeziorski, Przemysław, "Estimation of cost efficiencies from mergers: Application to US radio," *The RAND Journal of Economics*, 2014, *45* (4), 816–846.

Kim, E Han and Vijay Singal, "Mergers and market power: Evidence from the airline industry," *The American Economic Review*, 1993, pp. 549–569.

Krebs, Robert D, *Riding the Rails: Inside the Business of America's Railroads*, Indiana University Press, 2018.

Kwoka, John E and Lawrence J White, "Manifest destiny? The union Pacific-Southern Pacific merger," 1997.

Leeper, Sarah Elizabeth, "The game of radiopoly: An antitrust perspective of consolidation in the radio industry," *Fed. Comm. LJ*, 1999, *52*, 473.

McCullough, Gerard J, "US railroad efficiency: A brief economic overview," *Faculté d'économie appliquée, université du Minnesota, Minneapolis*, 2005.

McGuckin, Robert H and Sang V Nguyen, "On productivity and plant ownership change: New evidence from the longitudinal research database," *The RAND Journal of Economics*, 1995, pp. 257–276.

Oum, Tae Hoon, "Alternative demand models and their elasticity estimates," *Journal of Transport Economics and policy*, 1989, pp. 163–187.

Pesendorfer, Martin, "Horizontal Mergers in the Paper Industry," *RAND Journal of Economics*, 2003, *34* (3), 495–515.

Peters, Craig, "Evaluating the performance of merger simulation: Evidence from the US airline industry," *The Journal of law and economics*, 2006, *49* (2), 627–649.

Pittman, Russell W, "Railway Mergers and Railway Alliances: Competition issues and lessons for other network industries," 2009.

Prater, Marvin E, Ken Casavant, Eric Jessup, Bruce Blanton, Pierre Bahizi, Daniel Nibarger, and Isaac Weingram, "Rail competition changes since the Staggers Act," in "Journal of the Transportation Research Forum," Vol. 49 2012.

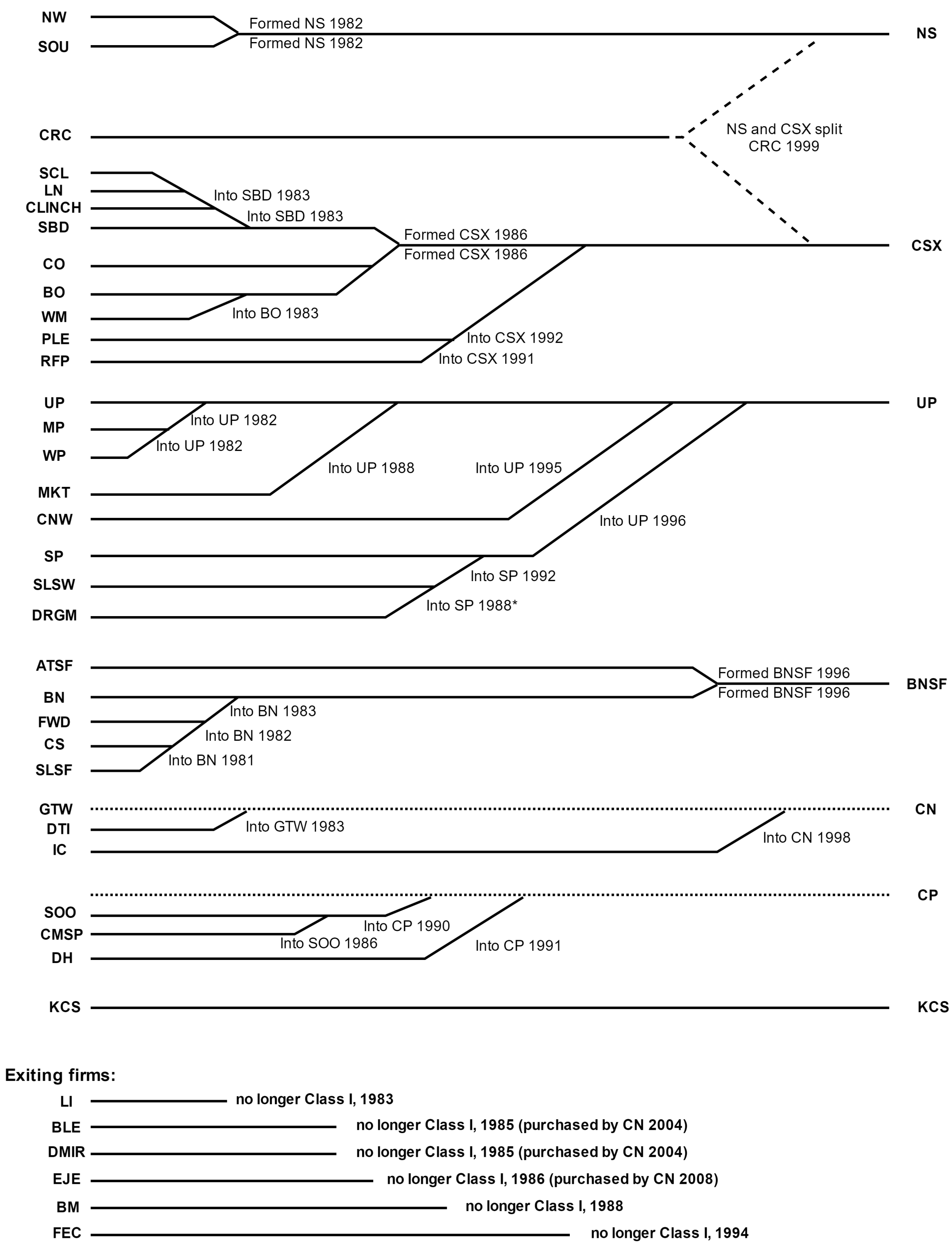
Williamson, Oliver E, "Economies as an antitrust defense: The welfare tradeoffs," *The American Economic Review*, 1968, *58* (1), 18–36.

Wilson, William W, Wesley W Wilson, and Won W Koo, "Modal competition in grain transport," *Journal of Transport Economics and Policy*, 1988, pp. 319–337.

