RAILROADS, THEIR REGULATION, AND ITS EFFECT ON EFFICIENCY AND COMPETITION

by

TAYLOR KLAGES MCKENZIE

A DISSERTATION

Presented to the Department of Economics
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2017

DISSERTATION APPROVAL PAGE

Student: Taylor Klages McKenzie

Title: Railroads, Their Regulation, and Its Effect on Efficiency and Competition

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Economics by:

Wesley W. Wilson	Chair
Van Kolpin	Core Member
Keaton Miller	Core Member
Jeremy Piger	Core Member
Diane Del Guercio	Institutional Representative

and

Scott L. Pratt	Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2017

DISSERTATION ABSTRACT

Taylor Klages McKenzie

Doctor of Philosophy

Department of Economics

June 2017

Title: Railroads, Their Regulation, and Its Effect on Efficiency and Competition

Railroads have been subject to federal regulation since 1887. Due to the development of competing modes of transportation and changes in types of products being shipped, regulation began to impede efficiency and viability of firms, leading to partial deregulation of the industry in 1980. Partial deregulation allowed railroads to reduce costs, notably through mergers and line abandonment, which were aggressively pursued following deregulation and led to dramatic efficiency gains. However, concerns remain over increased consolidation, lack of competition in the industry, and the ability of firms to continue to realize efficiency gains. This dissertation investigates more recent developments in the rail industry with an eye towards regulation's effect and role.

I begin with a study into the markups of price over marginal cost and elasticities of scale in the rail industry. Scale elasticities provide information on where firms are operating on their average cost curves, and markups provide a more theoretically appealing method of examining pricing behavior than the revenue-to-variable-cost measure currently used by regulators. I extend previously developed methods to identify markups and scales for each firm and in each year. I find prices well in excess of marginal cost, and evidence firms are operating near minimum efficient scale, indicating efficiency gains from deregulation may be fully realized.

I then present a study that examines productivity changes in the rail industry and the role of technological change. I extend stochastic frontier frameworks to allow productivity and the state of technology to evolve flexibly through time and vary across firms. I find firms turn towards technological innovation to realize productivity gains when other channels previously offered by deregulation are not available.

I finish with a study of allocative errors in the rail industry. I again extend a stochastic frontier model to include differences in production across firms and allow allocative errors to be correlated with competitive pressures. I find that incorporating flexibility into the description of firm production is crucial for obtaining unbiased estimates of allocative errors, overcapitalization is prevalent in the rate-regulated rail industry, and additional competition does not appear to reduce inefficiency.

This dissertation includes unpublished co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: Taylor Klages McKenzie

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
Willamette University, Salem, OR

DEGREES AWARDED:

Doctor of Philosophy, Economics, 2017, University of Oregon
Master of Science, Economics, 2013, University of Oregon
Bachelor of Arts, Economics and Mathematics, 2012, Willamette University

AREAS OF SPECIAL INTEREST:

Industrial Organization
Econometrics

GRANTS, AWARDS AND HONORS:

Ph.D. Research Paper Award, University of Oregon, 2015

ACKNOWLEDGEMENTS

I am extremely grateful for all of the opportunities and incredible people who have enabled me to achieve this goal. I could not be more thankful to all of the professors who made this experience thoroughly enjoyable, to my thoughtful and supportive advisor, Wes Wilson, and to all my classmates and friends who made the entire time great.

I am extraordinarily lucky to have a supportive family who puts up with far more than they should from me and who I could not have done this without. My parents, who are always there to support and encourage, Baba and Ammu for all their love, and my adoptive siblings, Boru Apu and Wafiq.

Finally, I could not be more grateful for my wife, Shamina, for everything she is and continues to be.

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

The railroad industry has had a long history of regulation that once promoted competition but over time, changes in the types of products being shipped and development of competing modes of transportation came to impede efficiency and innovation. Eventually, this led to bankruptcies and concerns about economic viability. While deregulation of the industry has certainly resulted in efficiency gains and the survival of the industry, massive consolidation and abandonment of unprofitable routes has resulted in less competition between firms and has left many shippers with few transportation options. This dissertation investigates the railroad industry in the context of its regulation to determine the gains earned through partial deregulation as well as negative effects caused by reduced competition.

This dissertation begins with an analysis of markups and scale elasticities for railroads that have differentiated networks that was co-authored with University of Oregon professor Dr. Wesley Wilson. Using a random coefficients framework, I extend the model of Klette (1999) to allow markups and scales to vary flexibly both across firms and through time and estimate the model using Bayesian methods. I find evidence that price significantly exceeds marginal costs and scale estimates that point to constant or increasing returns to scale. Further, I find that there are important differences in markups and scales both across firms and over time. Increases in track investment created excess capacity and increased returns to scale for Burlington Northern Santa Fe (BNSF) and CSX; while BNSF has maintained investment and its excess capacity, CSX appears to have filled its excess capacity by 2012.

I then present a study analyzing productivity changes in the railroad industry and decompose them into changes due to technological change and those due to other factors such as rail line abandonment. I extend the stochastic frontier framework to include time-varying parameters to both allow productivity to follow a flexible process unique to each railroad and to capture changes in the production technology over time and again estimate the model from a Bayesian approach. I first find that productivity has shown little growth in the industry since 1999, with the exception of the Canadian National railway (CN) and the Kansas City Southern railway (KCS), which both experienced sizable increases in productivity. The source of productivity growth varies across firms. On the one hand, BNSF, the Soo Line railway, and Union Pacific (UP) were all able to realize growth through innovations in production, but had little ability to increase productivity through other channels. On the other hand, firms like CN and KCS that still had the ability to abandon lines found large increases in productivity due to factors other than innovation, but tended to stagnate or even experience losses with respect to changes in the production technology.

Finally, in the third essay, I examine whether there are allocative inefficiencies and the effects of competitive pressures on allocative errors. As noted by Tsionas (2002), it is important to control for differences in productive capabilities across firms when estimating technical inefficiency because those differences would otherwise be attributed to how efficiently firms can transform inputs into output. Similarly, differences in input productivity across firms will have an impact on how errors in allocation of inputs are estimated; however, I am not aware of any published research that investigates the importance of incorporating differences in production across firms in obtaining unbiased estimated of allocative inefficiency. Next, as noted by Leibenstein (1966), firms that don't face sufficient competition may not only lack the incentive to keep prices low but also to minimize costs, resulting in so-called "X-inefficiencies." Precise allocation of inputs can reduce production costs, which can be important when firms face competition; however, firms that don't face competitive pressure don't need to rely on low costs of production to attract customers and may invest less in the allocative process. After conducting a review of the history of the industry and the literature relevant to study of inefficiency and its application to the railroad industry, I also develop and present my models that test for the importance of controlling for differences in production and for the existence of X-inefficiencies. I find that it is crucial to allow production to vary flexibly across firms to obtain consistent estimates of inefficiency, and that increased market power appears to decrease allocative errors on average, providing evidence against the existence of X-inefficiencies in the rail industry.

CHAPTER II

MARKUPS AND SCALE ELASTICITIES FOR DIFFERENTIATED RAIL NETWORKS

This chapter was co-authored with University of Oregon professor Dr. Wesley Wilson. Dr. Wilson provided initial motivation for this project and was instrumental with his in-depth institutional knowledge of the rail industry. From this impetus I extended the theoretical and empirical methodology to provide more granular analysis of markups and scales than any published of which I'm aware. I also carried out statistical estimation and wrote this chapter using those results.

Abstract

In this chapter, I develop and estimate a model that provides both markups and scale elasticities that vary across railroads and through time for the traffic on their networks. My model is based on a framework provided by Hall (1988) and Klette (1999) wherein markups and scale elasticities are estimated from production relations. In my model, I aggregate the shipments over each firm's network, which provides a mapping from inputs and network and shipment characteristics to aggregate outputs over the network. Markups and scale elasticities are taken to follow a multivariate distribution. This allows for differences in markups and scale across firms and through time, but also for covariances across firms in markups and scale. I estimate the model with Bayesian methods to find markups that are generally well in excess of marginal costs and scale elasticities that generally point to increasing or constant returns in the industry.

"With 90% of U.S. rail freight now controlled by only four companies, shippers claim the giants have unfairly banded together. Unapologetic railroads refuse to back down. An epic battle of business vs. business." (Fortune, March 16, 2015)

Introduction

The rail industry was in financial ruin in the 1970s, which led to partial deregulation of the industry in 1980. Since then, there have been tremendous declines in railroad costs and prices and increases in productivity. Indeed, the reversal over the last 35 years has been so successful that many are now concerned that the industry charges rates larger than what is necessary to recoup costs and that firms are earning excessive profits. While there have been a considerable number of studies that examine costs and prices, there are few, if any, studies that examine railroad markups.

Railroads each operate over massive and differentiated networks that each have different shippers, products, and operating characteristics. Over the network, they serve a multitude of markets, making standard empirical models of markups intractable. Instead, I develop a model of markups and scale elasticities based on the seminal work of Hall (1988) and Klette (1999) and introduce Bayesian methods to estimate the model. This model estimates rail specific markups, which can be interpreted as the average markup generated from traffic for each firm's network.[1] It also provides direct estimates of scale elasticities. In both cases, the Bayesian estimation provides distributions of markups and scales that are allowed to drift over time and vary across railroads. Further, this approach also allows for the correlation of markups and scale elasticities across firms that provide insight into industry structure. I find that railroad markups are well in excess of marginal

[1]These methods use a local mean value theorem approximation centered around the representative firm in each year, which can provide more accurate results than methods that assume a global functional form such as a Cobb-Douglas production function. However, this approximation is only accurate locally and thus cannot be used to evaluate total costs or examine minimum rates needed for firm viability.

costs and that most firms are operating with near constant returns to scale at most points in time.

The railroad industry has been federally regulated since 1887 and was partially deregulated in 1980. In the 1800s, there was high demand for rail transportation and few substitutes. During this period, the rail network grew reaching a maximum size in the early 1900s. Much of the existing network today is the result of this growth. However, due to the nature of railroad production and costs that include large fixed factor investments, the market has a tendency to be highly concentrated, especially on a local level (MacDonald and Cavalluzzo, 1996). The effects of concentrated transportation markets have long been a focus of regulators, and emphasis has been put on balancing efficiency advantage of large railroads with the harm caused by non-competitive pricing and shipper captivity (Boyer, 1987).

As the industry evolved and competing forms of transportation such as barges and trucks were introduced, there were dramatic changes to the structure of the industry and the regulations that govern it. In particular, innovations in transportation and changes in the types of goods being shipped led to a significant decrease in the demand for railroad transportation. Further, under regulation firms were slow to adjust and were unproductive. By the 1970s, the industry was failing financially and there were many highly publicized bankruptcies. The economic viability of the industry along with a failed regulatory regime led to partial deregulation of the industry with passage of the Railroad Revitalization and Regulatory Reform Act in 1976 and the Staggers Rail Act in 1980. The number of firms has fallen dramatically following deregulation, the rail network held by the Class I carriers has shrunk, shipments are traveling longer distances, and more shippers find themselves with few shipping options. These changes have certainly improved the efficiency and viability of railroads, but there remain many concerns that firms are price discriminating and charging excessive rates to many shippers (Wilson and Bitzan, 2003). While the remaining regulation over the industry has the power to limit excessive prices and non-competitive behavior, many sources have pointed to regulatory efficacy that is lacking both in its theoretical basis and in its execution.[2]

There are many studies that evaluate competition and structure in the railroad industry. These studies have highlighted the shortcomings of regulatory tools currently in use and have had success in determining important factors in production, modeling costs, and describing railroad operations. The most recent studies have continued with estimating costs of specific shipments and beginning to examine markers of non-competitive behavior and firm viability. Many studies have also looked towards precisely explaining differences in prices using detailed data describing costs and competitiveness.[3] Still, to my knowledge, there are few, if any, published studies that evaluate and consistently estimate both pricing behavior and scale of production; such research would provide insight into the structure of production and competition as well as non-competitive pricing.

This paper begins by covering the history of the railroad industry, its growth and regulation, the effects of partial deregulation, and the current state of the industry and regulation. I then highlight several methods from the industrial organization literature that look to evaluate structure, efficiency, and competition within an industry and consider the strengths and shortcomings of each. Among these methods are the

---

[2]Specifically, current regulatory practices compare revenue to variable costs to evaluate firm viability and excessive rates, but the methods used to estimate variable costs have been subject to significant criticism. Many studies have identified both broad and specific shortcomings in regulation, for further reading see Boyer (1987), MacDonald and Cavalluzzo (1996), Burton (1993), Winston et al. (1990).

[3]Most of these studies use waybill data and reduced form models to describe prices; unfortunately, these reduced form models fail to estimate parameters that are fundamental to firm behavior and can therefore produce inconsistent estimates and results. For further reading, see Casavant et al. (2012), Schmidt (2001), and Barnekov and Kleit (1990).

production-focused models developed by Solow (1957), Hall (1988), and Klette (1999), which I draw on and extend in this analysis. I continue on to describe how production and costs of railroads have been modeled in the literature and determine factors that are important to describing railroad operations.

This research goes on to derive a theoretical model of railroad production beginning with a commonly used specification of production. I then obtain an expression for firm output that involves firms' markups of price over marginal cost, scale elasticities, and network characteristics. Using this model and data from the United States Surface Transportation Board, I estimate two empirical models. The first, which replicates the methods of Hall (1988) and Klette (1999), assumes that markups and scales to production are constant across railroads and time. The results of this model give some evidence of deviations of price from marginal cost and production beyond minimum efficient scale in the industry.

I then propose a random-coefficients version of this model that allows for firm- and year-specific markups and scales, correlation among markups and scales across firms, and a central tendency for these parameters that varies flexibly through time.[4] Using Bayesian methods, I estimate this model and obtain density estimates of markups and scale elasticities. This method allows me to evaluate pricing behavior and scale economies for each firm and in each year and additionally provides information about the structure of the industry and the nature of competition between railroads through correlations in markups and scales. The results from this model provide evidence of significant markups of price over marginal cost and production near minimum efficient scale for most Class I railroads.

[4]Due to stark differences in the scope and scale of operations across railroads, it would not be surprising to find significant heterogeneity in pricing and returns to scale across firms. Further, though the industry has been relatively stable since 2001, investment in infrastructure and other operational changes could lead to variability in markups and scale elasticities across time.

Background

This section provides a brief history of the railroad industry, its regulation and partial deregulation, and how the industry has responded to these changes. The first subsection describes the beginning of the industry, the movement for its regulation, and the regulatory policies that were introduced. The second subsection describes how the industry changed over the 20th century as new forms of transportation were introduced, the impetus for deregulation of the industry, and how the it has changed after its partial deregulation. I also describe current regulatory practices, concerns over their efficacy, and worries that partial deregulation has made discriminatory pricing more common.

A Brief History of Railroads and Their Regulation

The railroad industry is one of the nation's oldest, and as the industry has grown and evolved, so has our understanding of how these firms compete and how to best regulate them. At the industry's beginning, railroads were massively successful. Not only was rail transportation much faster and more comfortable than previously existing modes, but it was effectively the only option for traveling long distances in a reasonable amount of time (Brown, 2013). Further, the ability to quickly transport goods and raw materials across the country allowed for rapid expansion of other industries like agriculture and energy production. As a result, railroads faced large demand for their service and the industry thrived.

At the same time, economic theory describing firm competition was growing and shifting as economists began considering failures in perfect competition and the consequences thereof (Brown, 2013). Leading thinkers in the area realized the extent of social damage that monopolies and other forms of imperfect competition can cause and began to seek regulation in industries where competition was evidently less prevalent.

Additionally, economists also began to recognize that firms in some industries might be able to realize economies of scale and increase efficiency by increasing their size. Given the large sunk capital investments in the industry and the vast amount of overlapping rail networks that resulted from decades of fierce competition, railroads were able to supply their service at a lower cost by combining or cooperating. This consolidation led to increased local market concentration and seller captivity, which allowed firms to more easily charge non-competitive rates.

Following this discussion, ideas about railroad regulation began to shift. As Brown (2013) states, prior to these realizations, "laissez-faire doctrines held that monopoly as an economic problem originated with explicit grants by governments to firms or individuals." The solution to these types of monopolies is clearly to reduce government involvement and purge the source of monopoly. However, in the case of the rail industry where local monopolies were more likely to emerge naturally, solutions had not been fully examined. At the same time, many shippers were exceedingly concerned about discriminatory and unfair pricing for railroad services. The Granger movement to regulate railroad operations began in Iowa and was the result of excessive rates being charged to less serviced shippers (Miller, 1954). Many farmers subsequently found that their operations were no longer viable under high shipping rates and it became necessary to control the prices charged by railroads. Many regulators like Cooley (1884) recognized the need to balance efficiency gains that would be realized from a natural monopoly and the social harm that can be inflicted by excessive rates and discriminatory pricing.