A History of American Railroads

A.1 Merger History

Figure A.1: Merger History of Railroads



A.2 Regulation Changes in the U.S. Railroad Industry

Here I document a brief history of regulation changes in the U.S. railroad industry. The information is collected from multiple sources by the Surface Transportation board and other government resources.

History: 1887 - 1980

- 1887, the Interstate Commerce Act: create ICC value of service pricing (VOS pricing)
- 1973, The Regional Rail Reorganization Act ("3R" act): Establishment of US Railway Association, abandoning designated portions of the Northeast system
- 1976, Railroad Revitalization and Regulatory Reform Act ("4R" act): Creation of Conrail, permitting a railroad to adjust its rates up or down within a "zone of reasonableness", which is initially within 8 percent of the existing ICC tariff and widened over time. Accelerate the legal procedure dealing with abandoning unprofitable lines and expedite processing of merger
- 1980, The Staggers Act: The most important change is the removal of inefficient commodity rate regulation Enhanced the ability of abandoning lines and merge with others

Most recent: 1980 - current

- After the deregulation of 1980, ICC/STB no longer sets fixed prices for the railroad industry. Instead, it implements a constrained market pricing strategy, in which railroads are not allowed to set rates that are "too high". The STB does not have jurisdiction over the reasonableness of a rate for rail transportation unless the rail carrier providing the service has "market dominance". By statue, a necessary but not sufficient condition for a railroad to be considered to have market dominance is that the revenue produced by the rate is greater than 180% of its variable cost of providing the service as determined under the STB's Uniform Rail Costing System. When the rate goes beyond this 180% threshold, shippers are able to request STB to evaluate whether the service has "market dominance". There are three methods that STB allow shippers to use to evaluate market dominance of rail carriers: Stand-alone cost constraint (the most frequently used tools in law suits, invented in 1985), the three-benchmark procedure (invented in 1996), and the simplified-SAC (invented in 2007).
- 1985, ICC's Coal Rate Guidelines: ICC implements the requirement of constrained market pricing, in which the rate set by rail carriers needs to satisfy three constraints:
  - Revenue adequacy constraint: intended to ensure that railroads earn enough revenue to make normal profits, but not more (3 rate law cases invoked this principle since 1980 but all settled between shipper and railroad company)
  - Management efficiency constraint: prevent the shippers from paying avoidable costs that result from the inefficiency of the railroad (0 cases invoked this principle since 1980)

- Stand-alone cost constraint (SAC): Simulate the competitive rate that would exist in a contestable market by assuming a new highly efficient competitor railroad. The shipper must demonstrate that the "new" competitor would fully cover its costs, including a reasonable return on investment (full-SAC) (the most frequently used principle in rate cases. 50 rate cases invoked this principle after 1996, according to STB data base)

- 1995, ICC Termination Act
- 1996, The Three-benchmark procedure (only applies to cases where the total revenue of service is under \$1 million over five years)
  - Revenue shortfall allocation method, the uniform mark-up above variable cost that would be needed from every shipper in the captive group ($R/VC > 180$) to cover the URCS fixed cost
  - R/VC for comparative traffic
  - $R/VC_{>180}$ average captive price: calculate the average price of all the "captive" shippers

  Only 3 rate cases use three-benchmark from 1996 - 2007, while 25 rate cases use full-SAC in the same period.

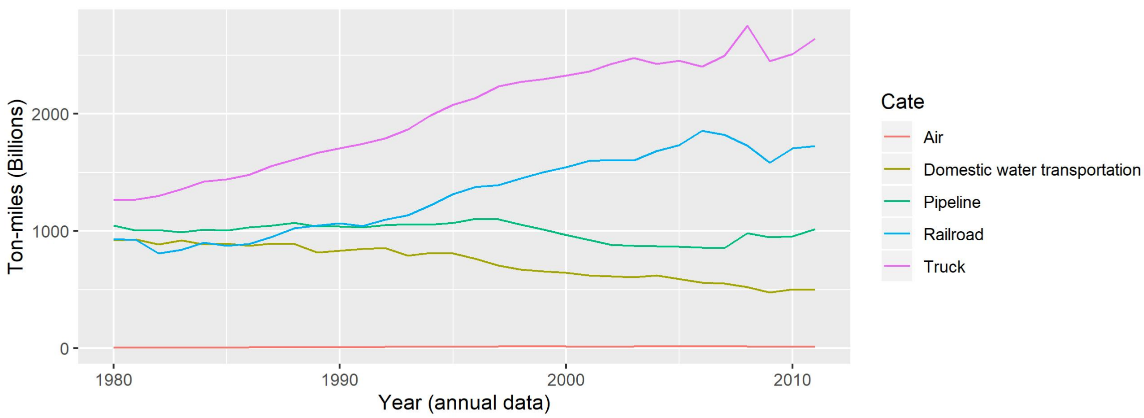
- 2007, Simplified SAC (only applies to cases where the total revenue of service is under \$5 million over five years): this allows shippers to use the existing infrastructure that serves the traffic, instead of coming up with a hypothetical stand-alone railroad to prove the market dominance of current service provider. Only 2 rate cases use simplified-SAC from 2007, while 20 cases use full-SAC in the same period
- 2013, Rate Regulations Reforms: remove limit of simplified-SAC, raise limit of three-benchmark to \$4 million (we see 6 rate cases after 2016, but all are using full-SAC method)

B Details of U.S. Railroad Industry

B.1 Industry Statistics

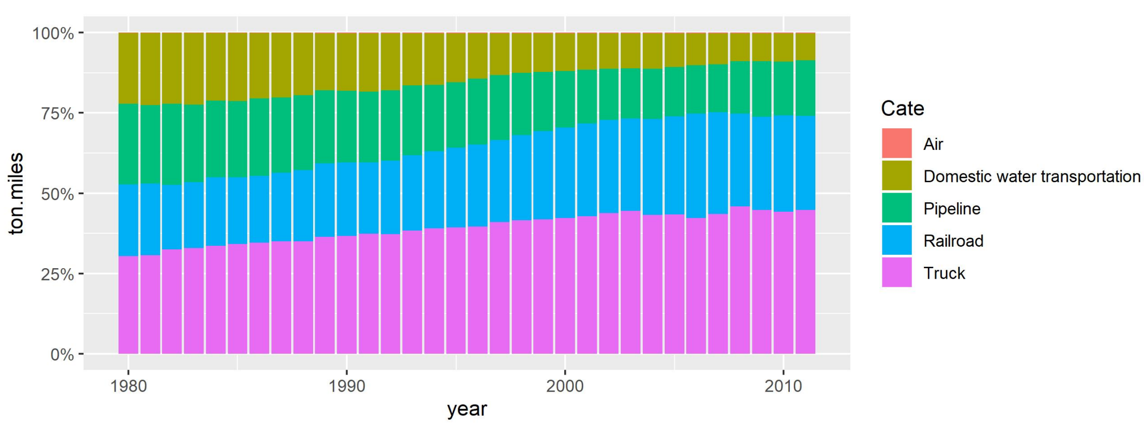
Figure B.1 and B.2 plots the total ton-miles of freight carried by mode, from 1980 to 2011 Importance of railroad industry among all the transportation mode. We can see that the U.S. railroad industry only accounts fo a small proportion of total ton-miles of freight (around 20%), and even accounts for a smaller proportion than pipelines at the start of 1980s. But it keeps increasing after the deregulation of 1980 and reaches 33% before the financial crisis. According to American Association of Railroads, if we only look at the intercity ton-miles, the railroad industry accounts for about 40 percent of the total shipping, more than any other transportation mode.

Figure B.1: U.S. total ton-miles of freight by mode



Source: U.S. Department of Transportation, Bureau of Transportation Statistics

Figure B.2: U.S. total ton-miles of freight by mode (percentage)



Source: U.S. Department of Transportation, Bureau of Transportation Statistics

A majority of the freight revenues in the U.S. freight railroad industry is generated by Class I railroads (see figure B.3, Class I railroads generate around 95% of the total freight revenues).

Figure B.3: U.S. Freight Railroad Industry 2012

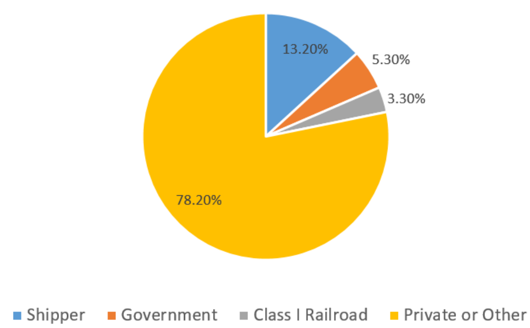
U.S. Freight Railroad Industry 2012

Railroad	Number	Miles	Employees	Revenue ($ bil)
Class I	7	95,264	163,464	$67.6
Regional	21	10,355	5,507	1.4
Local	546	32,858	12,293	2.6
Total U.S.	574	138,477	181,264	$71.6
Canadian	1	47		
Grand Total	575	138,524		

Source: Association of American Railroads, Analysis of Class I Railroads Annual Issues and Railroad Ten-Year Trends

Generally, the regional and local rails serve following purposes: first, shipping commodities from the main rails to end-users: for example, shipping coals from main rails to power plants; Second, serve as parking lot of railcars. A large proportion of regional and short lines are owned by leasing companies (Figure B.3) and serve this purpose; Third, used as interchanging places for different railroad companies to interchange the shippings.

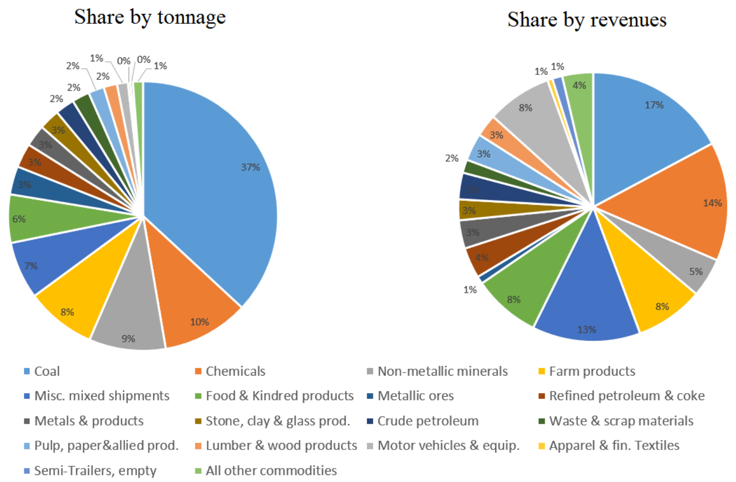
Figure B.4: Ownership of Regional and Short Lines



Source: American Short Line and Regional Railroad Association

The U.S. railroad industry ships various types of commodities, but the majority of carloads are generated by coal, chemical, and farm products. Although, the revenues generated by shipping coal are decreasing and the carloads of intermodal shippings (Misc. mixed shipments) are increasing nowadays.

Figure B.5: Carloads originated by commodity



Source: Association of American Railroads, Analysis of Class I Railroads Annual Issues and Railroad Ten-Year Trends

B.2 Documentation of Interviews

Below I document the interviews I conducted with people working in this industry (under confidentiality review, details to be filled in later):

1. Interview with Terminal Superintendent at Conrail, about how interchange works and why it is costly

2. Interview with Train & Terminal Operations Manager at Lake State Railway Company, about why interchange is costly and the incentive problem in exchanging equipment with another railroad

3. Interview with business development manager of Canadian National:

    - How do firms make pricing decision?

      - The pricing department gets an estimate of operational cost from the Costing department about how much money it costs to serve each origin-destination market. Then based on these cost estimates, the pricing department maximizes profits by charging a reasonable price margin as much as the market allows.

      - (The downward spiral) The service of a particular origin-destination market will be reduced if the operational cost outweighs the generated profits. However, sometimes this happens only because the operational cost is mismeasured. For example, the actual miles run by the train may not necessarily be fully related to the service it is providing. As a consequence, once a service is reduced, the volume of shipment decreases thus the operational cost further increases on a per-car basis, and more services get reduced.

    - How is interchange contract negotiated?

      - Usually the origin railroad has the bargaining power, but it depends. For example, there was a time when CN needs to make some shipment from Vancouver to New York, and they ask for a quote from the connecting railroad on shipment from Buffalo to New York. However, the Marketing representative from the other railroad only agree to give quote from Chicago to New York, rather than from Buffalo to New York, in order to maximize their revenue. "The hot stuff of one person is not the hot stuff of the other."