Regulation of railroads began with the passage of the Interstate Commerce Act (ICA) of 1887. This act and its successors allowed for significant oversight of railroad operations, including authority over rates charged, entry into and exit from the industry, and mergers. Specifically, the ICA made it explicit that railroads were legally bound to providing services at reasonable rates and without undue discrimination in prices or availability of services. The ICA also established the Interstate Commerce Commission (ICC), which provided regulatory oversight of railroads, and the Elkins Act, Hepburn Act, and Mann-Elkins Act of the early 1900s further strengthened regulation and the abilities of the ICC to enforce those regulations; by 1920, the ICC was able to set minimum and maximum rates, preside over and even encourage mergers, and control and punish collusive behavior (Keeler, 1983). These regulations improved outcomes for shippers and railroads alike; excessive rates were abolished, shipper coverage increased, rate wars and excessive competition between firms diminished, and railroad profits tended to improve.

Partial Deregulation and Its Effect on the Industry

As the 20th century progressed, rail transportation waned in popularity with the introduction and improvement of competing forms of transportation such as barges, trucks, and airplanes. As these new technologies were adopted, competition in the rail industry changed dramatically; demand for rail transportation fell,[5] the market became much less viable for firms, and the regulations put in place in the early part of the century became less relevant and arguably hindered efficiency in the industry (Keeler, 1983). There were not only many concerns about the difficulty and cost of implementing regulation but also about how those regulations made it difficult for firms to remain viable by, for example, forcing railroads to continue operating on routes that were massively unprofitable (Waters, 2007). In response to worries about the future of firms and stability of the industry, the government began to deregulate the railroads, primarily with the passage of the Railroad Revitalization and Regulatory Reform (4R) Act of 1976, which

[5]This drop in demand can be attributed both to shippers substituting other modes of transportation for rail shipments and a change in the types of products being shipped. With the introduction of plastics came shipments that were lighter on average, resulting in fewer revenue ton miles.

reduced price regulation and made it easier for firms to enter and exit the market, and the passage of the Staggers Act of 1980, which further reduced regulations on pricing and mergers.

The major effect of partial deregulation of the railroad industry was allowing firms to set their own rates with minimal ICC control for most commodities. Instead, the ICC exercises control over movements it deems to be market dominant, defined as a movement where the ratio of revenue to variable costs exceeded a given threshold.[6] Movements that fit this criterion are then investigated by the ICC to determine if competition among railroads is present and if the rates charged are "reasonable". The prices charged to a given shipper are determined to be unreasonable if they exceed the stand-alone cost of servicing only that shipper. The ICC rests the burden of proof on shippers; in order to demonstrate unfair rates, the shipper must construct a hypothetical railroad and show that the cost of servicing the shipper is greater than the rate charged. While stand-alone costs exclude the variable and separable fixed costs of servicing other shippers, they include the common costs of supporting an entire railroad network, which decrease as the railroad services more shippers. As a result, rates charged by a stand-alone railroad would necessarily be "absurdly high" because they largely ignore economies of scope and scale, and very few shippers would be able to demonstrate that rates are unreasonable (Roberts, 1983). While this regulatory policy has been heavily scrutinized and criticized, it remains today as the primary tool for identifying and penalizing non-competitive pricing.

The Staggers Act further set minimum rates at average variable cost and allowed railroads to more easily abandon routes and merge with one another (Winston et al., 1990). These regulatory changes had a dramatic effect on the structure and operations of the industry; massive consolidation resulted in the number of firms decreasing from 40

[6] This threshold began at 1.6 but is currently 1.8.

in 1980 to seven in 1999, abandonment of routes led to a 57.8% decrease in miles of road between 1980 and 2013, and the average length of haul increased appreciably from 615 miles in 1980 to 973 in 2013 (AAR, 2013). The effects of these changes are clearly seen in the time series of output and network size, as shown in Figures 1, 2, and 3. Following partial deregulation in 1980, railroads began abandoning unprofitable routes, leading to a massive decrease in the total amount of network operated by Class I railroads. At the same time, due to efficiency gains, the total output of Class I railroads grew as costs fell. Given the massive consolidation in the industry, the average railroad's network grew sizeably through 1999, by which time all major consolidations had occurred. Since 1999, network size has remained relatively constant, illustrating a stabilization of the industry. These regulatory changes not only had a large effect on the industry but also achieved many of their goals; operating costs per ton mile fell by 60% between 1980 and 1998, return on equity increased from 3% to 10.7%, and shippers have benefited from lower rates, more reliable service, and faster shipment times (Peltzman and Winston, 2000).[7]

While it initially appears that partial deregulation has provided an improvement for railroads and shippers alike, there are still many concerns about non-competitive pricing behavior, especially in captive markets. Anecdotal evidence of harm to captive shippers is compelling, with shippers noting that "when you're captive to one of these railroads, the idea of negotiating a contract is pretty laughable" (Bowman, 2013). Recent research has found similar trends; the number of captive shippers has been increasing since deregulation, and railroads appear to be exerting market power to charge higher rates, especially when shippers have few transportation options to choose from. Henrickson and Wilson (2014) find that rates can be as much as 13.9% lower when shippers have access to

[7]Winston et al. (1990) found that the benefit to shippers from changes in regulatory policy was approximately $12 billion annually in 1999.

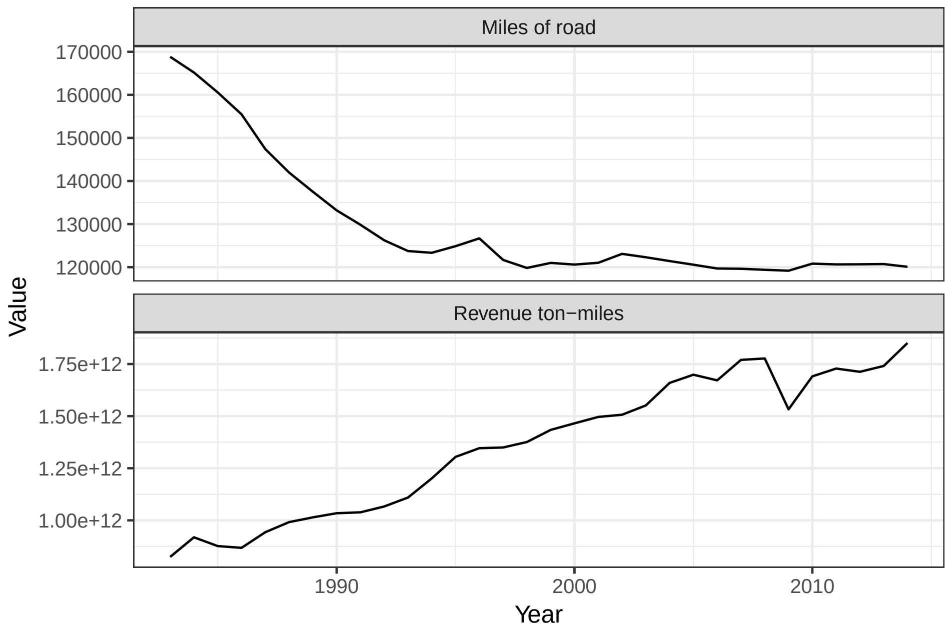

FIGURE 1. Revenue Ton-miles and Miles of Road

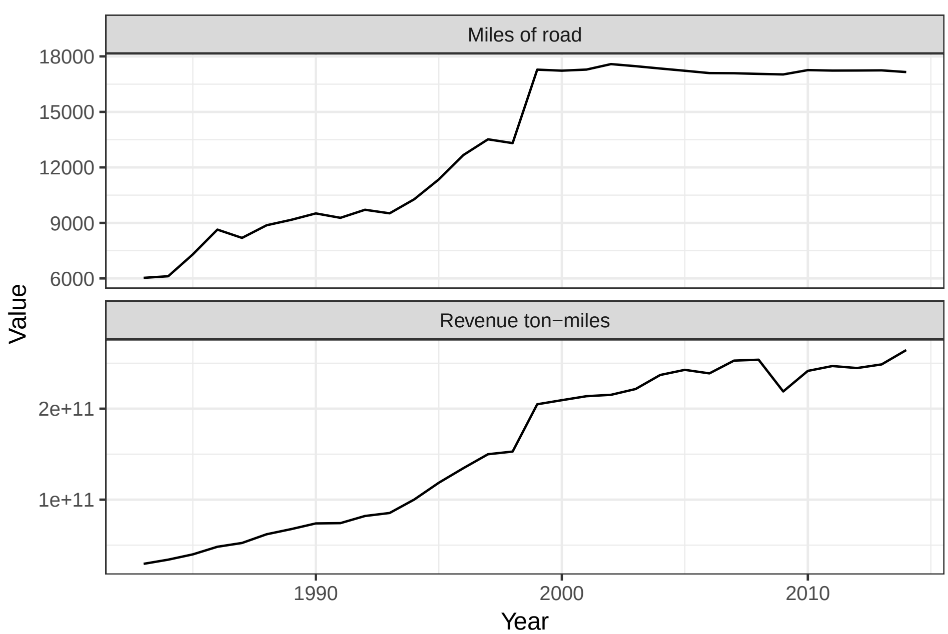

FIGURE 2. Average Revenue Ton-miles and Miles of Road

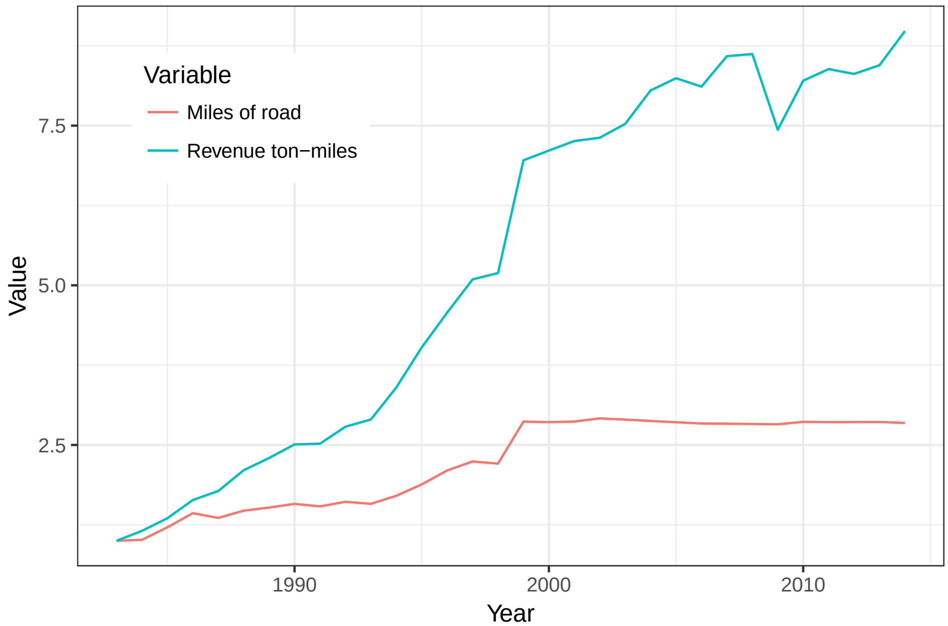

FIGURE 3. Average Revenue Ton-miles and Miles of Road Relative to 1983

multiple methods of transportation, pointing to the market dominance and discriminatory pricing policies of many railroads.

There have also been many concerns over regulatory efficacy. The Christensen report, a study commissioned by the United States Surface Transportation Board (STB), examined competition in the rail markets in the period that followed partial deregulation of the industry. Laurits R. Christensen Associates, Inc. (2009) first found major flaws in the methods regulators use to identify non-competitive behavior. The report found that the STB's policy of using the ratio of revenue to variable costs of individual movements as an indicator of market dominance was flawed for two reasons. First, the study found evidence that methods used to estimate variable cost for a given movement were not theoretically sound and produced poor results; estimates of the ratio of revenue to variable cost ranged anywhere from over 300% to paradoxically below 100%. Additionally, the Christensen report found that the constructed ratio of revenue to variable costs was only weakly correlated with more preferred measures of market captivity and dominance. As a result, the report recommended an overhaul in the methods used by the STB to identify non-competitive behavior, noting that "more appropriate measures of captivity should focus on the effects of the transportation market structure on rail rates and, by extension, markups" (Laurits R. Christensen Associates, Inc., 2009). While many studies have estimated cost functions that are used to obtain markup estimates, I am not aware of any published work that directly estimates markups and scale elasticities through estimation of the production technology.

Literature Review

This section highlights existing literature related to modeling railroad operations and evaluating structure and competition. Since there is little research that directly evaluates the competitive behavior of railroads, this review separately focuses on methods used to evaluate structure and identify non-competitive behavior and those used to model railroad operations and production. The first subsection examines the extensive literature in industrial organization that has developed many tools to evaluate the structure of industries and the conduct of firms and identifies a model that is appropriate for the study of railroads. The second subsection focuses on research that has modeled railroad operations and determines which factors are important to describing costs and production of railroads.

Identifying Non-competitive Behavior

The field of industrial organization has had a long history of evaluating competition between firms, beginning with Mason's industry studies, continuing with many inter-industry studies of the 1950s and 1960s, and culminating in the development of methods used to evaluate competition and pricing behavior within an industry (Schmalensee, 2012). As the field has matured, theory describing competition between firms and more robust statistical methods needed to consistently estimate these models have been developed and successfully used. These studies tend to either investigate the type of competition between firms or estimate observable markers of competition, such as the ratio of price to marginal costs. Several of these models have been criticized for their lack of theoretical basis or difficulty in estimation. Other methods, however, have been shown to produce consistent and credible estimates when applied appropriately.

Kadiyahi et al. (2001) provides a broad exposition into several methods of estimating the form of competition and highlights the successes and failures of each. One method, colloquially known as the menu approach, estimates a statistical model that nests several theoretical competitive models and performs statistical tests on parameters that coincide with each type of competition. Unfortunately, the conclusions of this method were often seemingly paradoxical; it is possible that one might either fail to reject hypotheses of exclusive forms of competition or, in a more troubling circumstance, one might reject hypotheses of all types of competition built into the model. In either case, the result is not particularly illuminative, and it can be difficult to make policy recommendations based on these conclusions.

Another more frequently used method, which includes the conjectural variations approach, focuses on estimating parameters that will provide indications of competition. Corts (1999) uses a conduct parameter in his analysis that describes the extent to which firms set prices above marginal cost. The theoretical value of this conjectural variation can be calculated for several forms of competition and compared to actual estimates. Unfortunately, this approach suffers from the same blight that affected the menu approach; it's entirely possible that one could fail to reject multiple hypotheses of conflicting types of competition, and it's also possible that one could reject hypotheses of all forms of competition built into the model. Further, this approach typically requires estimating both demand and supply functions, raising possible concerns over specification errors as well as any problems associated with estimating of systems of equations like endogeneity and instrument choice. However, since this approach treats firm conduct as a continuous variable, the results of these models can offer some idea of how "uncompetitively" firms behave; still, this interpretation has faced criticism because it fails to accurately model the exact form of competition between firms.

Current research has began to favor estimating an indicator of competition, as in the conduct parameter approach, but has focused on making that analysis more robust. Specifically, attention has turned towards estimating the ratio of price to marginal cost, commonly referred to as the markup and focuses less on describing the exact form of competition. Since perfectly competitive firms set price equal to marginal cost, one can establish a competitive benchmark and compare this with estimates of firm markups; markups significantly greater than one indicate deviation from competitive behavior. There have been several methods developed to estimate markups; these models either estimate models of cost and demand or consumer preferences directly, or estimate production functions to find marginal costs and use an assumption of profit maximization to obtain markup estimates.

Berry et al. (1995) examines markups by estimating a model of costs and a flexible model of consumer preferences. The authors develop a model of consumer preferences that allows for preference heterogeneity and aggregate decisions to the market level. Using product characteristics and time series for sales, the authors estimate the distribution of individual consumer preferences and derive a measure of aggregate demand for each product as well as hypothetical products. The authors additionally develop a model of costs for multiproduct firms and pair it with their model of aggregate demands to obtain markup estimates. It is important to note that this model requires well-defined definitions of markets and an assumption of Bertrand competition in each market to obtain consistent estimates. The authors apply their model to automobile markets; not only do estimates provide information about demand such as preferences for certain characteristics and cross price elasticities, but they also provide reasonable markup estimates, ranging from 18.4% for the Nissan Sentra to 48.4% for the Lexus LS400 (Berry et al., 1995).