Documentation of importance of routing and network planning:

1. From former president of Southern Pacific and former CEO of Burlinton Northen Santa Fe, Robert Krebs, *Riding the Rails*

2. From former CEO of Canadian National and former CEO of CSX Transportation, E. Hunter Harrison, *How We Work and Why*

3. From Chief Operating Officer of CSX Transportation, Cindy Sanborn, *Reciprocal Switching, Opening Comments of CSX Transportation, INC.*

B.3 More Summary Stats of Waybill Data

Table B.1: Summary Statistics of Market Competition (Annually)

Year	Number of Waybills	Percentage of Interchange Lines	Number of Competitors in an O-D Market			Number of O-D Market (at BEA-to-BEA level)
			mean	25th percentile	75th percentile	
1984	262,626	41%	3	1	3	12,135
1985	262,703	41%	3	1	3	12,088
1986	276,177	38%	3	1	3	11,907
1987	300,324	35%	3	1	3	11,957
1988	322,257	35%	3	1	3	11,905
1989	324,936	36%	3	1	3	11,846
1990	323,570	35%	2	1	3	11,835
1991	314,705	32%	2	1	3	11,583
1992	346,632	31%	2	1	3	11,695
1993	373,868	29%	2	1	3	11,849
1994	426,092	27%	2	1	3	11,899
1995	453,802	26%	2	1	3	11,632
1996	457,505	25%	2	1	3	11,510
1997	473,070	23%	3	1	3	11,740
1998	496,856	20%	2	1	3	11,675
1999	524,856	15%	2	1	3	11,573
2000	544,738	14%	2	1	2	11,732
2001	522,927	14%	2	1	2	11,514
2002	535,722	13%	2	1	2	11,381
2003	554,967	13%	2	1	2	11,473
2004	580,572	12%	2	1	2	11,474
2005	611,033	11%	2	1	2	11,611
2006	632,748	11%	2	1	2	11,327
2007	611,421	10%	2	1	2	11,025
2008	568,584	10%	2	1	2	10,964
2009	477,526	10%	2	1	2	10,242
2010	533,364	10%	2	1	2	10,485

Source: STB, Carload Waybill Sample

Table B.1 shows that the year-by-year change tells the same story as in Table 2. First, total number of waybills in the waybill sample (the waybill sample is 2% of total waybills) changed from around $263,000$ to $533,000$ from year 1985 to 2010. This shows that total volume of railroad shipment doubled from 1985 to 2010. This is consistent with the story I show in figure 1. Meanwhile, the percentage of interchange lines decreased from 41% to 10% while the total traffic volume doubled. It shows that following the wave of mergers from 1985 to 2010, there is a significant decrease of interchanges. The number of O-D markets is relatively stable across the years, with a small decrease from $11,835$ to $11,611$ from year 1990 to 2005. Therefore, the change of extensive margin after the mergers does not seem to be a big concern. Last, the average number of competitors in each O-D market slightly decreased from 3 to 2 from 1985 to 2010. It indicates that firms conduct oligopolistic competition in the local markets.

C Illustration of Network Measures

Here is how to understand the four network measures. First, figure C.1 shows the average degree centrality of each firm in year 2004. Panel (a) shows the average degree centrality (average number of neighbors of each node) of the seven railroad firms in 2004. We can see that NS has the highest average degree centrality of about 5 while CN has the lowest average degree centrality of about 3. Panel (b) and (c) show the network of NS and CN respectively. We can see that NS are more "connected" than CN, in the sense that for each node in NS, there are more connections to other adjacent nodes.

[Insert Figure C.1 here.]

Figure C.2 shows the average betweenness centrality of each firm in year 2004. Intuitively, average betweenness centrality measures the availability of "alternative routes" in traveling among node-pairs within a network. Panel (a) shows that KCS has the highest average betweenness centrality of 0.2, which means 20% of all shortest routes between node-pairs within the network of KCS pass through the same nodes. Put this in a real context, it means something like 20% of all traffics within the network of KCS needs to go through Kansas City. Panel (a) also shows that UP has the lowest betweenness centrality of 0.06. Therefore, the hubness of nodes in UP is much smaller than the hubness of nodes in KCS. But it also means that there are less alternative routes to travel among node-pairs within the network of KCS. To better explain this point, I will show a more extreme example in figure C.3, where I compare the network of UP and CN in 1986.

[Insert Figure C.2 here.]

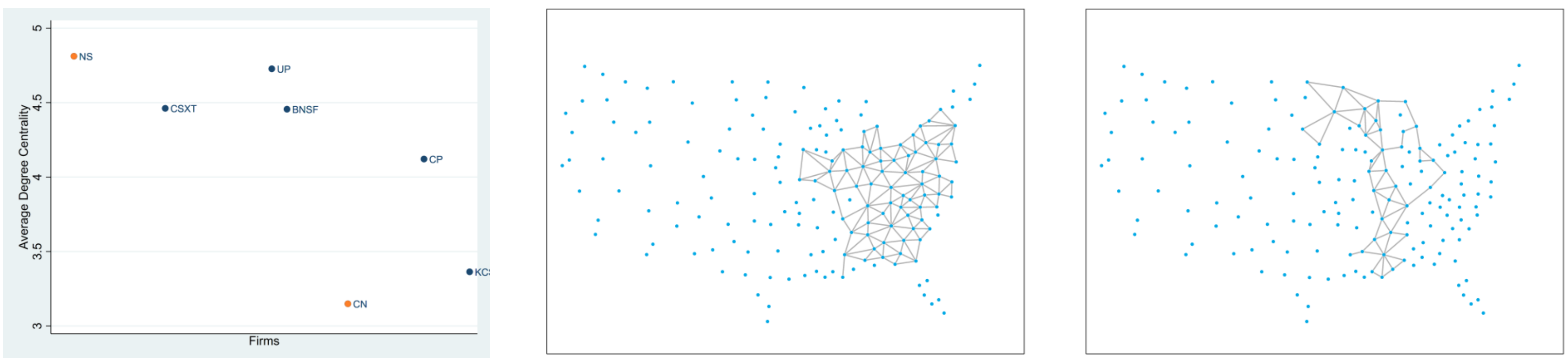
In the example of figure C.3, UP has an average betweenness centrality equals to 0.1 and CN has an average betweenness centrality equals to 0.4, four times larger. It means among all the node pairs within CN's network, 40% of the routes go through the same nodes on average. By comparing panel (a) and panel (b) in figure C.3, you can see that the network of CN is quite "linear". That is to say, the more "linear" the network is, the larger the average betweenness centrality will be.

[Insert Figure C.3 here.]

Changes over the years. You may already notice that from 1986 to 2004, the betweenness centrality of UP decreased from 0.1 to 0.06. Figure 13 shows how degree centrality and betweenness centrality change across years. Throughout the studied period, the average degree centrality increased, and the average betweenness centrality decreased. This is because waves of mergers between railroad firms caused the network of individual firms to be larger and larger. As a result, the number of connected neighbors to each node increased (degree centrality), but the importance of particular individual node decreased (betweenness centrality), because there are more alternative routes in a larger network. In brief, the individual network of each firm expanded in the past years, and the dependence of routing on a single node decreased.

Figure C.1: Average Degree Centrality of Firms in 2004

Note: figure C.1 shows the average degree centrality of each firm in year 2004. Panel (a) shows the average degree centrality (average number of neighbors of each node) of the seven railroad firms in 2004. We can see that NS has the highest average degree centrality of about 5 while CN has the lowest average degree centrality of about 3. Panel (b) and (c) show the network of NS and CN respectively. We can see that NS are more "connected" than CN, in the sense that for each node in NS, there are more connections to other adjacent nodes.



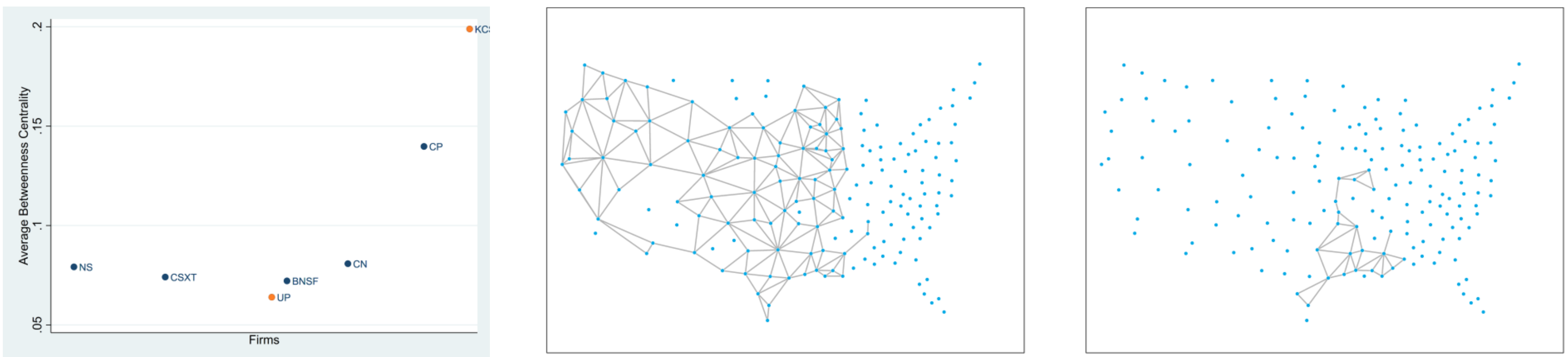
(a) Average Degree Centrality

(b) Network of NS

(c) Network of CN

Figure C.2: Average Betweenness Centrality of Firms in 2004

Note: figure C.2 shows the average betweenness centrality of each firm in year 2004. Intuitively, average betweenness centrality measures the availability of “alternative routes” in traveling among node-pairs within a network. Panel (a) shows that KCS has the highest average betweenness centrality of 0.2, which means 20% of all shortest routes between node-pairs within the network of KCS pass through the same nodes. Put this in a real context, it means something like 20% of all traffics within the network of KCS needs to go through Kansas City. Panel (a) also shows that UP has the lowest betweenness centrality of 0.06. Therefore, the hubness of nodes in UP is much smaller than the hubness of nodes in KCS. But it also means that there are less alternative routes to travel among node-pairs within the network of KCS.