Rather than relying on estimating a cost function and model of demand directly, many studies have looked towards using variability in productivity to explain costs, which can then be used to obtain more robust estimates of the ratio of price to marginal cost by limiting problems that can occur in demand estimation. Solow (1957) set the groundwork for this type of analysis by noting that given the assumption that firms have Hicks neutral demand, output can be expressed in terms of input prices, technological growth, and productivity of the various inputs. Solow used his results for a macroeconomic analysis of how technological change and labor productivity has evolved over time in the United States and was able to determine the average rate of technological change and how much of the growth in aggregate output was due to these technological improvements. Hall (1988) recognized that this type of analysis would be enormously valuable if carried out at the industry level; he extends Solow's theory to describe individual firms and estimates the relationship between input usage and output separately for a number of industries. Hall further derives a method of estimating the markup of price over marginal cost using Solow's estimation framework. Hall (1988) found markup estimates that generally matched expectations. For example, the markup on non-durable goods was estimated to be approximately 104% greater than the markup on durable goods, likely because of the relatively low elasticity of demand for the former.

Klette (1999) further extended and solidified Hall's contributions. Klette follows Solow and Hall by assuming a Hicks neutral production process for firms. However, Klette goes on to make several important original contributions. First, rather than using a Taylor approximation as in Solow and Hall, Klette uses a mean value theorem approximation. Since variables in this analysis include individual firm output and input usage, which can vary wildly across an industry, linear Taylor approximations, which are accurate around the point they are centered but less so away from that point, tend to perform poorly. The mean value theorem approximation, on the other hand, is "*a priori* suitable for samples with any magnitude of cross sectional differences in output, productivity, and inputs" (Klette, 1999). Next, it is possible that firms might experience different input qualities and therefore different input prices; capturing this variation in input quality and prices is important to properly identify productivity differences across firms. Finally, while Klette measures average parameter values and markups and scale elasticities for various industries similar to Hall's results, he also develops and utilizes a framework to estimate the distribution of firm markups, a result that would enormously useful in regulating industries. Additionally, Klette's econometric approach appears to produce more reasonable estimates of markups than did Hall's; Klette's markup estimates were generally between 1 and 1.1, while many of Hall's estimates were greater than 2 and many exceeded 3.

Modeling Railroad Operations

Describing the functions, costs, and operations of railroads is difficult for a number of reasons. First, unlike many other extensively studied industries, railroads produce many types of outputs; not only are there many different commodities that railroads ship, but firms also serve many shippers that want their goods transported to many different locations. Modeling railroad operations for each of these outputs would be onerous even with complete data; unfortunately, only aggregate measures of output are typically observed, such as the total number of tons shipped. Further, given the scope of railroad operations, it can be arduous to accurately describe the cost of transporting goods.[8] Not only do costs depend on direct factors like use of a locomotive or consumption of fuel, but also on indirect factors that control operation and coordination of movements along a railroad's network. Finally, since railroads serve many geographically separated points on their networks and since many inputs to production aren't sufficiently divisible, many markets experience varying amounts of excess capacity, giving rise to very complex pricing behavior to cover the cost of excess capacity. Many techniques and methods have been developed and refined to more accurately describe railroad operations, costs, and productive capabilities.

Describing output of railroads, especially aggregated output, has been a long standing problem in the study of the industry. Given the many outputs that firms produce that depend on both the commodity shipped and the origin and destination of the shipment, one would ideally use completely disaggregated data that contain amounts of each commodity shipped, rates charged, and complete information for each movement across the network to identify all factors that influence costs. However, given the difficulty in obtaining this data, few studies investigating individual movements across railroad networks have been conducted. Notably, Wilson and Bitzan (2003) use disaggregated industry data to investigate the costs of individual rail movements and had success in using this data to estimate shipment specific costs over time. Using properties about shipments, these methods can be used to find, for example, marginal shipping costs for various commodity groups that depend on location and other shipment characteristics.

[8] This is made even more difficult when working with aggregated data since shipment specific characteristics can't be directly controlled for.

However, complete disaggregated data concerning railroad operations has not been available historically and is currently confidential, so most studies have relied on using aggregated data instead.

Given the complexity of outputs that railroads produce, identifying an appropriate measure of aggregate output was a focal concern in early studies. Studies have used many different measurements of railroad output, but eventually researchers, recognizing the importance of both weight and distance traveled, predominantly began using revenue ton-miles, defined as one ton that generates revenue shipped one mile, as a measure of output (Waters, 2007). Collapsing rich data describing highly specific measures of output into this aggregate measure of output discards a wealth of information, making it difficult to accurately describe the relationship of costs and aggregate output and necessary to consider many factors when modeling costs and production. Specifically, total revenue ton miles reflect flows of shipments over the railroad's network, but due to the nature of its aggregation, its effect on costs is clearly correlated with other characteristics of the railroad's network. For example, since coordination of movements necessarily becomes more difficult as a railroad's network size increases, the marginal cost of revenue ton miles will depend on network size. Friedlaender and Spady (1981) identified major factors of railroad costs and labeled them as fixed inputs, variable inputs, and technological conditions. Variable inputs include labor as well as use of equipment and fuel, the fixed input was total value of track, measured using replacement costs, was used as the fixed input, and technological conditions were described with characteristics of the railroad's network such as network size, average length of haul, and the mix of various commodities being shipped. It is crucial to include these factors as well as properly specify how each works together to produce output in order to obtain consistent estimates of production and costs.

In addition to determining the factors that affect output and costs, researchers have struggled with describing the shape and functional form of costs and production. Researchers began by using familiar functional forms that are relatively easy to work with; for example, Keeler (1974) uses Cobb-Douglas production function that accounts for both inputs used and network characteristics to estimate returns to scale and density of railroads. Recognizing the restrictiveness of assuming particular functional forms, many researchers have had success in estimating more flexible cost function forms such as the translog form; these studies have generally found that using flexible forms that relax important assumptions like subadditivity yield more accurate results and predictions (Waters, 2007). Overall, researchers have seen success in using more flexible functional approximations to describe production and costs, and, fundamentally, many approximations such as higher order Taylor approximations, log-linearization, and mean value theorem approximations could also be used.

Running parallel to the field of industrial organization, research of railroad operations and competition has focused on estimating the markup of price over marginal cost as an indicator of competitive behavior. In order to identify railroad markups, researchers must fundamentally have a model of costs to derive an estimate of marginal costs and a model of market demand to identify equilibrium prices.[9] Railroad studies have largely focused on simultaneously estimating cost functions and demand relations (Waters, 2007). To my knowledge, no production based methods for estimating markups like those found in Hall (1988) and Klette (1999) have been applied to the railroad industry.

Notably, researchers have used conduct parameter approaches like those proposed in Corts (1999), and some have focused on more robust demand estimation like that found

[9]While costs must generally be estimated to obtain estimates of markups, assumptions about demand can range from complex specifications that estimate preferences to simple assumptions like profit maximization.

in Berry et al. (1995). Specifically, Ivaldi and McCullough (2005) assumes heterogeneous preferences amongst shippers and receivers and models aggregate demand for a number of markets[10] using a random coefficients framework. The authors pair these estimates of demand with a model of costs and the assumption that firms engage in Bertrand competition to obtain estimates of efficiency and market power. Using this framework, the authors examine the effect of mergers on efficiency and consumer surplus. The authors find strong evidence of scale economies and conclude that efficiency gains from mergers outweigh the costs of increased market power and non-competitive pricing; the authors estimate that over the period from 1986 to 2001, consumer surplus increased by about 30%. Further, the authors estimate that the ratio of price over marginal cost ranged from 1.378 for freight services to 1.85 for intermodal services. Finally, the authors find that returns to density for railroads were approximately 1.08, indicating that firms are operating near their minimum efficient density. However, as previously noted, the methods used in Berry et al. (1995) require a clear definition of markets in which the firms compete, which can be difficult in the railroad industry where firms not only provide different types of outputs but also service many distinct geographic areas.

Several other studies have also looked towards estimating returns to scale for railroads. Keeler (1974) estimates a cost function for railroads and finds that average returns to scale were approximately 0.993 in the period from 1968 to 1970. Caves et al. (1981b) estimates an extensive translog cost function for railroads and found that returns to scale remained relatively constant and near one over the period from 1955 to 1974.[11] More recently, Bereskin (2009) estimates a model of costs and technological variation to

---

[10] These markets were defined over three different types of rail services: Bulk, general freight, and intermodal transportation.

[11] Specifically, the authors estimate that scale elasticities of 1.012 for 1955 and 1963 and an elasticity of 1.036 in 1974.

explain economies of scope, density, and scale and finds that the average railroad operates at minimum efficient scale. Overall, the literature finds overwhelming evidence that railroads are no longer able to realize economies of scale and are operating at capacity. While many studies have examined railroad scales, few have looked towards estimating the markup of price over marginal cost.[12]

Finally, research in railroads and other network industries has attempted to draw a clear distinction between returns to scale and returns to density. While returns to scale measures the degree to which output is affected by an increase in the overall scale of a firm's operations including the size of its network, returns to density holds network size constant and measures the effect of increased concentration of input use. Since returns to scale considers an increase in all inputs and network size while returns to density considers an increase to all inputs other than network size, these returns can be markedly different depending on how much of an effect network size has on output. Caves et al. (1984) separate returns to density from returns to scale by estimating a trans-log cost function that directly controls for network size and apply their model to the airline industry. Since airline costs increase considerably as network size increases, the authors estimate that returns to density significantly exceeded returns to scale, with elasticities of 1.18 and 0.99, respectively. The cost literature regarding railroads similarly finds that larger networks are associated with higher average costs; Wilson (1994) estimates that the cost elasticity of network size is 0.22, and thus finds a significant difference in returns to scale and returns to density, with estimates of 0.99 and 1.34, respectively.[13] As a result, it is important to draw the distinction between these measures and clearly identify which is being used.

[12]Notable exceptions are Ivaldi and McCullough (2005), as previously mentioned, and Bitzan (2000), which simulates polar Ramsey markups for different levels of shipper captivity.

[13]Wilson (1997) estimates the cost elasticity of network size is even higher at 0.74, leading to an even greater gap between scale and density estimates.

Theory

I begin with a production technology that has been used in a plethora of previous studies and specifically assume that production follows a Hicks-neutral process. In this specification, I characterize production with a mapping of inputs into an aggregate output. I also, however, recognize that the mapping depends on a set of operating characteristics that vary across outputs and networks. Hence, I condition on the variable φ to reflect these differences.[14] Overall, I let the output of firm i in year t depend on a productivity factor A_{it}, input use $\{X_{it}^{j}\}_{j\in M}$, where M is the set of inputs, and network characteristics $\varphi_{it} = \{\varphi_{it}^{\xi}\}_{\xi\in\Xi}$, where Ξ is the set of network characteristics:

$$Q_{it} = A_{it}F_t(\{X_{it}^{j}\}; \varphi_{it}).$$

Now, I wish to make minimal assumptions on the shape or exact form of F_t, so I log-linearize the production technology around the representative firm in each year. Letting variables with t subscripts denote values for the representative firm (e.g., Q_t is aggregate output for the representative firm in year t), firm i's output can be approximated with

$$\begin{aligned}\ln(Q_{it}) - \ln(Q_t) \;&\approx\; (\ln(A_{it}) - \ln(A_t)) \\ &+ \sum_{j\in M}\frac{\partial \ln F_t}{\partial \ln X_{it}^{j}}\left(\ln(X_{it}^{j}) - \ln(X_{t}^{j})\right) \\ &+ \sum_{\xi\in\Xi}\frac{\partial \ln F_t}{\partial \ln \varphi_{it}^{\xi}}\left(\ln(\varphi_{it}^{\xi}) - \ln(\varphi_{t}^{\xi})\right).\end{aligned}$$

Now, for notational convenience, I will use lower case variables with a hat to indicate log-deviations of that variable from the representative firm in a given year (e.g., $\widehat{q}_{it} = \ln(Q_{it}) -$

[14] Empirically, of course, I have a set of variables for operating characteristics and networks as well as a set of firm dummies that control for these effects.

$\ln(Q_t)$). Then, this approximation can be rewritten as

$$\widehat{q}_{it} \approx \widehat{a}_{it} + \sum_{j \in M} \frac{\partial \ln F_t}{\partial \ln X_{it}^j} \widehat{x}_{it}^j + \sum_{\xi \in \Xi} \frac{\partial \ln F_t}{\partial \ln \varphi_{it}^\xi} \widehat{\varphi}_{it}^\xi.$$

Standard log-linearization evaluates the above derivatives at the point centered around (i.e., around the representative firm at the point (A_t, X_t, φ_t)). However, the multivariate version of the mean value theorem tells us that there exists a point $(\overline{A}_{it}, \overline{X}_{it}, \overline{\varphi}_{it})$ between $(A_{it}, X_{it}, \varphi_{it})$ and (A_t, X_t, φ_t) such that this is no longer an approximation but is in fact exact:

$$\widehat{q}_{it} = \widehat{a}_{it} + \sum_{j \in M} \widehat{x}_{it}^j \cdot \left. \frac{\partial \ln F_t}{\partial \ln X_{it}^j} \right|_{(\overline{A}_{it}, \overline{X}_{it}, \overline{\varphi}_{it})} + \sum_{\xi \in \Xi} \widehat{\varphi}_{it}^\xi \cdot \left. \frac{\partial \ln F_t}{\partial \ln \varphi_{it}^\xi} \right|_{(\overline{A}_{it}, \overline{X}_{it}, \overline{\varphi}_{it})}.$$

Unfortunately, the mean value theorem only tells us about the existence of this point but not its exact location. As discussed in the Model section, the best *a priori* approximation is likely where the derivatives are evaluated at the midpoint between $(A_{it}, X_{it}, \varphi_{it})$ and (A_t, X_t, φ_t), but I also investigate evaluating the derivatives at (A_t, X_t, φ_t) and find little difference in results. I will continue to use the bar notation to indicate evaluation of these derivatives at the interior point whose existence is guaranteed by the mean value theorem.

Next, notice that $\frac{\partial \ln F_t}{\partial \ln X_{it}^j}$ is the elasticity of output with respect to input j and $\frac{\partial \ln F_t}{\partial \ln \varphi_{it}^\xi}$ is the elasticity of output with respect to network characteristic ξ. For notational convenience, denote these elasticities with $\overline{\alpha}_{it}^j$ and $\overline{\zeta}_{it}^\xi$, where the bar notation indicates evaluation of these elasticities at the point $(\overline{A}_{it}, \overline{X}_{it}, \overline{\varphi}_{it})$. One would like to have more information about the exact form of these elasticities; I begin by noticing that the elasticity of output with respect to input j can be expressed as

$$\alpha_{it}^j = \frac{X_{it}^j}{Q_{it}} \frac{\partial Q_{it}}{\partial X_{it}^j}.$$

Now, if firms are price-takers in input markets, then the first-order condition for profit maximization with respect to input j is

$$\frac{\partial Q_{it}}{\partial X_{it}^{j}} = \frac{W_{it}^{j}}{P_{it} + (\partial P_{it}/\partial Q_{it})Q_{it}} = \frac{W_{it}^{j}}{(1+1/\varepsilon_{it})P_{it}},$$

where W_{it}^{j} is the price of input j, P_{it} is the price of output, and ε_{it} is the elasticity of demand for using firm i's network in year t. Notice that, similar to Klette, I aam allowing railroads to experience different input prices, reflecting the possibility that the quality of inputs might vary across firms. Then, the elasticity of output with respect to input j can be written as

$$\alpha_{it}^{j} = \frac{X_{it}^{j}}{Q_{it}}\frac{\partial Q_{it}}{\partial X_{it}^{j}} = \left(\frac{1}{1+1/\varepsilon_{it}}\right)\left(\frac{X_{it}^{j}W_{it}^{j}}{Q_{it}P_{it}}\right) = \mu_{it}\frac{X_{it}^{j}W_{it}^{j}}{Q_{it}P_{it}},$$

where μ_{it} is the ratio of price to marginal cost, which I will refer to as the markup. For notational convenience, I denote the cost share of input j to total revenue with

$$\overline{s}_{it}^{j} := \frac{\overline{X}_{it}^{j}\overline{W}_{it}^{j}}{\overline{Q}_{it}\overline{P}_{it}}.$$

As a result, the elasticity of output with respect to input j can be expressed as $\overline{\alpha}_{it}^{j} = \mu_{it}\overline{s}_{it}^{j}$. As noted by Klette, due to various rigidities with fixed inputs, this relationship likely doesn't hold for capital inputs K; as a remedy, notice that scale to production is given by

$$\eta_{it} = \sum_{j \in M} \alpha_{it}^{j} = \alpha_{it}^{K} + \sum_{j \neq K} \alpha_{it}^{j}.$$

Solving this expression for α_{it}^{K}, one can arrive at the conclusion that

$$\alpha_{it}^{K} = \eta_{it} - \sum_{j \neq K} \alpha_{it}^{j} = \eta_{it} - \mu_{it}\sum_{j \neq K} s_{it}^{j}.$$

Returning to the equation for firm production and using the above work,

$$\begin{aligned}\widehat{q}_{it} &= \widehat{a}_{it} + \widehat{x}_{it}^{K}\overline{\alpha}_{it}^{K} + \sum_{j\neq K}\widehat{x}_{it}^{j}\overline{\alpha}_{it}^{j} + \sum_{\xi\in\Xi}\widehat{\varphi}_{it}^{\xi}\overline{\zeta}_{it}^{\xi}\\ &= \widehat{a}_{it} + \eta_{it}\widehat{x}_{it}^{K} + \mu_{it}\sum_{j\neq K}\overline{s}_{it}^{j}(\widehat{x}_{it}^{j} - \widehat{x}_{it}^{K}) + \sum_{\xi\in\Xi}\overline{\zeta}_{it}^{\xi}\widehat{\varphi}_{it}^{\xi}.\end{aligned}$$

Finally, since input use, input prices, and output are observed, one can be able to calculate

$$\widehat{x}_{it}^{V} := \sum_{j\neq K}\overline{s}_{it}^{j}(\widehat{x}_{it}^{j} - \widehat{x}_{it}^{K}).$$

Thus, firm i's output in year t can be described with

$$\widehat{q}_{it} = \widehat{a}_{it} + \eta_{it}\widehat{x}_{it}^{K} + \mu_{it}\widehat{x}_{it}^{V} + \sum_{\xi\in\Xi}\overline{\zeta}_{it}^{\xi}\widehat{\varphi}_{it}^{\xi}.$$

Data

The data used for this analysis primarily come from R1 forms, which contain various financial information and operating statistics for each Class 1 Railroad and are published annually by the United States Surface Transportation Board (STB). These forms contain aggregate measures of output, input use and prices, measures of capital depreciation, and statistics describing various network characteristics. I additionally supplement these data with the annualized version of the Quarterly Freight Commodity Statistics (QCS), also published by the STB. The QCS describes shipment revenue and tonnages for individual commodity groups, which are useful for precise measurement of prices and describing network characteristics. The time span of the sample has been restricted to the period from 2001 to 2012 for two reasons. First, though my preferred model allows for changing parameters over time, it may not accurately capture massive structural changes[15] to the industry that could be present in longer samples; as noted in the Background section, railroad operations have been relatively stable since 2001.[16] Second, all mergers of Class I railroads occurred prior to 1999,[17] no entry or exit occurred after 1999, and firms were fully consolidated by 2001, meaning that this sample constitutes a balanced panel, providing less complication in estimation. The Class I Railroads in the sample are the Burlington Northern and Santa Fe Railway (BNSF), the Canadian National Railway (CN), CSX Transportation (CSX), the Kansas City Southern Railway (KCS), the Norfolk Southern Railway (NS), the Soo Line Railroad (SOO), and the Union Pacific Railroad (UP).

The dependent variable in this analysis is aggregate revenue ton-miles, which is defined as one ton of product shipped one mile that generates revenue. I explain variation in output with input usage and variables describing characteristics of the network. Friedlaender and Spady (1981) find that costs, and by extension production, depend on way-and-structures capital, variable inputs, and technological conditions. The literature has used many different way-and-structures and technology variables to describe costs, but the authors generally find that total miles of road, average length of haul, traffic mix, and total value of track and other capital are key to modeling costs related to fixed inputs. Since I wish to measure scale elasticities relative to the value of capital used, I use value of track as a fixed input and the remaining variables as indicators of technological conditions.

[15]For example, the industry continued to experience consolidation and significant technological change through the late 1990s.

[16]While the industry was certainly more stable in the 2000s than in the period immediately following partial deregulation, there were still significant changes ranging from shocks in demand to substantial increases in investment over the course of the sample. Since these changes affected firms in different ways at different times, I still expect to find heterogeneity in model estimates across both firms and years.

[17]It is worth noting that while the Canadian Pacific Railway has owned the Soo Line Railroad since 1990, Soo underwent a change of name to Canadian Pacific in the early 2000s; I will still refer to this firm as the Soo Line Railroad.

Also, similar to Friedlaender and Spady (1981), I treat labor, equipment, and fuel use as variable inputs and I use rental rates as the opportunity cost of equipment inputs.

I define variable inputs for railroads to be labor, fuel, and the amounts of cars and locomotives used. Labor use is measured in hours worked, with the price of labor defined as the average wage for a railroad in a given year and fuel use is measured in gallons of diesel; also, while the R1 forms provide the total numbers of cars and locomotives used, it does not tell directly tell us about the opportunity cost of these inputs. Following Wilson and Bitzan (2003), I define the annual per unit opportunity cost for an equipment input j as

$$\frac{\text{Annual Depreciation}_{it}^{j} + \text{ROI}_{it}^{j}}{X_{it}^{j}},$$

where $\text{ROI}_{it}^{j} = \left(\text{Investment}_{it}^{j} - \text{Accumulated Depreciation}_{it}^{j}\right) \times \text{Cost}_{it}^{K}$ and Cost_{it}^{K} is the cost of capital for firm i in year t. Investment, depreciation, and input use can all be found in the R1 forms, and for the cost of capital I use the Rail Cost Adjustment Factor (RCAF) from the Association of American Railroads. I use total investment in road as the fixed input in this analysis. Since this variable measures the value of track, estimates of scale elasticities can best be interpreted as the elasticity of output with respect to the value of way-and-structures capital, holding network characteristics constant.[18] As with equipment variables, I calculate the opportunity cost of road investment using formulas similar to those used above.

Finally, I include into my analysis several variables describing network characteristics. First, as noted in Wilson and Bitzan (2003) and Ivaldi and McCullough (2005), the type of shipment (i.e., way-, through-, or unit-train shipments) and the type of product being shipped (e.g., bulk or specialty) could each have a large impact on costs

[18]As noted in the following paragraph, network size is included as a network characteristic, so this scale parameter is the elasticity of output given an increase in way-and-structures investment, holding network size constant. Thus, this elasticity of scale can better interpreted as elasticity of network density.

and are therefore important in production decisions. To control for these factors, I include the percentage of total ton-miles that are shipped via unit train and the percentage of ton miles that ship bulk items[19] into my regressions. Further, to capture the effective size of the network, I also include the average length of haul and total miles of road as controls in my regressions.

Descriptive statistics for each of the variables used in this regression are given in Table 1; nominal variables have been adjusted for inflation, and each of these statistics has been averaged across the time span of the sample.

Many of these variables including output, labor use, fuel use, numbers of locomotives, and investment in road have remained relatively constant over the course of the sample, except for a common negative shock induced by the 2009 recession. The number of cars used by railroads has been steadily decreasing over time due to an increasing number of shippers owning their own cars in recent years. Real prices of inputs and output have been increasing since 1999 except during the recession. Measures of network characteristics have remained largely constant over for each firm over this time frame, illustrating the stability of the industry since consolidation occurred.

Empirical Models

Common Markups and Scales

Recall the expression for firm i's output in year t:

$$\widehat{q}_{it} = \widehat{a}_{it} + \eta_{it}\widehat{x}_{it}^{K} + \mu_{it}\widehat{x}_{it}^{V} + \sum_{\xi\in\Xi}\overline{\zeta}_{it}^{\xi}\widehat{\varphi}_{it}^{\xi}.$$

[19]Bulk items belong to one of the following commodity groups: Metallic ores, nonmetallic minerals (not fuels), waste/scrap metals, clay/concrete/glass/stone, farm products.

TABLE 1. Descriptive Statistics

	BNSF	CN	CSX	KCS	NS	SOO	UP	Total
Output								
Revenue ton-miles (billions)	589.625	50.471	232.188	26.437	190.25	27.312	526.688	234.71
	(77.897)	(8.422)	(14.829)	(4.91)	(11.716)	(6.972)	(29.993)	(221.554)
Price	0.026	0.04	0.04	0.034	0.047	0.036	0.03	0.036
	(0.003)	(0.006)	(0.004)	(0.003)	(0.005)	(0.004)	(0.005)	(0.008)
Inputs								
Gallons of fuel (millions)	1309.606	102.26	547.779	63.283	482.172	54.176	1235.647	542.132
	(113.96)	(17.879)	(53.895)	(6.05)	(33.314)	(12.274)	(127.394)	(504.555)
Hours of labor (millions)	81.29	12.9	57.64	5.726	57.407	6.274	96.616	45.408
	(5.997)	(1.163)	(5.838)	(0.511)	(4.212)	(1.262)	(8.105)	(35.03)
Locomotives	6180.875	541	3889.562	545.75	3787.688	421.438	7913.688	3325.714
	(895.528)	(107.54)	(248.982)	(53.432)	(266.951)	(90.79)	(646.264)	(2814.122)
Cars	83837.5	24652.75	94733.75	12846.688	96078.125	14690.812	90103.75	59563.339
	(7253.763)	(6450.657)	(20162.264)	(1585.578)	(11392.32)	(1985.792)	(15344.713)	(38578.147)
Investment per mile of road (thousands of $)	794.122	1149.204	710.699	562.209	713.604	343.864	896.354	738.579
	(93.956)	(281.994)	(193.44)	(186.585)	(219.347)	(121.28)	(84.433)	(294.336)
Input Prices								
Fuel price	1.901	1.877	1.919	1.907	1.859	2.063	1.952	1.925
	(0.884)	(0.9)	(0.857)	(0.882)	(0.869)	(0.884)	(0.86)	(0.855)
Average wage	41.468	43.269	42.492	38.266	30.479	36.537	38.887	38.771
	(3.092)	(5.43)	(2.328)	(2.373)	(2.194)	(2.391)	(2.617)	(5.074)
Opportunity cost of locomotives	98703.232	80472.921	112343.754	44226.118	103682.365	58312.829	93820.205	84508.775
	(19954.892)	(35218.273)	(7479.84)	(32882.05)	(6127.92)	(22277.111)	(21263.086)	(32399.188)
Opportunity cost of cars	2390.587	3302.028	3520.771	1171.713	3203.51	2621.953	2817.87	2718.347
	(405.62)	(1167.264)	(495.391)	(767.3)	(329.293)	(1000.944)	(815.184)	(1047.773)
Opportunity cost of road invesment	5219.833	830.266	3478.481	490.814	3305.812	641.805	5274.224	2748.748
	(561.06)	(151.667)	(514.214)	(54.776)	(476.209)	(147.135)	(556.998)	(1995.899)
Network Characteristics								
Average length of haul	1046.849	288.825	533.266	363.193	464.402	440.834	921.168	579.791
	(67.523)	(22.837)	(40.975)	(43.623)	(21.087)	(38.505)	(24.854)	(271.546)
% unit shipments	0.476	0.17	0.333	0.414	0.252	0.277	0.415	0.334
	(0.031)	(0.054)	(0.022)	(0.095)	(0.02)	(0.046)	(0.021)	(0.111)
% bulk shipments	0.196	0.295	0.194	0.161	0.156	0.294	0.183	0.211
	(0.009)	(0.1)	(0.011)	(0.012)	(0.007)	(0.034)	(0.014)	(0.068)
Miles of road	32482	5847.812	21783.938	3096.312	20889.375	4123.125	32447.312	17238.554
	(424.674)	(1353.077)	(1038.289)	(163.18)	(693.435)	(1325.3)	(575.021)	(12034.503)
N	14	14	14	14	14	14	14	98

I begin by assuming that a firm's productivity relative to the median firm is constant across time in expectation, so that $\widehat{a}_{it} = \widehat{a}_i + u_{it}$, where u_{it} is a mean zero error term. Then, this expression for output can be rewritten as

$$\begin{aligned}\widehat{q}_{it} &= \widehat{a}_i + \eta_{it}\widehat{x}^K_{it} + \mu_{it}\widehat{x}^V_{it} + \sum_{\xi\in\Xi}\overline{\zeta}^{\xi}_{it}\widehat{\varphi}^{\xi}_{it} + u_{it} \\ &= \widehat{a}_i + \eta\widehat{x}^K_{it} + \mu\widehat{x}^V_{it} + \sum_{\xi\in\Xi}\overline{\zeta}^{\xi}\widehat{\varphi}^{\xi}_{it} + v_{it},\end{aligned}$$

where $v_{it} = u_{it} + (\eta_{it} - \eta)\widehat{x}^K_{it} + (\mu_{it} - \mu)\widehat{x}^V_{it} + \sum_{\xi}(\overline{\zeta}^{\xi}_{it} - \overline{\zeta}^{\xi})\widehat{\varphi}^{\xi}_{it}$. Now, I assume that μ and η are the mean markup and scale for the industry, so v_{it} is also a mean zero error term. Since $\widehat{q}_{it}$, $\widehat{x}^K_{it}$, $\widehat{x}^V_{it}$, and $\widehat{\varphi}_{it}$ are all calculable from the data, if v_{it} is assumed to be uncorrelated with explanatory variables, this equation could potentially be estimated with OLS. However, if shocks to output could be correlated with changes in input allocation, then my estimates would suffer from endogeneity bias.[20] Similar to Klette (1999), I propose using an instrumental variables approach to obtain consistent parameter estimates. Fortunately, I also have a set of relevant instruments that were assumed to be exogenous. In the Theory section, I assumed firms take input prices as given, meaning that shocks to output cannot be correlated with changes in input prices. I therefore instrument for $\widehat{x}^K_{it}$ and $\widehat{x}^V_{it}$ with input prices. I also assume network characteristics are exogenous to firm output[21] and thus do not instrument for those variables.

[20]In fact, one would expect this to occur since firms will likely adjust input usage to most efficiently produce a different quantity.

[21]Given that network characteristics are often difficult, if not impossible, for railroads to control (e.g., firms might have some control over what proportion of its shipments carry bulk products, but that characteristic is largely driven by demand for bulk product shipments), and since network characteristics have remained relatively constant over the sample period, I find this to be a reasonable assumption.

Firm- and Year-Specific Markups and Scales

Returning again to the equation

$$\widehat{q}_{it} = \widehat{a}_i + \eta_{it}\widehat{x}_{it}^K + \mu_{it}\widehat{x}_{it}^V + \sum_{\xi\in\Xi}\overline{\zeta}^{\xi}\widehat{\varphi}_{it}^{\xi} + u_{it}.$$

Given the firm- and year-specific markups and scales that appear in this equation, it is not identified in its current form; however, if it is assumed that these parameters come from a common distribution, then this equation can become estimable. A common approach for random-coefficient methods is to assume these parameters are independent and identically distributed. However, due to the nature of competition between firms and how technologies are adopted, I expect that markups and scale elasticities should be correlated, so that the independence assumption fails to hold. Instead, I assume that markups and scales have a common central tendency across firms, but allow these parameters to be correlated in a flexible way. Additionally, since markups and scales have likely changed over time for the industry as a whole, I allow for a flexible trend in the central tendency of the distribution over time. Overall, letting the number of firms be denoted by F, I assume that for each year t,

$$[\mu_{1t} - 1, ..., \mu_{Ft} - 1, \eta_{1t}, ..., \eta_{Ft}]' \sim \ln N\left([\mu_t, ..., \mu_t, \eta_t, ..., \eta_t]', \Sigma_{2F}\right),$$

where the central tendency $[\mu_t, \eta_t]'$ is assumed to independently and identically distributed across time and is also allowed to vary for each year. The distributional assumption on each μ_{it} and η_{it} enforces the theoretical restriction that scale elasticities must be greater than zero and markups must be greater than one for firms to be profit maximizing. Also,

Σ_{2F} is a square covariance matrix with dimensions equal to twice the number of firms that allows for flexible correlation between firm markups and scales.

Results

Common Markups and Scales

I first estimate the model where markups and scales are assumed to be constant across firms and years. Once again, in order to protect against endogeneity bias of the parameters μ and η, I instrument the variables $\widehat{x}^V$ and $\widehat{x}^K$ with input prices. The first stage of this regression is given in Table 2.

From these results, the instruments used appear to be relevant; in fact, given the importance of the instruments in the first stage, one might be concerned about their excludability. Fortunately, since there are more instruments than endogenous variables, I can test this identifying assumption with an overidentification test. It is important to note that if I find that these instruments are endogenous to output, then firms are not price takers in input markets; since this assumption was used in the derivation of the theoretical model, it is necessary that I don't observe contrary effects in the data. The results of the Sargan-Hansen overidentification test are given in Table 3, along with the second stage results.