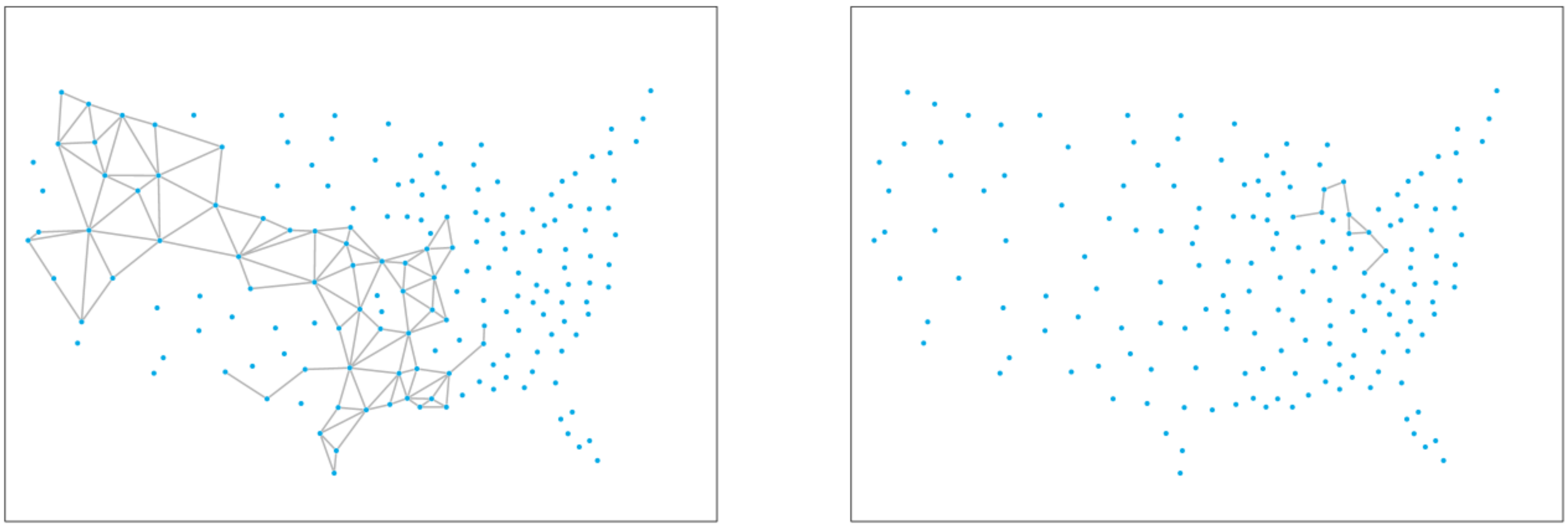
(a) Average Betweenness Centrality

(b) Network of UP

(c) Network of KCS

Figure C.3: Average Betweenness Centrality: UP vs. CN in 1986

Note: figure C.3 aims to show how network structure affects average betweenness centrality. UP has an average betweenness centrality equals to 0.1 and CN has an average betweenness centrality equals to 0.4, four times larger. It means among all the node pairs within CN's network, 40% of the routes go through the same nodes on average. By comparing panel (a) and panel (b) in figure C.3, you can see that the network of CN is quite "linear". That is to say, the more "linear" the network is, the larger the average betweenness centrality will be.



(a) Network of UP (with average Betweenness Centrality= 0.1)

(b) Network of CN (with average Betweenness Centrality= 0.4)

D Robustness Check for Reduced-form Analysis

First as a robustness check of the price effect of mergers, I run the price regression for each type of commodity. Table D.1 shows a complete summary statistics of commodities shipped by rail from waybill data.

Table D.1: Descriptive Statistics of Commodity Types and Car Ownership Category

	Number of Waybills	Percentage
Commodities		
Field Crops	466,584	3.85%
Forest Products	5,361	0.04%
Marine Products	2,138	0.02%
Metallic Ores	93,371	0.77%
Coal	1,002,580	8.28%
Crude Petroleum	2,855	0.02%
Nonmetallic Minerals	371,109	3.06%
Ordnance or Accessories	1,838	0.02%
Food or Kindred Products	882,352	7.28%
Tobacco Products	1,222	0.01%
Textile Mill Products	9,533	0.08%
Apparel or Other Textile Products	46,414	0.38%
Lumber or Wood Products	487,386	4.02%
Furniture or Fixtures	34,101	0.28%
Pulp, Paper or Allied Products	483,980	4.00%
Newspapers and Books	15,933	0.13%
Chemicals	635,119	5.24%
Petroleum or Coal Products	158,794	1.31%
Rubber or Miscellaneous Plastics Products	62,202	0.51%
Leather Products	2,484	0.02%
Clay, Concrete, Glass or Stone Products	323,923	2.67%
Primary Metal Products	354,360	2.93%
Fabricated Metal Exc.	24,387	0.20%
Machinery Exc.	23,351	0.19%
Electrical Machinery	70,893	0.59%
Transportation Equipment	1,098,439	9.07%
Instruments, Optical Goods	3,192	0.03%
Miscellaneous Products	21,965	0.18%
Waste or Scrap Materials	342,374	2.83%
Miscellaneous Freight Shipments	60,474	0.50%
Containers	660,513	5.45%
Mail	43,970	0.36%
Freight Forwarder	3,689	0.03%
Shipper Association	48,529	0.40%
Miscellaneous Mixed Shipments	3,434,269	28.35%
Small Packaged Freight Shipments	62,495	0.52%
Waste Hazardous	7,329	0.06%
Other	762,855	6.30%
Car Ownership Category		
Privately Owned	5,349,791	44%
Railroad Owned	3,621,221	30%
Trailer Train	2,202,838	18%
Non-Categorized	939,731	8%
Waybills (Carrier-Origin-Destination-Date)	12,113,581	

Source: STB, Carload Waybill Sample

Then I study effect of merger on price changes case-by-case. Table D.2 shows the estimation result of shipment price changes. The result suggests that on average a railroad merger reduces the shipment price by 9.4%. If we look at the merger effect case by case, we find that most of the large mergers result in a price reduction of more than 10%, such as the merger of the Burlington Northern and Santa Fe, the merger of the Southern Pacific and Union Pacific, and the merger of the Chicago and North Western Railway (CNW) and Union Pacific. The only exception is the merger of the Seaboard System Railroad (SBD), Chesapeake and Ohio Railway (CO), and Baltimore and Ohio Railroad (BO) which occurred in 1986. Mergers involving smaller railroad firms have an insignificant impact on shipment price, likely because these mergers affect only a small fraction of routes.

Table D.2: Effect of Merger on Price Change

	(1) Log Price	(2) Log Price
Indicator of Merger	-0.094***	
	(0.014)	
SBD		0.106***
		(0.021)
BNSF		-0.114***
		(0.023)
LA		-0.043
		(0.058)
MSRC		0.052
		(0.059)
IC		-0.025
		(0.041)
CNW		-0.162***
		(0.039)
MKT		0.009
		(0.044)
DRGW		0.018
		(0.043)
SP		-0.119***
		(0.021)
SSW		-0.227***
		(0.040)
WC		0.006
		(0.058)
Log Weight	-0.259***	-0.260***
	(0.015)	(0.015)
Private Railcars	-0.112***	-0.110***
	(0.009)	(0.009)
Trailer Train Railcars	-0.052***	-0.053***
	(0.009)	(0.009)
Observations	12,110,107	12,110,107
Number of marketID	22,510	22,510
Adjusted R-squared	0.361	0.363
Year FE	Y	Y
Firm FE	Y	Y
Commodity FE	Y	Y
O-D Route FE	Y	Y

Standard errors in parentheses, clustered at O-D route level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Surface Transportation Board, Carload Waybill Sample

As a robustness check, I run the price regression for each type of commodity (defined in STCC):

Table D.3: Effect of Merger on Price Change (by Commodities)