First, the p-value of the Sargan test is greater than 0.05, indicating that there isn't significant evidence of instrument endogeneity. Next, I estimate that the average industry markup is 1.366 and that the average industry scale is 0.718. This indicates that, over the course of the sample, firms charge prices that are 36.6% greater than marginal cost on average and that firms are operating beyond minimum efficient scale. I additionally find that the elasticity of output with respect to average length of haul is negative and that the

TABLE 2. First Stage Results

	Dependent variable:	
	$\widehat{x}^V$	$\widehat{x}^K$
	(1)	(2)
Fuel price	0.044^{***}	-0.053^{***}
	(0.010)	(0.017)
Labor price	0.009^{***}	-0.020^{***}
	(0.002)	(0.004)
Road cost	1.811^{***}	-3.354^{***}
	(0.380)	(0.644)
Car cost	-0.00003^{*}	0.00004^{*}
	(0.00001)	(0.00002)
Locomotive cost	-0.00000^{**}	0.00000^{**}
	(0.00000)	(0.00000)
Average length of haul	0.003^{*}	-0.005
	(0.002)	(0.003)
Percent unit	-0.045	0.178^{**}
	(0.051)	(0.087)
Percent bulk	-0.006	-0.129^{*}
	(0.041)	(0.069)
Network size	-0.244^{***}	1.098^{***}
	(0.045)	(0.076)
Observations	84	84
R^2	0.853	0.996
Adjusted R^2	0.821	0.995
Residual Std. Error (df = 68)	0.049	0.083
F Statistic (df = 15; 68)	26.343^{***}	$1{,}059.620^{***}$
Note:	*p<0.1; **p<0.05; ***p<0.01	

elasticities with respect to other network characteristics are not differentiable from zero at the 5% level.

Firm- and Year-Specific Markups and Scales

In this section I present results from the model that allows markups and scales to vary across firms and across time. This model also allows for correlations between

TABLE 3. Common Parameter Regression

	$\widehat{q}$
$\widehat{x}^V$	1.366*
	(0.814)
$\widehat{x}^K$	0.718
	(0.479)
Average length of haul	-0.007***
	(0.002)
Percent unit	0.122*
	(0.072)
Percent bulk	0.065
	(0.089)
Network size	0.134
	(0.34)
Sargan Test Statistic	8.735
Sargan Test p-value	0.068
Railroad FE	Yes
Observations	98
R^2	0.998
Note:	*p$<$0.1; **p$<$0.05; ***p$<$0.01

markups and scales across firms, which can shed light on market structure. Rather than assuming parameters are given and data are random, as in a classical framework, Bayesian methods assume the data are given and parameters are random. As a result, these methods will yield probability distributions for each parameter that depend on prior beliefs and the likelihood of the data.

To conduct this estimation, I used a Metropolis-Hastings sampler with 500,000 burn-in iterations and 1,000,000 sampling iterations. Moderate autocorrelation of the Monte Carlo chain makes it necessary to use many samples and independent chains to obtain a representative sample of the posterior distribution. To ensure each chain has converged, I have varied initial parameter values and produced trace plots for each chain, ensuring that draws from the chain appear stationary. Additionally, I have varied prior distributions and compared results to the specification used in this paper to ensure prior assumptions aren't driving results. The full model and prior assumptions are presented in the Appendix of this paper.

This model produces a posterior probability distribution for markups and scales for each firm and in each year; since this constitutes a mass of information, I begin by examining these parameters for the most recent year for which results are available. Table 4 contains distribution quantiles for each firm's markup and scale in 2012 as well as elasticities with respect to network characteristics. I also present density plots for each firm's markup and scale in Figure 2.

TABLE 4. Bayesian Estimation Results

		Quantiles:				
	Mean	5%	25%	50%	75%	95%
2014 Markups						
BNSF	1.608	1.219	1.391	1.548	1.748	2.203
CN	1.406	1.238	1.296	1.358	1.466	1.745
CSX	1.558	1.237	1.354	1.49	1.648	2.183
KCS	1.508	1.174	1.304	1.478	1.664	1.998
NS	1.422	1.168	1.257	1.344	1.545	1.896
SOO	1.51	1.185	1.295	1.483	1.657	2.035
UP	1.518	1.245	1.346	1.46	1.572	2.072
2014 Scales						
BNSF	1.445	1.265	1.403	1.459	1.504	1.571
CN	0.968	0.865	0.927	0.978	1.012	1.047
CSX	1.136	0.509	0.781	1.034	1.483	1.965
KCS	0.911	0.883	0.902	0.912	0.921	0.939
NS	0.827	0.609	0.701	0.823	0.93	1.089
SOO	0.845	0.766	0.792	0.828	0.895	0.942
UP	1.025	0.888	0.964	1.047	1.087	1.13
Average length of haul	0	-0.008	-0.003	0	0.002	0.006
Percent unit	-0.221	-0.419	-0.275	-0.228	-0.142	-0.041
Percent bulk	0.134	0.023	0.091	0.121	0.181	0.263
Network size	-0.261	-0.394	-0.323	-0.263	-0.19	-0.159

One can first notice that there is significant heterogeneity in markup estimates across firms. There also appears to be some connection of markups with firm size and,

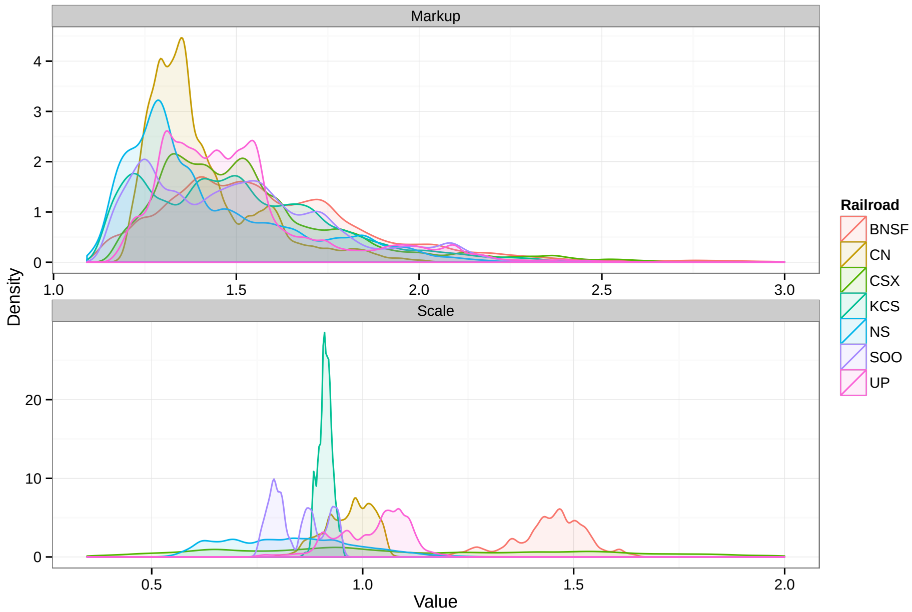

FIGURE 4. Distributions of Markups and Scales in 2012

by extension, market power; for example, BNSF, the largest firm in 2012, has the highest markup estimate while CN, a much smaller firm, has a lower markup. The posterior markup densities show that firms are generally pricing in excess and sometimes well in excess of marginal cost, with median markup estimates between 34.4% and 54.8%.[22] I find even greater heterogeneity in scales among firms. Smaller railroads like CN, KCS, NS, and SOO appear to be producing beyond their minimum efficient scale, while larger firms like BNSF, CSX, and UP show evidence of economies of scale. I estimate that BNSF is operating the furthest below minimum efficient scale in 2012, with mean and median scale estimates of 1.445 and 1.459, respectively. The heterogeneity observed in these

[22]It is important to note that markups in excess of one aren't necessarily indicative of non-competitive pricing. In particular, firms might charge prices greater than marginal costs in order to remain viable. Unfortunately, because this model is accurate only local to the median firm, I cannot estimate total costs and therefore cannot address firm viability.

parameter densities makes it clear that the assumption of common markups and scales is not appropriate, even for a given year.

This model allows for variation in markups and scales both across firms and across time. To show these dynamic results, I present line plots of median markup and scale estimates for each firm across time in Figure 3. While firm markups in 2012 are approximately the same as in 2001, there are many industry wide and firm specific fluctuations that have occurred over the sample. Overall, however, I don't observe a significant trend in average markups. Similarly, with the exception of BNSF and CSX, scale elasticities have remained relatively constant over the sample and tend to be centered around one, indicating that firms are producing at approximately minimum efficient scale on average. BNSF has shown a persistent increase in returns to scale that began in 2007, while CSX temporarily produced well below minimum efficient scale from 2009 to 2011; both of these increases in returns to scale appear to be driven by greater investment in these railroads' networks that led to excess capacity.[23] CSX began to fill that capacity in 2012, but BNSF has maintained investment and, thus, its economies of scale.

The United States Surface Transportation Board currently investigates shipments for which the ratio of price to estimated average variable cost is greater than 1.8; these shipments are scrutinized by the regulatory agency and the firm faces consequences if sufficient supplementary evidence is found. The Bayesian estimation framework used in this analysis allows me to estimate the probability that markups lie in some interval; thus, these results can be used to obtain the probability that a firm's markup of the price for network services in a given year is greater than the STB's designated threshold. I

[23]For example, over the period from 2004 to 2008, BNSF's investment in its network increased by 7.1% per year on average while the rest of the industry increased investment by only 5.2%. These investments improved the efficiency of inputs used on BNSF's network; in the period from 2006 to 2009, BNSF's output grew by 10%, yet it used 18.8% less fuel and 19% less labor.

FIGURE 5. Markups and Scales Over Time

present these probabilities for each firm in each year in Figure 4.[24] While this plot is similar to the plot of markups, it does not exactly match because of finer intricacies in the distribution of each parameter not fully described by the median. I observe strong evidence of excessive markups in the results, with this probability exceeding 0.5 for at least one year for each firm in the sample. Additionally, 2004 and 2007 appear to have the highest overall propensity for excessive markups.

As mentioned previously, this model allows markups and scales to be correlated across firms in each year, which can give information about market structure. I expect that markups and scales might be correlated across firms because of non-competitive pricing behavior, overlapping networks, and shared technologies. If two firms have a positive correlation in markups for a given year, then if one firm realizes a markup above

[24]It is important to note that these probabilities are not unconditionally independent from one another and only reflect the probability that a given firm's markup exceeds 1.8 in a given year.

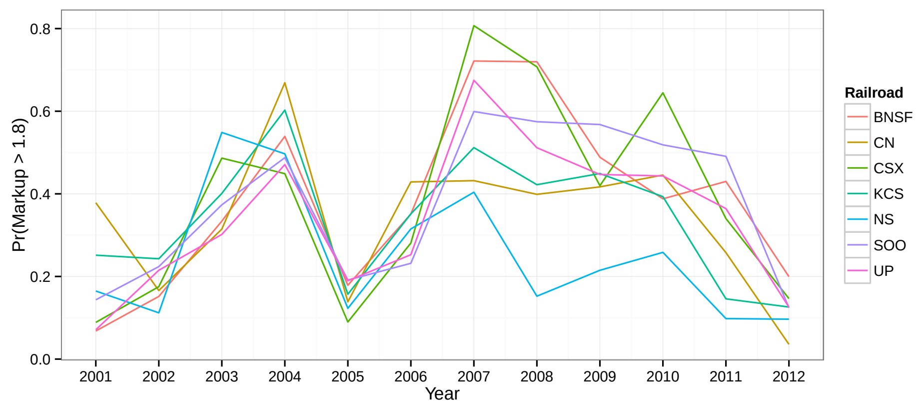

FIGURE 6. Probability of Markups Exceeding 1.8

its mean for that year, the other firm will tend to as well. While I can't attribute a cause to this correlation, it does provide information about the interaction between firms. These markup and scale correlations have been calculated for the most recent year available and are presented in Tables 5 and 6.

TABLE 5. Markup Correlations in 2012

	BNSF	CN	CSX	KCS	NS	SOO	UP
BNSF	1.0000						
CN	0.30235	1.0000					
CSX	0.20687	0.37846	1.0000				
KCS	0.31561	0.35442	0.34472	1.0000			
NS	0.17591	0.22838	0.24227	0.56877	1.0000		
SOO	0.0731	0.13855	0.21197	0.37733	0.35825	1.0000	
UP	0.35227	0.31953	0.40456	0.39826	0.49082	0.26694	1.0000

First, one will notice that each of the correlations between markups is estimated to be positive, indicating that a positive shock to any one firm's markup will tend to increase every other firm's markup. This observed correlation could be due to competition between firms, because as one firm lowers its price others will do the same to remain competitive and maintain customers, but could also be attributed to various other external causes that affect all firms in a similar way. For example, a positive shock to fuel prices will

TABLE 6. Scale Correlations in 2012

	BNSF	CN	CSX	KCS	NS	SOO	UP
BNSF	1.0000						
CN	0.33768	1.0000					
CSX	0.16234	0.19833	1.0000				
KCS	-0.06246	-0.35413	0.06353	1.0000			
NS	-0.41378	0.21033	0.20415	-0.29922	1.0000		
SOO	-0.48628	-0.47378	-0.52418	0.29651	-0.29331	1.0000	
UP	0.76765	0.33826	0.17917	-0.13196	-0.22497	-0.69003	1.0000

tend to increase marginal costs for each firm, thereby lowering markups. Given that I don't directly control for the effect of fuel prices on markups, this shock would induce an observed correlation among markups.

I observe even greater heterogeneity in scale correlations among firms. Again, I are not able to attribute a cause to these correlations, but they do offer some insight on market structure. There could be positive correlations in returns to scale if, for example, the adoption of a new technology affects the productivity of inputs in a similar way for each firm, or if natural disasters or economic conditions affect railroad networks comparably. A negative correlation in scale elasticities could be the result of some firms consistently realizing the benefits of new technology before others or if shocks to production allow some firms to attract more efficient inputs than others. Overall, since these estimates of correlation are generally large and vary in magnitude, allowing correlation among these parameters is important to properly describe markups and scales.

Robustness Checks

Table 7 presents estimates of markups, scale elasticities, and elasticities of network characteristics for four separate specifications. The OLS and 2SLS specifications both assume that markups and scales are constant across both firms and time, while the Linear and Flexible Trend specifications allow these parameters to vary by both firm and year. The median of the average markup and scale for 2012 are presented in the Linear and Flexible Trend results. The 2SLS model is the primary common-parameter model presented in Section 7.1 and the OLS specification simply runs only the second stage and ignores instrumentation. The Flexible Trend model is my primary model that was presented in Section 7.2, and the Linear Trend model is similar but allows the central tendency of markups and scales to drift linearly through time. Further explanation of the Linear Trend model and complete results can be found in the Appendix. Estimates are generally similar across specifications but differ in some important ways. First, the OLS estimates of markups and scales are lower than the 2SLS estimates, indicating the need for an instrumental variables approach to correct for endogeneity. Second, including firm heterogeneity in markups and scale elasticities yields higher average markups and scales, presumably because of skew in the distribution of markups and scales across firms and because these results only describe average markups in 2012, which showed evidence of higher than average markups in previous years. Finally, the Linear Trend model produces higher estimates of markups and scales than the Flexible Trend model because it attempts to apply a trend to these parameters when no clear trend may exist, forcing more recent estimates to be higher. Overall, because of its flexibility and theoretical basis, I prefer the Flexible Trend specification presented earlier in the paper.

TABLE 7. Comparison of Results

	OLS	2SLS	Linear	Flexible
	(all years)		(2012 only)	
μ	1.278	1.366	1.557	1.49
η	0.627	0.718	0.904	1.018
Average length of haul	-0.007	-0.007	0.001	0
Percent unit	0.129	0.122	-0.165	-0.228
Percent bulk	0.049	0.065	0.096	0.121
Network size	0.208	0.134	0.063	-0.263

Conclusion

The railroad industry has undergone massive changes since its partial deregulation. With the introduction and improvement of competing forms of transportation, regulatory change was needed to keep firms viable and led to massive consolidation of firms, abandonment of routes, and increased flexibility over pricing. While these changes lowered costs and improved the outlook of the industry, there have been many concerns that railroads are charging excessive rates, especially to captive shippers.

The methods used by railroad regulators to identify evidence of non-competitive pricing among firms have been heavily scrutinized in recent years. Due both to the lack of theoretical foundation and practical application of these methods, there has been a growing need for robust techniques to investigate pricing behavior and market structure. Many successful studies have been conducted that bring sound theoretical models derived from economic principles together with robust econometric techniques to investigate various phenomena in rail markets, but to my knowledge no published work has been able to successfully obtain consistent estimates of markups and scale elasticities.

This research estimates a model of production to obtain estimates of markups and scales for each firm and in each year. I first find that these parameters show significant heterogeneity across firms and time, indicating the need to model this variation. Next, I find that most firms charge prices well in excess of marginal costs; while recent markups are lower than for the majority of the 2000s, I still find markup estimates between 34% and 55% in 2012. Finally, I find that some firms have made efforts to increase capacity, but most firms have filled excess capacity and are operating near minimum efficient scale. While these results provide broad insight into the productive capabilities of firms, it would is useful to know specifics of how those capabilities change and factors that drive that change. In order to further investigate the production capabilities of railroads and how those abilities have evolved over time, I turn to my third chapter, "Decomposing Changes in Productivity Using Bayesian Methods."

CHAPTER III

DECOMPOSING CHANGES IN PRODUCTIVITY USING BAYESIAN METHODS

Abstract

Productivity and its growth are central to the long-term growth, and long-term viability of firms and industries. Partial deregulation of railroads was led by concerns that existing regulation and changes to the industry led to stagnation in productivity. Policy changes made it easier for firms to increase productivity through broad organizational changes like mergers and abandoning unprofitable routes as well as specific technological innovation through the 1980s and early 1990s. However, as the industry has become increasingly consolidated and as more lines have been abandoned, firms may need to rely on technological change to increase productivity. I develop and estimate a model that separates changes in productivity due to innovation and those caused by non-innovative factors and use Bayesian estimation. This allows productivity and technology to evolve flexibly across firms and through time, allowing an examination of changes in railroad productivity and identification of its driving component. I find that every Class I railroad has experienced growth in productivity since 1999. Improvements in technology were the driving factor in the growth of BNSF, KCS, Soo Line, and UP, while CN, CSX, and NS saw significant growth due to broad organizational changes. Finally, I develop a metric that determines whether firms substitute inputs towards factors that innovation makes more productive. I estimate the probability that each firm takes that action to be around 50% with no discernible pattern over time, providing evidence that firms don't anticipate technological change or don't adjust input allocation to take advantage of innovations.

Introduction

The productivity of firms is the amount of real output that can be produced with a marginal increase in real inputs. It has long been of interest to researchers, regulators, and industry analysts alike. Productivity allows firms to produce their products at lower cost and is also the source of long-term economic growth. Further, the level and growth of productivity informs regulators and is central in their regulatory mandate. Prior to its partial deregulation, the railroad industry was faced with many concerns of viability and low levels of productivity. While productivity growth was rapid through the mid-1900s, it had slowed dramatically by the mid-1970s due to the rise of competing modes of transportation and changes to the types of products being shipped. With the goal of reducing costs and increasing productivity, the industry was partially deregulated in 1980. The immediate effects of this policy have been studied extensively,[1] and it is clear that there has been rapid productivity growth since partial deregulation (Wilson, 1994).

The more recent effects of partial deregulation have not been examined. Immediately following partial deregulation, it was relatively easy for firms to merge (thereby taking advantage of economies of scale) and ceasing service on unprofitable routes (Bitzan and Wilson, 2007). The number of Class I railroads fell from 40 in 1980 to just 7 in 1999, and the total size of the network controlled by these carriers dropped from 164,822 miles in 1980 to 95,391 miles in 2013 (United States Surface Transportation Board, 2015). While these changes have dramatically reduced costs in the industry and improved its viability, there is relatively little room for to continue realizing productivity growth through these broad changes (Bitzan and Keeler, 2007). Instead, firms may need to improve their production technologies through innovation and substitute towards more productive

[1]For further reading, see Bitzan and Keeler (2007), Winston et al. (1990), and Barnekov and Kleit (1990).

inputs in order to realize continued growth, which is vital to the sustained viability of the industry.

In order to separately identify changes in productivity, I develop a model that incorporates inefficiency and also allows productivity and technology to vary across firms and evolve in a flexible way over time. Using a theoretical framework, I decompose changes in production into increased use of inputs, input substitution, increased productivity due to technological change, and increased productivity due to non-innovative factors. I then estimate my model using Bayesian methods. This allows me to identify and decompose productivity changes for each firm and each year. I am not aware of any published research that empirically separates growth in productivity due to innovation and that due to factors other than innovation. Productivity growth due to technological change becomes increasingly important as an industry matures and other methods of increasing productivity like merging with or acquiring other firms become less feasible. Further, firms can take potentially take greater advantage of innovations by substituting towards inputs that changing technology makes more productive. I identify a condition under which firms substitute towards more productive inputs and estimate the probability that firms take that action for each year. Finally, these models provide estimates of total factor productivity and its growth, which are key values for informing regulation and give insight into developments in the industry.

The models I estimate allow productivity growth to vary flexibly both across firms and over time by imparting structure on its dynamics; specifically, I allow productivity and technological parameters to follow random-walks with drift. Ignoring the effect of technological change, I find that the Canadian National (CN) railway showed the strongest productivity growth since 1999 at a rate of 3.551% per year. All other railroads exhibited modest productivity growth, between 0.235% and 2.474% per year. After accounting for technological innovation, I am able to identify how much of productivity growth is due to changing technology and how much is due to neutral shifts in the production technology. I find that CN, CSX, and Norfolk Southern (NS) railways experienced strong growth in productivity caused by factors other than technological growth; all other railroads showed decreasing productivity due to these non-innovative factors. Burlington Northern Santa Fe (BNSF), Kansas City Southern (KCS), Soo Line, and Union Pacific (UP) found significant increases in productivity due to technological change, with growth between 30% and 60% between 1999 and 2014. CN, CSX, and NS experienced smaller productivity gains due to changing technology. Overall, when considering total productivity due to all factors, CN and KCS have shown the strongest growth in productivity driven mostly by factors other than technological innovation, though both have seen decreases since 2011. BNSF showed modest total productivity growth due to technological change, and all other firms had constant total productivity. Finally, I find that firms have about a 50% chance of shifting resources towards inputs that innovation makes more productive. This provides evidence that firms don't anticipate technological changes, aren't able to substitute inputs fast enough to capitalize on innovations, or that input prices tend to offset changes in technology.

I estimate three different models with varying degrees of flexibility in the dynamics of productivity change and technological growth. The first model assumes the productivity of each firm follows a simple linear trend, the second allows productivity to follow a random-walk with drift while holding technology constant, and the third allows both productivity and technology to follow a random-walk with drift. Using Bayesian model selection, I find that the model allowing both productivity and technology to evolve flexibly over time has the greatest probability of being the true model, indicating the importance of controlling for technological change. Using estimated model probabilities, I conduct Bayesian model averaging of the results of each model and find that each firm likely experienced modest growth in productivity between 1999 and 2014, with median estimates ranging from 0.296% to 0.719% per annum. However, I estimate the probability that all firms experienced productivity growth is 46.783%, indicating that at least one firm likely saw a decrease in productivity over the sample period.

This paper begins with an overview of the railroad industry and its regulation. Following this, I provide a review of the relevant literature, covering both the methods used to measure productivity and how productivity has been studied in the railroad industry. I then develop my theoretical model and proceed to present the data used in this analysis. I cover each of the three empirical models presented in this paper, then show and explain my results. A conclusion of my findings follows.

Institutional Background

The railroad industry has been federally regulated since 1887. The Interstate Commerce Commission (ICC) was created in response to concerns of excessive rates, market power, and discriminatory pricing in the industry with the passage of the Interstate Commerce Act (ICA) of 1887. This policy gave the ICC control over collective rate making and oversight over mergers and provided a channel through which the reasonability of rates charged by railroads could easily been questioned by shippers. Through most of the 1900s, these regulations helped promote competition in the industry and kept shipping rates low while still allowing railroads to be profitable.

By the 1970s, the regulations that once promoted competition impeded firms in the industry. Not only had new competing modes of transportation such as air, barge, and trucking been developed and improved, but plastics, which are much less dense than goods shipped in the past, constituted a greater proportion of all goods shipped.

Consequentially, railroad costs rose to the point that rate regulation prevented firms from being cost viable. In an effort to save the industry, railroads were partially deregulated with the passage of the 4R Act in 1976 and the Staggers Act in 1980. These policies allowed railroads to merge more easily to reduce costs through economies of scale, gave firms the ability to negotiate contracts and generally provided greater pricing flexibility, and allowed firms to more easily abandon lines on which operations were not profitable.[2]

Partial deregulation resulted in many drastic changes to the industry. The number of Class I railroads fell from 40 in 1980 to just 7 in 1999, mostly through acquisitions and mergers. The total size of the network controlled by Class I railroads fell from 164,822 miles in 1980 to 95,391 miles in 2013, largely through the abandonment and sale of unprofitable lines. The average length of haul increased from 615 miles in 1980 to 973 miles in 2013 (United States Surface Transportation Board, 2015). Overall, individual railroad networks were larger, the total size of the network grew smaller, and shipments were traveling longer distances. As a result of these changes, rail shipping rates fell dramatically, from \$0.0646 per revenue-ton-mile in 1980 to \$0.0329 in 2014 (United States Surface Transportation Board, 2015). The reduction in prices is largely reflective of a reduction in railroad costs and improvements in productivity (Bitzan and Keeler, 2007).

Following rapid changes that occurred in the railroad industry through the 1980s and early 1990s, the general structure of the industry has mostly remained constant since 1999. Only seven Class I railroads remained in 1999, and additional mergers have not occurred. Most of the lines on which operations were unprofitable have been abandoned or sold to short-line regional railways (Tretheway et al., 1997). As a result, there is little room for railroads to improve their productivity on those fronts. Thus, to remain viable, firms have

[2]There have been a considerable number of studies that describe these policies and their effects. See, for example, Bitzan and Wilson (2007), Schmalensee et al. (2015), Wilson (1997), and Winston et al. (1990).

turned towards other channels, such as technological progress, to realize productivity gains and further reduce costs. As an example, the elimination of cabooses, a remnant of the age of steam locomotives that required a crew to operate, resulted in a reduction in costs by between 5% and 8% between 1983 and 1997 (Bitzan and Keeler, 2003). Railroads have also invested $575 billion in infrastructure and equipment since 1980; recently, nearly 2700 new locomotives were purchased between 2008 and 2012, and many innovations have been made in safety, fault detection, and performing maintenance that preempts equipment failure (AAR, 2013). To my knowledge, there has been no published research that considers the effects of these recent innovations.

Literature Review

This research investigates productivity of the railroad industry using a stochastic frontier model. In this section, I provide a history and review of studies and methods used to estimate productivity in general. I then describe research that has investigated the productivity of railroads and the effects of the industry's partial deregulation. Finally, I describe stochastic frontier models and how they have been used to study productivity and separate it from inefficiency.

Total Factor Productivity

Productivity has rightfully long been a focal point in many branches of economics; various aspects of productivity can inform on the effectiveness with which inputs can be transformed into outputs as well as the overall efficiency of production. Total factor productivity has been studied extensively and provides a useful metric: The number units of real output a firm can produce with one unit of real inputs (Jorgenson and Griliches, 1967). The value of this measure can be easily seen; it can be used to evaluate economies of scale, trends provide information about growth rates, and heterogeneity across firms can shed light on factors that affect productivity and costs.

The notion of total factor productivity was created to explain economic growth. Growth can either be the result of increased use of inputs, usually called capital accumulation,[3] or growth in productivity. In light of limited resources, increases in productivity are the only way to sustainably promote economic growth.[4] Empirical findings have shown that productivity is the main cause of changes in economic growth; in his seminal paper, Solow (1957) found that between 1909 and 1949, approximately one-eighth of the variation in output was due to capital accumulation and seven-eighths was due to changes in productivity. Further, he estimated that annual productivity growth rates ranged from -7.6% to 7.2%, at an average of 1.5% per annum. Finally, Solow estimates several forms for the production function. Using a log-linear (i.e., Cobb-Douglas) form, he estimates that the level of productivity was approximately 0.482.

Productivity has been estimated using a variety of methods. Solow's seminal work on productivity suffered from a number of practical issues. Most notably, any deviations from the empirical model (i.e., residuals) were assumed to be the result of differences in productivity (Solow, 1957). Of course, there are many additional sources of error including differences in efficiency and measurement error. Further, Solow used a linear approximation in his analysis, which does not allow inputs to exhibit complementarity or substitutability and can result in large approximation errors. Several models have been

---

[3]While non-capital inputs (e.g., labor) can also increase output, economists have historically not attributed long-term growth to those factors since the stock of those inputs tends to grow at a relatively slow rate.

[4]If resources are limited, capital cannot be endlessly accumulated, so growth must come from another source.

developed and extended to address these issues and largely fit into two groups, either parametric or non-parametric.[5]

Parametric models assume a specific form for the production function and aim to decompose shifts in the production frontier into changes in productivity and efficiency and measurement error. There has been an abundance of research that estimate translog cost functions and infer changes in productivity from shifts in the cost function. Since translog cost functions are a second-order approximation of the true cost function, this method reduces the approximation error presents in Solow's work. Caves et al. (1981b) used this framework to derive an expression for productivity growth that depends on the change in costs and change in output over time. The authors estimate these parameters in a cost function and use them to calculate productivity growth in the U.S. passenger and freight rail industry. Cost function frameworks similar to this have been used to study productivity in a variety of contexts.[6] While this framework is very flexible, it cannot separately identify productivity from inefficiency. Stochastic frontier (SF) models, which are further explained in Section 3.3, extend the standard translog estimation framework to include inefficiency; by noting that efficiency must lie between 0% and 100%, structure can be imparted on the model that allows productivity and efficiency to be separately identified (Aigner et al., 1977).

Non-parametric methods remain agnostic of the specific functional form of the production function and instead rely on non-parametric methods to infer its shape. Data envelopment analysis (DEA) is most commonly used to infer the production frontier. This method assumes that production plans lie on the frontier and uses linear programming

[5]Of course, semi-parametric models, which have some parametric and some non-parametric components, have also been used. For further reading, see Jondrow et al. (1982) and Park and Simar (1994).

[6]For more examples of studies of the railroad industry that use translog cost functions, see Bitzan and Wilson (2007), and Bitzan and Keeler (2007).

techniques to trace out the exact location of the frontier (Friesner et al., 2006).[7] Lim and Lovell (2009) use DEA to investigate short-run profit changes in the rail industry. By using non-parametric methods to identify the production frontier, the authors decompose changes in profit into changes in price and productivity.

Both parametric methods like SF and non-parametric methods like DEA commonly appear in the literature. Eisenbeis et al. (1999) compares the two in the context of the banking industry and finds that while the level of estimated inefficiency is higher under DEA, the two measures are highly correlated, indicating they capture similar information. However, the authors also find that estimates from SF analysis more accurately capture efficiency in management and preferences for risk than do linear programming methods.

Stochastic Frontier Models

Stochastic frontier (SF) models extend the productivity estimation framework in two important ways. First, they assume that firms are not necessarily efficient; these analyses distinguish maximum possible output from actual output and term the deviation between the two inefficiency. The seminal work of Aigner et al. (1977) presents a commonly used formulation of the stochastic frontier framework:

$$q_i = f(x_i; \beta) + \varepsilon_i - \delta_i. \tag{3.1}$$

In equation (3.1), q_i represents log output, x_i are inputs, β are parameters describing production, ε_i is a productivity, and δ_i is inefficiency. Under usual error assumptions, it would be impossible to separately identify ε_i from δ_i; however, inefficiency, defined

[7] More specifically, the basic DEA model assumes *non-dominated* plans, for which no other plan produces more output with fewer inputs, lie on the frontier. Extensions of this model have been made to include inefficiency.

as deviation from maximum output, is inherently one-sided. Using this assumption and modeling δ_i as, for example, half-normal or log-normal allows productivity shocks and inefficiency to be separately identified.

Similar to other parametric methods of estimating production, the problem of estimating the form of the production function remains. Unlike non-parametric methods like DEA, parametric models require a the researcher to assume a specific form of the production function; in order to maintain minimal assumptions on the exact shape of the production function, researchers have historically utilized some form of functional approximation to address this problem. Early research into productivity, such as Solow (1957), used a first-order Taylor approximation on the log-production function, also known as the Cobb-Douglas form. While this form provides a good starting approximation, it does imply the assumption that production is additive in the inputs; that is, the Cobb-Douglas form assumes that the productivity of any given input depends only on the amount of that input being used and not on any other inputs. Christensen et al. (1973) test simple functional forms that assume additivity and constant returns to scale against a more flexible second-order Taylor approximation of the log-production function, also known as the translog form. The authors use data describing private production in the United States from 1929 to 1969 and find that the assumption of additivity is clearly not satisfied, leading to bias when a first-order approximation is used. As a result, it is safer to use a more flexible functional form such as translog or even higher-order approximations if the assumption of additivity is not clearly satisfied.

Stochastic frontier models have been extended in a number of ways and have been used to identify differences or changes in productivity and efficiency. Schmidt and Lovell (1979) separately estimate the production frontier and the cost function to separate technical inefficiency, which originates in the transformation of inputs into outputs, from allocative inefficiency, which occurs when inputs are not used in the optimal proportions. Applying this model to steam-electricity generation, there was evidence that both types of inefficiency were significant: Technical inefficiency raised costs by about 8.5% while allocative inefficiency raised costs by about 9.2%.