	(1) Field Crops	(2) Metallic Ores	(3) Coal	(4) Nonmetallic Minerals	(5) Food or Kindred Products	(6) Apparel or Textile Products	(7) Lumber or Wood Products	(8) Furniture or Fixtures	(9) Pulp, Paper	(10) Newspapers
Indicator of Merger	-0.009	0.048	-0.179***	-0.036	-0.052***	0.049	0.016	-0.023	-0.013	-0.057
	(0.014)	(0.062)	(0.028)	(0.029)	(0.014)	(0.047)	(0.013)	(0.036)	(0.015)	(0.042)
Observations	466,222	93,316	1,002,552	371,035	882,066	46,409	487,275	34,095	483,952	15,933
Number of marketID	6,982	1,086	1,360	3,697	10,766	1,210	8,145	1,694	8,441	780
Adjusted R-squared	0.178	0.144	0.251	0.234	0.266	0.767	0.274	0.829	0.299	0.667
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
O-D Route FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

	(11) Chemicals	(12) Petroleum	(13) Plastics Products	(14) Clay, Concrete, Stone Products	(15) Primary Metal Products	(16) Fabricated Metal Exc.	(17) Machinery	(18) Electrical Machinery	(19) Transportation Equipment	(20) Miscellaneous Products
Indicator of Merger	-0.114***	-0.115***	-0.056**	-0.045***	-0.112***	-0.104***	-0.035	-0.146***	-0.022	0.022
	(0.031)	(0.024)	(0.028)	(0.015)	(0.027)	(0.035)	(0.037)	(0.038)	(0.035)	(0.035)
Observations	634,684	158,774	62,197	323,910	354,322	24,385	23,343	70,889	1,097,641	21,963
Number of marketID	10,462	4,175	2,087	6,670	6,639	2,009	2,072	1,996	7,177	1,163
Adjusted R-squared	0.157	0.210	0.730	0.262	0.163	0.558	0.413	0.569	0.321	0.765
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
O-D Route FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

	(21) Waste or Scrap Materials	(22) Miscellaneous Freight Shipment	(23) Containers	(24) Mail	(25) Shipper Association	(26) Miscellaneous Mixed	(27) Small Packaged	(28) 49
Indicator of Merger	-0.000	-0.103*	0.049	0.036*	-0.179***	-0.170***	0.019	-0.072***
	(0.022)	(0.055)	(0.052)	(0.019)	(0.035)	(0.034)	(0.041)	(0.018)
Observations	341,973	60,443	660,061	43,965	48,523	3,434,108	62,487	762,745
Number of marketID	7,843	2,775	2,747	903	1,245	5,821	633	9,562
Adjusted R-squared	0.195	0.412	0.426	0.611	0.468	0.569	0.670	0.224
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
O-D Route FE	Y	Y	Y	Y	Y	Y	Y	Y
O-D Route Cluter	Y	Y	Y	Y	Y	Y	Y	Y

The results show that price reduction following railroad mergers is consistent across different types of commodities. If we look particularly at commodities that are largely shipped by rail, such as coal, chemicals, and construction materials (clay, concrete, etc.), there is a large and significant price reduction following railroad mergers.

E Robustness Check for Demand Estimation

The first concern is that the demand parameters changed a lot across years. To address this concern, I replicated the demand estimation in year 1997, 2002 and 2007 (when CFS data is available). Table E.1 shows the estimated results of 2007.

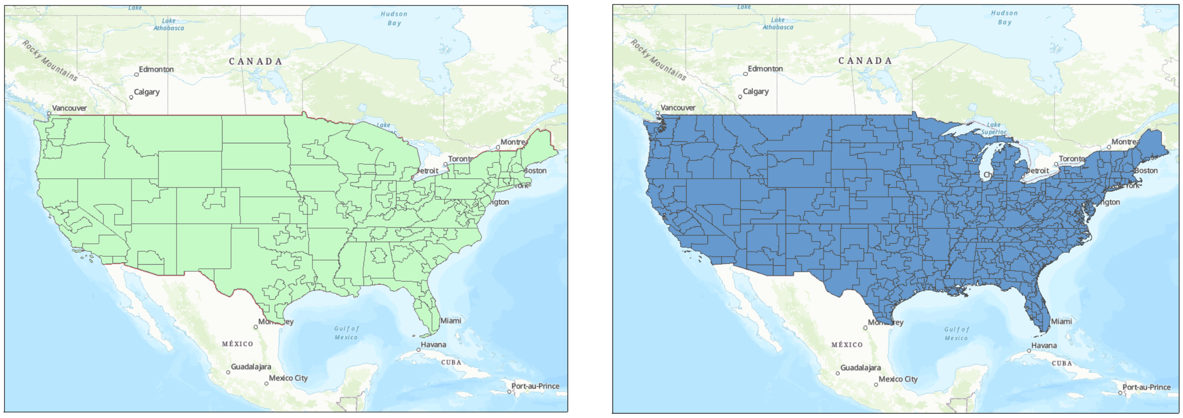
Table E.1: Results of Demand Estimation Using 2007 Data

	(1)	(2)	(3)
	OLS	Fixed Effects	Fixed Effects
Price	-0.280***	-0.419***	-0.410***
	(0.032)	(0.031)	(0.031)
Log Travel Distance	1.340***	-0.179**	-0.178**
	(0.051)	(0.087)	(0.086)
Log Amount of Track Miles	0.535***	0.408***	0.378***
	(0.038)	(0.041)	(0.044)
Constant	-15.373***	-4.156***	-4.250***
	(0.379)	(0.602)	(0.614)
Observations	5,323	5,323	5,323
R-squared	0.152	0.069	0.083
Number of marketID		1,440	1,440
Market FE		Yes	Yes
Firm FE			Yes

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In my main demand estimation, I define each market to be state. The definition of market is broader than what I use in the model. In the model I define market as BEA area to BEA area rather than state to state. The reason I used the definition of state to state in demand estimation is because of data limitation of CFS data. To address the concern that state-to-state level demand parameters are significantly different from BEA-to-BEA level, I replicated the main analysis in finer definition of market, as in CFS area to CFS area. CFS area is the finest definition of an area in the CFS data. As a comparison, there are 50 states, 170 BEA areas, and 130 CFS areas. Figure E.1 illustrates the CFS areas and BEA areas in a topographic map. Shipment information at CFS area level is only available in 2012, therefore I only do this analysis using 2012 data.