Kumbhakar (1988) adapts the technical/allocative inefficiency framework to panel data, assuming that productivity is constant across all firms and time but that technical inefficiency varies by firm. The author applies this model to Class I Railroads and estimates input demand to correct for possible endogeneity and separate technical and allocative inefficiency. As expected, the author finds sizable variation in inefficiency across firms.

Many researchers have found success in using Bayesian methods to estimate SF models. Generally, other estimation procedures such as maximum likelihood estimation can produce unstable estimates of SF model parameters (van Den Broeck et al., 1994). Further, the parameter uncertainty expressed by standard methods may not be accurate, especially for small sample sizes (Koop et al., 1995). Bayesian methods are able to properly express parameter uncertainty for large and small samples alike and tend to produce more stable estimates. Recent research by Yan et al. (2009) have introduced Bayesian estimation and extended the SF framework to analyze to analyze panel data and models productivity and inefficiency in a flexible manner. The authors assume productivity follows a deterministic trend shared by all firms and inefficiency is a random effect across firms with structural breaks across time. In using this model to analyze container ports, the author finds productivity increases by about 4.4% per year and that inefficiency showed heterogeneity both other firms and across time.

Productivity of Railroads

Productivity of railroads has long been a topic of interest to regulators and researchers. Worries about the efficiency and productivity of railroads was a major impetus for the partial deregulation of the railroad industry in the 1970s (Bitzan and Wilson, 2007). Proponents of deregulation argued that because of the development and improvement of other modes of shipping like planes, barges, and trucks and because of changes in the mix of products being shipped, existing regulation intended to promote competition for the majority of the 20th century were hindering efficiency and limiting the cost-viability of the industry (Winston et al., 1990). Following deregulation, firms were more easily able to take advantage of economies of scale by merging and could reduce costs by abandoning rail lines that were unprofitable (Bitzan and Keeler, 2007).

Of course, there has been much interest in how these changes have affected productivity. Additionally, in the light of the different characteristics and actions that firms took after deregulation, there is interest in how firms differentially progressed following deregulation and what factors led to those differences. Finally, the ultimate prospects for the industry remain unclear; recent declines in aggregate demand have further cut into firm profits and other modes of transportation continue to improve.

Many studies that examine railroad productivity have been conducted, and questions about many aspects of the industry have been addressed. Caves et al. (1981b) began the investigation into changes in the industry and its effect on productivity and viability. The authors found that the industry was quickly becoming more productive prior to 1963; productivity growth was estimated to be 3.5% per year on average during the period from 1955 to 1963. However, in the following period from 1963 to 1974, productivity grew at a much slower rate, only 0.6% per year on average. The authors posit that in the early period, many firms began small and were able to exercise economies of scale as they grew through the late 1950s and early 1960s. The growth of the size of these firms slowed and most excess capacity was filled by the mid 1960s, leading to slower productivity growth through the mid 1970s. In accompaniment with changes in the industry, this slowing growth led many to worry about its ultimate survival and spurred its partial deregulation.

A crucial question for regulators is whether and by how much partial deregulation helped the industry. Tretheway et al. (1997) indirectly address this question by examining productivity and performance of Canadian railways, which underwent partial pricing deregulation in 1967, and compared with U.S. railroads, which were partially deregulated later in 1976 and 1980. While Canadian railroads had significantly higher productivity growth than U.S. railroads prior to the deregulation of the U.S. rail industry, U.S. railroads saw productivity growing between 1.3% and 1.5% per annum faster than Candian railways between 1981 and 1988. The authors conclude that this increased growth was due to reductions in the amount of inputs being used as well as higher traffic density in the U.S. While partial deregulation was not *necessarily* responsible for these changes, it did provide an environment where firms could more easily merge, thereby taking advantage of economies of scale, and had greater flexibility in abandoning routes, which could have lead railroads to find a more advantageous traffic density.

Further, while it was clear that some kind of intervention was needed to ensure the viability of the U.S. rail industry, there have been questions over exactly what type of intervention would be most beneficial to firms. Apart from deregulation, which aims to utilize free-market principles to improve the efficiency and viability of firms, the most commonly suggested intervention is public ownership of railways. Public ownership of the entire rail industry has not been investigated since such a program has not been enacted, but several studies have looked into public ownership of firms and appropriation of public funds towards private railways. Caves and Christensen (1980) compare the publicly-owned

Canadian National Railway (CN) with the privately owned Canadian Pacific Railway (CP). Opponents of public ownership worry that firms won't face the proper incentives to minimize costs and improve efficiency. The authors found that competitive pressures both between CN and CP and from other modes of transportation were very strong; as a result, both railways experienced similar productivity growth. Due to an abundance of unprofitable lines, CN initially had lower productivity at the beginning of the sample in 1956; however, both CN and CP aggressively abandoned track through 1967 and saw their productivities converge and continue to grow at a similar rate through 1974.

Rather than owning railways outright, governments can appropriate funds towards supporting rail operations. Similar to completely publicly-owned firms, railroads that receive subsidies may not have the incentive to minimize costs and maximize efficiency absent sufficient competitive pressures. Oum and Yu (1994) consider railways in nineteen OECD countries[8] and investigate the effect of public funding and firms' autonomy from their public funders on efficiency. Since many of the firms did not see competition from other railroads or other forms of competition, efficiency tended to be higher for less publicly funded firms and for firms that had a greater degree of autonomy. In all, whether public funding or ownership is beneficial or detrimental is extremely dependent on whether firms will face competition; when they don't, railways will not have the incentive to improve and will tend not to do so as a result.

Conceptual Framework

Fundamentally, productivity research compares production plans and determines how much of the difference in output is due to increased use of inputs and how much is due to changes in productivity. An example of this decomposition is shown graphically in Figure

[8]Contrary to many studies of the U.S. rail industry, the railways in this study transport mostly passengers rather than freight.

1. In this example, I consider an output that uses two inputs X_i and X_j. In practice, the

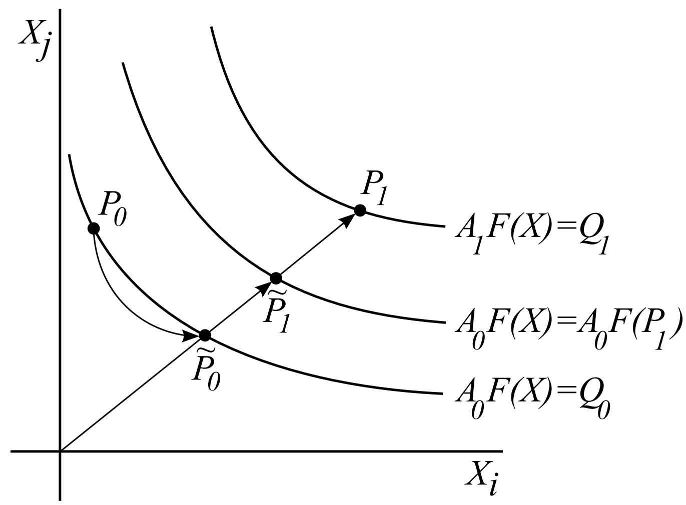

FIGURE 7. Input Substitution and Productivity

researcher is faced with two production plans, given by P_0 and P_1 in the figure, and knows their associated levels of output Q_0 and Q_1. The two plans are on two different isoquant curves, given by $A_0F(X) = Q_0$ and $A_1F(X) = Q_1$. These differ only by a productivity factor, and the researcher's goal is to estimate the growth of productivity from A_0 to A_1. In changing production from P_0 to P_1, the firm can first change the composition of inputs it uses to most efficiently produce P_1; this is called input substitution and is shown in the graph by the shift from P_0 to $\widetilde{P}_0$. Since P_0 and $\widetilde{P}_0$ are on the same isoquant, they produce the same level of output, i.e., $A_0F(P_0) = A_0F(\widetilde{P}_0)$. The firm can also increase the amount of inputs it uses; this is shown in the graph by the movement from $\widetilde{P}_0$ to $\widetilde{P}_1$, where $A_0F(\widetilde{P}_1) = A_0F(P_1)$. Then, the proportional growth in output due to increased inputs is

$$\frac{A_0F(\widetilde{P}_1)}{A_0F(\widetilde{P}_0)} = \frac{F(P_1)}{F(P_0)}. \tag{3.2}$$

Finally, output can increase due to changes in productivity, shown in the graph by the shift from $\widetilde{P}_1$ to P_1. The proportional growth in output due to the increase in productivity is given by

$$\frac{A_1 F(P_1)}{A_0 F(\widetilde{P}_1)} = \frac{A_1 F(P_1)}{A_0 F(P_1)} = \frac{A_1}{A_0}. \tag{3.3}$$

Overall, the total proportional change in production is given by

$$\frac{Q_1}{Q_0} = \frac{A_1}{A_0}\frac{F(P_1)}{F(P_0)}. \tag{3.4}$$

To reiterate, the ratio A_1/A_0 represents the proportional increase in productivity while the quotient $F(P_1)/F(P_0)$ represents the proportional increase in output due to increased inputs.

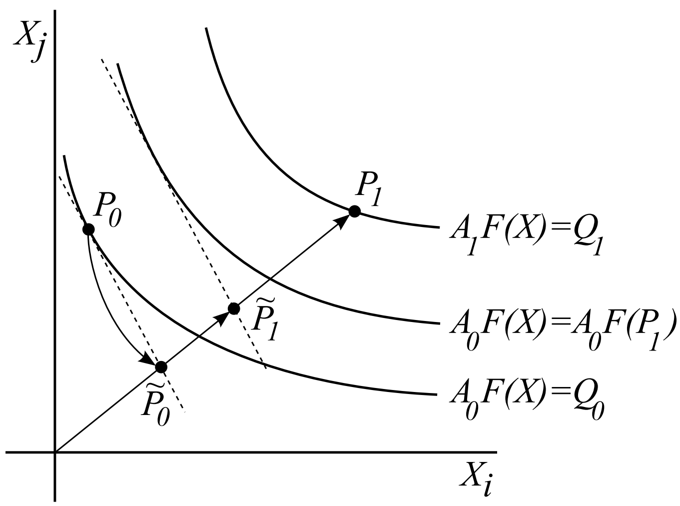

FIGURE 8. Estimating Productivity With Linearization

The above decomposition assumes the researcher knows the shape and position of the isoquants (and therefore also knows the production function). In practice, this is rarely true, and researchers have relied on a number of techniques to approximate or infer the shape of the production function. The use of Taylor approximations has been very prevalent in productivity estimation. These approximations have been popular largely because of their flexibility; apart from differentiability, Taylor approximations make no assumptions on the shape of the production function and as a result may be used in a variety of contexts. Further, by including more terms in the approximation, it can be made as accurate as desired, subject to data restrictions. To see how this affects estimation of productivity, consider Figure 2; the dotted line represents the linearization of the production function around the point P_0. In practice, the researcher does not observe the isoquant $A_0F(X) = Q_0$ but instead can only approximate its form. The researcher would then estimate input substitution as the shift from P_0 to $\widetilde{P}_0$. The proportional change in output due to increased use of inputs would be estimated as the shift from $\widetilde{P}_0$ to $\widetilde{P}_1$. Finally, the change in output due to increased productivity is estimated to be the shift from $\widetilde{P}_1$ to P_1. Comparing these results to Figure 1, we can see that the researcher would overestimate productivity in this example. Naturally, more complex Taylor approximations can be used, which would decrease the approximation bias in productivity estimates.

While having been used extensively in the literature, the simple framework presented above suffers from theoretical and practical issues. First, as illustrated above, productivity estimates could be biased due to errors in approximating the production function. Many have worked to reduce these errors by using higher order approximations, but unfortunately approximation error can never be eliminated because the approximations are never exact. As noted previously, some researchers have found success in using data envelopment analysis (DEA) to non-parametrically estimate the production frontier (and therefore production function).

Additionally, the above framework has encountered issues in estimation. Specifically, researchers have historically used a deterministic production function and have inferred that any deviations from that function (i.e., residuals) are due to differences in productivity. Of course, there are many channels through which errors can propagate. For example, in addition to approximation and measurement error, inefficiency, defined as deviation from maximum possible output, can affect the level of production independent of changes in productivity. As discussed in greater length in Section 3.3, stochastic frontier (SF) analysis works to separately identify these various sources of error by imparting structure on their form.

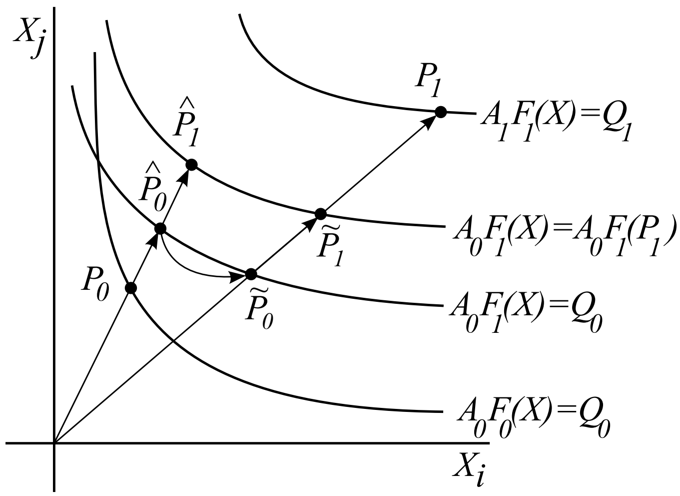

FIGURE 9. Productivity and Changing Technology

Finally, the method of identifying productivity described above ignores the possibility that the production technology could change. This will lead to a bias in the estimate of productivity. As an example, consider Figure 3, which shows two isoquants that describe different production technologies. The function F_1 represents the new production technology and F_0 represents the original. Since the isoquant is relatively less steep under F_1 than F_0, the input X_j is more productive under the new production technology. The firm increases production from Q_0 to Q_1 through a few different channels. First, the production technology changes, which changes the productivity of inputs, which in turn affects output. The proportional change in productivity due to innovation is shown in Figure 3 by the movement from P_0 to $\widehat{P}_0$. Notice that in this example, the production technology is becoming less efficient since it requires more of both inputs to produce the quantity Q_1, holding productivity constant. Then, the firm can substitute inputs to more efficiently produce Q_1, which is shown by the movement from $\widehat{P}_0$ to $\widetilde{P}_0$. Next, the firm can increase the amount of inputs it uses, shown in the shift from $\widetilde{P}_0$ to $\widetilde{P}_1$. In practice, the researcher cannot separately identify the change in technology from the increased use of inputs because those changes occur simultaneously. However, the researcher can observe the sum of these effects, shown in the graph by the movement from P_0 to $\widehat{P}_1$. The proportional change in output due to innovations and increases in inputs is quantified by

$$\frac{A_0 F_1(\widehat{P}_1)}{A_0 F_0(P_0)} = \frac{F_1(P_1)}{F_0(P_0)}. \tag{3.5}$$

Finally, the change in productivity due to factors other than technological change is shown by the movement from $\widetilde{P}_1$ to P_1, which can be written as

$$\frac{A_1 F_1(P_1)}{A_0 F_1(\widetilde{P}_1)} = \frac{A_1 F_1(P_1)}{A_0 F_1(P_1)} = \frac{A_1}{A_0}. \tag{3.6}$$

The total change in production can then be expressed as

$$\begin{aligned} \frac{Q_1}{Q_0} &= \frac{A_1}{A_0}\frac{F_1(P_1)}{F_0(P_0)} &(3.7)\\ &= \frac{A_1}{A_0}\frac{F_1(P_1)}{F_1(P_0)}\frac{F_1(P_0)}{F_0(P_0)}. &(3.8) \end{aligned}$$

Here, A_1/A_0 represents the proportional change in productivity, $F_1(P_1)/F_1(P_0)$ represents the proportional change in output due to increasing inputs, and $F_1(P_0)/F_0(P_0)$ represents the proportional change in output due to changing technology using the inputs P_0. Thus, conditional on approximating the production function and how it evolves over time, the researcher can separately identify changes in output due to increased inputs, improvements in technology, and increases in productivity due to non-innovative factors.

In my analysis of changing productivity, there are a few key values of interest derived above. First, $F_1(P_0)/F_0(P_0)$ denotes the change in productivity due to innovation and A_1/A_0 represents the proportional change in productivity due to factors other than technological change. As a result, the product $A_1F_1(P_0)/A_0F_0(P_0)$ is the total change in productivity due to any factor, which I refer to as *technology-inclusive productivity growth.* I also focus on another value, $F_1(P_1)/F_0(P_1)$, which measures technology's contribution to productivity growth using the new inputs P_1. By comparing this value to $F_1(P_0)/F_0(P_0)$, inferences can be made about the benefits of input substitution:

– If $F_1(P_0)/F_0(P_0) > F_1(P_1)/F_0(P_1)$, the new technology increases output more for the original plan than for the new plan. Thus, the firm substituted towards inputs that innovation made less productive.