Figure E.1: CFS area v. BEA area



(a) CFS areas of 2012 (131 areas in total)

(b) BEA areas (170 areas in total)

Table E.2: Results of Demand Estimation in CFS Area

	(1) OLS	(2) Fixed Effects	(3) Fixed Effects
Price	0.114***	-0.211***	-0.198***
	(0.021)	(0.028)	(0.028)
Log Travel Distance	1.396***	-1.070***	-1.013***
	(0.036)	(0.132)	(0.133)
Log Amount of Track Miles	0.286***	0.175***	0.188***
	(0.025)	(0.037)	(0.040)
Constant	-11.862***	5.853***	5.091***
	(0.258)	(0.888)	(0.893)
Observations	11,825	11,825	11,825
R-squared	0.158	0.037	0.050
Number of marketID		7,422	7,422
Market FE		Yes	Yes
Firm FE			Yes

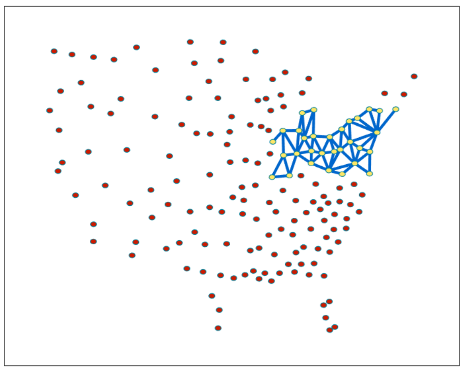
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table E.2 shows the estimated results of demand in CFS area. The estimated coefficients are very similar to those estimated in the main specification in Table 6. Therefore, it means that using a broader definition of market does not significantly change the results in demand estimation.

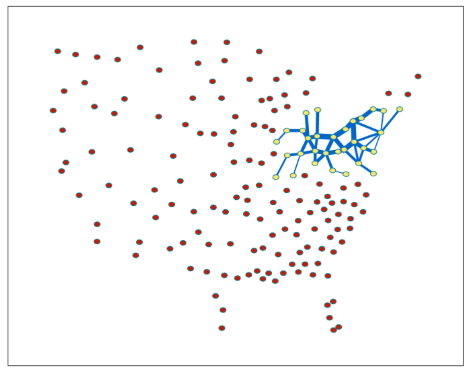
F Comparative Statistics

Figure F.1: How value of γ affects allocation and routing decisions.

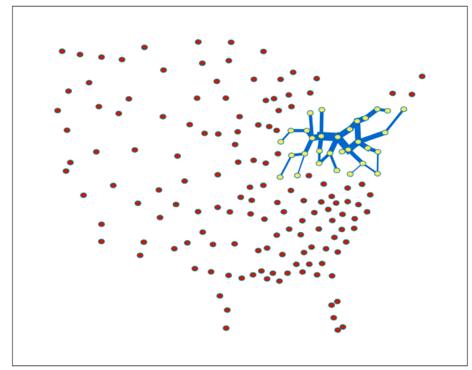
Note: Figure F.1 shows how the routing and allocation decision of the Conrail changes when the value of γ changes. Panel (a) to (c) shows the allocation decision and the thickness of the line shows the amount of resources. Panel (d) to (f) shows the routing decision of origin 3 to destination 7 when the value of γ changes. When $\gamma = 0$, there is no economy of scope, and resources are evenly attributed to all arcs owned by Conrail in panel (a), and routing of 3 to 7 takes the shortest route in distance $3 \to 10 \to 7$ in panel (d). When γ is non-negative, there is economy of scope so it is more cost-efficient to consolidate traffic and allocate more resources to routes with larger traffic volume. From panel (b) to (c), we see that when γ increases the allocation is more consolidated to “major” routes. Regarding routing, when γ increase, the optimal routing is no longer the shortest route in distance, because the routing now takes advantage of the lower cost of going through “major” routes. From panel (d) to (c), the optimal routing changed from $3 \to 10 \to 7$ in panel (d) to $3 \to 10 \to 8 \to 7$ in panel (e) and to $3 \to 10 \to 9 \to 8 \to 7$ in panel (f).



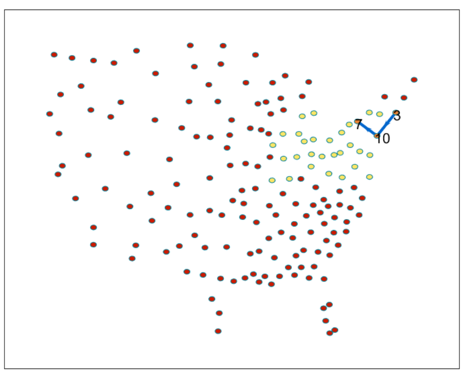
(a) Allocation of resources, $\gamma = 0$



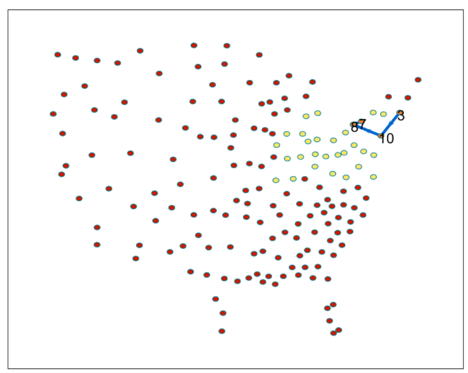
(b) Allocation of resources, $\gamma = 0.5$



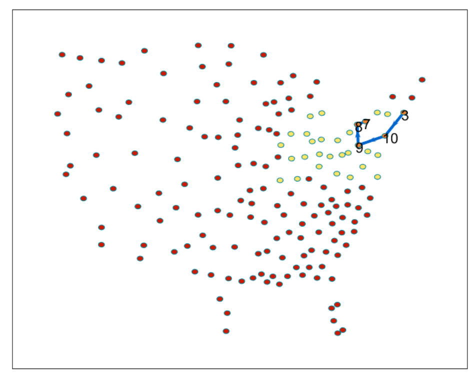
(c) Allocation of resources, $\gamma = 0.85$



(d) Routing of $3 \to 7$, $\gamma = 0$



(e) Routing of $3 \to 7$, $\gamma = 0.5$



(f) Routing of $3 \to 7$, $\gamma = 0.85$