– If $F_1(P_0)/F_0(P_0) < F_1(P_1)/F_0(P_1)$, the new technology increases output more for the new plan than for the original plan. This indicates that the firm substituted towards factors that technology change made more productive.

Next, I turn to explaining how I approximate the shape of the production function. Consider the output of a firm i in year t, given by Q_{it}. Suppose that the firm's production technology is described by

$$Q_{it} = A_{it}F_t(X_{it}; \Phi_{it})\Delta_{it}, \tag{3.9}$$

where α_{it} is a productivity factor, X_{it} is a vector of inputs, φ_{it} is a vector of network characteristics, and Δ_{it} is a constant between zero and one describing efficiency. While it is not possible to determine the exact shape of F_t, one can approximate it using the second order Taylor approximation of $\ln Q_{it}$ around zero:

$$
\begin{aligned}
q_{it} \approx\ & \alpha_{it} - \delta_{it} + \sum_j \frac{\partial \ln F_t}{\partial \ln X^j} x_{it}^j + \sum_j \frac{\partial \ln F_t}{\partial \ln \Phi^j} \varphi_{it}^j \\
& + \frac{1}{2!} \sum_j \sum_k \frac{\partial^2 \ln F_t}{\partial \ln X^j \partial \ln X^k} x_{it}^j x_{it}^k \\
& + \frac{1}{2!} \sum_j \sum_k \frac{\partial^2 \ln F_t}{\partial \ln \Phi^j \partial \ln \Phi^k} \varphi_{it}^j \varphi_{it}^k \\
& + \frac{1}{2!} \sum_j \sum_k \frac{\partial^2 \ln F_t}{\partial \ln X^j \partial \ln \Phi^k} x_{it}^j \varphi_{it}^k. \qquad (3.10)
\end{aligned}
$$

Here, lower-case variables are log-transformed versions of upper case variables, and superscripts index vectors of variables. As an exception, $\Delta_{it} = \exp(-\delta_{it})$, and δ_{it} is restricted to be positive to ensure that the efficiency term Δ_{it} is between zero and one. I first assume that inputs and network characteristics are separable in production, so that $\frac{\partial^2 \ln F_t}{\partial \ln X^j \partial \ln \Phi^k} = 0$ for all j and k. Also, under a modest assumption on F_t,[9] the second derivatives of F_t will be symmetric.[10] Using these assumptions, equation (4.2) becomes

$$
\begin{aligned}
q_{it} \approx\ & \alpha_{it} - \delta_{it} + \sum_j \frac{\partial \ln F_t}{\partial \ln X^j} x_{it}^j + \sum_j \frac{\partial \ln F_t}{\partial \ln \Phi^j} \varphi_{it}^j \\
& + \frac{1}{2} \sum_j \frac{\partial^2 \ln F_t}{\partial (\ln X^j)^2} (x_{it}^j)^2 + \frac{1}{2} \sum_j \frac{\partial^2 \ln F_t}{\partial (\ln \Phi^j)^2} (\varphi_{it}^j)^2 \\
& + \sum_j \sum_{k>j} \frac{\partial^2 \ln F_t}{\partial \ln X^j \partial \ln X^k} x_{it}^j x_{it}^k \\
& + \sum_j \sum_{k>j} \frac{\partial^2 \ln F_t}{\partial \ln \Phi^j \partial \ln \Phi^k} \varphi_{it}^j \varphi_{it}^k. \qquad (3.11)
\end{aligned}
$$

[9]Specifically, I assume that the second derivatives of F_t are continuous in a neighborhood of zero.

[10]That is, $\frac{\partial^2 \ln F_t}{\partial \ln X^j \partial \ln X^k} = \frac{\partial^2 \ln F_t}{\partial \ln X^k \partial \ln X^j}$ for all j and k.

Renaming partial derivatives with respect to inputs β^t and those with respect to network characteristics θ^t, equation (3.11) can be rewritten to arrive at the familiar translog form:

$$
\begin{aligned}
q_{it} \approx \alpha_{it} - \delta_{it} + \sum_j \beta_j^t x_{it}^j + \sum_j \theta_j^t \varphi_{it}^j \\
+ \sum_j \beta_{jj}^t (x_{it}^j)^2 + \sum_j \theta_{jj}^t (\varphi_{it}^j)^2 \\
+ \sum_j \sum_{k>j} \beta_{jk}^t x_{it}^k x_{it}^j + \sum_j \sum_{k>j} \theta_{jk}^t \varphi_{it}^k \varphi_{it}^j.
\end{aligned}
\tag{3.12}
$$

Let x_{it} be the matrix of all log-inputs, all log-inputs squared, and all of the interactions between log-inputs (i.e., containing each x_{it}^j, $(x_{it}^j)^2$, and $x_{it}^j x_{it}^k$), and let φ_{it} be similarly defined. Then, equation (4.3) can be expressed in vector form as

$$q_{it} \approx \alpha_{it} - \delta_{it} + x_{it}\beta_t + \varphi_{it}\theta_t. \tag{3.13}$$

Naturally, there is some error incurred in the approximation and measurement of q_{it}. I label this approximation error ε_{it}, so that

$$q_{it} = \alpha_{it} - \delta_{it} + x_{it}\beta_t + \varphi_{it}\theta_t + \varepsilon_{it}. \tag{3.14}$$

Data

The data used in this analysis come from R1 forms, collected and presented by the United States Surface Transportation Board (STB). These forms are published annually and contain financial information and operating statistics for all Class I railroads, including aggregate output and input use and characteristics of each firm's network. The time span of the sample has been restricted to the period from 1999 to 2014; this analysis is interested in how productivity has evolved since Class I railroads fully merged in 1999.

The Class I railroads in this sample are Burlington Northern Santa Fe (BNSF), the Canadian National Railway (CN), CSX Transportation (CSX), the Kansas City Southern Railway (KCS), the Norfolk Southern Railway (NS), the Soo Line Railroad (SOO),[11] and the Union Pacific Railroad (UP).

The dependent variable in this analysis is aggregate revenue-ton-miles, which are defined as one ton of product shipped one mile that generates revenue. Production of revenue-ton-miles is described by input use and network characteristics. I use amounts of locomotives and cars, quantity of fuel consumed, and total hours of labor worked, and investment per mile of road as inputs.[12] Following Friedlaender and Spady (1981), I opt to include investment in firms' networks as an input and include network size as a characteristic of output. The authors found that including network size as input results in negative output elasticities with respect to the network due to economies of density.

Characteristics of each firm's network are crucial in describing production, especially aggregate production, in the railroad industry. Tretheway et al. (1997) investigate the effect of aggregation on the estimates of productivity in the rail industry. The authors estimate productivity using both aggregate and disaggregate data and found significant differences. Using aggregate output assumes that the mix of products being shipped remains constant over time. If, for example, firms instead shift to shipping products that require fewer inputs, productivity would appear to increase even if productive capability remained constant. One would ideally use disaggregated data in their analyses; however, only aggregated data is publicly available, so it is important to control for other factors

[11]While Canadian Pacific Railway has owned the Soo Line Railroad since 1990, Soo changed in name to Canadian Pacific in the early 2000s; I will continue to refer to this railroad as SOO.

[12]Miles of road is defined as the total length of non-redundant track operated by a railroad. Investment was deflated using the GDP price deflator.

TABLE 8. Descriptive Statistics

	BNSF	CN	CSX	KCS	NS	SOO	UP	Total
Output								
Revenue ton-miles (millions)	589625	50471.232	232187.5	26437.171	190250	27312.274	526687.5	234710.097
	(77897.261)	(8422.489)	(14828.885)	(4909.945)	(11716.086)	(6972.383)	(29992.707)	(221554.046)
Inputs								
Locomotives	6180.875	541	3889.562	545.75	3787.688	421.438	7913.688	3325.714
	(895.528)	(107.54)	(248.982)	(53.432)	(266.951)	(90.79)	(646.264)	(2814.122)
Cars	83837.5	24652.75	94733.75	12846.688	96078.125	14690.812	90103.75	59563.339
	(7253.763)	(6450.657)	(20162.264)	(1585.578)	(11392.32)	(1985.792)	(15344.713)	(38578.147)
Investment per mile of road	1210.875	1425.349	1119.413	911.214	1137.301	612.706	1362.493	1111.336
	(127.323)	(289.596)	(221.539)	(192.393)	(233.919)	(68.127)	(139.185)	(319.765)
Gallons of fuel (millions)	1309.606	100.021	547.779	63.283	482.172	54.176	1235.647	541.812
	(113.96)	(24.114)	(53.895)	(6.05)	(33.314)	(12.274)	(127.394)	(504.872)
Labor hours (thousands)	81289.551	12899.74	57640.454	5726.081	57406.789	6274.362	96616.222	45407.6
	(5996.77)	(1162.888)	(5837.744)	(510.784)	(4212.059)	(1262.238)	(8104.72)	(35030.257)
Network characteristics								
Average length of haul (miles)	1046.849	288.825	533.266	363.193	464.402	440.834	921.168	579.791
	(67.523)	(22.837)	(40.975)	(43.623)	(21.087)	(38.505)	(24.854)	(271.546)
Average train speed (MPH)	18.925	20.5	18.411	17.504	19.066	19.612	22.355	19.482
	(1.74)	(1.085)	(0.706)	(2.571)	(0.7)	(1.904)	(1.566)	(2.135)
Miles of road	32482	5847.812	21783.938	3096.312	20889.375	4123.125	32447.312	17238.554
	(424.674)	(1353.077)	(1038.289)	(163.18)	(693.435)	(1325.3)	(575.021)	(12034.503)
Percent unit train	0.476	0.17	0.333	0.414	0.252	0.277	0.415	0.334
	(0.031)	(0.054)	(0.022)	(0.095)	(0.02)	(0.046)	(0.021)	(0.111)
Percent bulk shipments	0.196	0.295	0.194	0.161	0.156	0.294	0.183	0.211
	(0.009)	(0.1)	(0.011)	(0.012)	(0.007)	(0.034)	(0.014)	(0.068)
N	16	16	16	16	16	16	16	112

that could be correlated with aggregate input use but not necessarily with productive potential.

There are several network characteristics that are important to consider. First, as noted in Tretheway et al. (1997), traffic mix is a crucial feature of railroad networks. The mix of traffic is dependent both on the types of goods being shipped as well as types of shipments that traverse the network. In this analysis I include the percentage of shipments that carry bulk products[13] as a measure of product mix and the percentage of shipments that are unit train shipments as a measure of shipment mix. Bitzan and Keeler (2007) additionally find that the shipment distance is important in describing costs.[14] The percentage of shipments that are unit train shipments partially captures aspects of shipment distances, and I additionally include the average length of haul into this analysis. Friedlaender and Spady (1981) find that network size is a crucial factor in transportation costs, so I also include miles of road for each firm. Finally, the quality of a railroad's track will determine how efficiently trains can traverse the network and can also influence maintenance costs. Following Wilson (1997), I use average locomotive speed as a measure of network quality.

Descriptive statistics for each of these variables are presented in Table 1. Means and standard deviations are given for each firm and as an average over all firms. The sample spans sixteen years, and the descriptive statistics are averaged over time.

[13] I define bulk products as belonging to one of the following categories: Metallic ores, nonmetallic minerals (not fuels), waste/scrap metals, clay/concrete/glass/stone, farm products.

[14] This comes at no surprise since, at the very least, short shipments require more fuel per revenue-ton-mile.

Empirical Models

In the analyzing the productivity and efficiency of firms, we are most interested in estimating their respective parameters α_{it} and δ_{it} in equation (3.14). However, all of the parameters of equation (3.14) cannot be separately identified in a standard regression framework. There are a number of ways to manipulate this model so that productivity and efficiency can be identified, and in this section I describe the three models I use in this paper.

Deterministic Trend in Productivity, Constant Technology

This model first assumes that technology and the effect of network characteristics are constant across time, so that $\beta_t = \beta$ and $\theta_t = \theta$. I also assume that each firm has its own initial productivity which then follows a deterministic linear trend shared by all firms. Further, the model assumes that inefficiency is constant across time (but is allowed to vary by firm); as a result, $\delta_{it} = \delta_i$. Finally, recall that δ_i was restricted to be greater than zero; following the majority of the stochastic frontier literature, I assume δ_i has a half-normal distribution centered and truncated at zero. Inefficiency can be separately identified from productivity both because they have different dynamics across time[15] and because inefficiency is strictly greater than zero. The model can be expressed in the following relations.

$$
\begin{aligned}
q_{it} &= \alpha_{it} + x_{it}\beta + \varphi_{it}\theta - \delta_i + \varepsilon_{it} \\
\alpha_{it} &= \alpha_i + \tau t \\
\alpha_i &\sim N(\mu_\alpha, \sigma_\alpha) \\
\varepsilon_{it} &\sim N(0, \sigma_\varepsilon) \\
\delta_i &\sim N^+(0, \sigma_\delta)
\end{aligned}
$$

[15] That is, inefficiency is assumed to be static while productivity follows a linear trend.

Conditional on having prior distributions over the parameters, this model can be estimated using Gibbs sampling. For a detailed description of the sampler, see Section 9.1.1 in the Appendix.

Random Walk in Productivity, Constant Technology

This model is similar to the previous model, but focuses on modeling productivity in a more flexible way than with a deterministic trend. Specifically, I assume that productivity follows a random walk with drift that is independent for each firm. This type of process has been used in a number of applications and can model many processes, especially those that exhibit persistence, flexibly and effectively.[16] Importantly, productivity likely exhibits persistence because firms don't tend to change their exact methods of production by a significant on an annual basis, and as a result, productivity in one year will be dependent on productivity in the previous year.

The previous model also assumed that productivity growth was constant across all firms; as a result, the estimates of productivity growth in that model are best viewed as the average productivity growth in the industry. It is more likely that each firm follows its own trend in productivity due to differences in how firms operate. This model relaxes the common trend assumption and allows each firm to have its own productivity trend.

I maintain all of the other assumptions of the model, which is written below.

[16] For some examples of how time-varying parameters have been used in a variety of contexts, see Leybourne (1993), Mazzocchi (2003), and Del Negro and Otrok (2008).

$$
\begin{aligned}
q_{it} &= \alpha_{it} + x_{it}\beta + \varphi_{it}\theta - \delta_i + \varepsilon_{it} \\
\alpha_{it} &= \alpha_{it-1} + \tau_i + \eta_{it} \quad ; \quad t > 0 \\
\alpha_{i0} &\sim N(\mu_\alpha, \sigma_\alpha) \\
\varepsilon_{it} &\sim N(0, \sigma_\varepsilon) \\
\delta_i &\sim N^+(0, \sigma_\delta) \\
\eta_{it} &\sim N(0, \sigma_\eta)
\end{aligned}
$$

Given assumptions of the prior distributions of each parameter, which can be found in Section 9.1.2 of the Appendix, this model can be estimated using Gibbs sampling. Exact evaluation of the likelihood is complicated by the random-walk process in productivity, but is made possible via the Kalman filter. A review of the methodology for using the Kalman filter to estimate standard regression models with time-varying parameters in a Bayesian framework is given in Sarris (1973). While stochastic frontier models and time-varying parameter models have been estimated, I am not aware of any published research that combines the two to examine dynamic changes in productivity.

Random Walk in Productivity and Technology

This model presents an additional extension of the previous model. I maintain the assumption that productivity follows a random walk with drift but relax the assumption that technology remains constant over the time frame of the sample. There are a couple of perspectives that justify relaxing this assumption. First, firms are constantly striving to reduce costs and make innovations to their production technology to further that goal. As discussed in Section 2, firms have invested large amounts in improving improving their networks and pursuing innovation. Ignoring these innovations would lead to a bias in productivity, as discussed in Section 3.1.

A second line of reasoning refers back to the original definition of productivity: The marginal amount of output that can be produced using an additional unit of real resources. As the production technology changes, the combination of inputs that constitutes one unit of real resources will change; not only will the amount of output that can be produced with one unit of real inputs change, but firms will alter the composition of inputs they use as factor productivities change at differing rates. Assuming that technology remains constant over time, an increase in real expenditures will increase all inputs by a constant fixed amount, which will increase output by a fixed amount, after controlling for productivity.

Instead of assuming a technology that is constant across firms and time, I assume that all firms share the production technology (up to their multiplicative productivity) in a given year, but that technology is allowed to change over time. The primary estimating equation then becomes

$$q_{it} = \alpha_{it} + x_{it}\beta_t + \varphi_{it}\theta - \delta_i + \varepsilon_{it}. \tag{3.17}$$

The data prevent the separate identification of β_t for each year; instead, I propose that β_t follows a random walk with drift. Once again, I expect that β_t will exhibit persistence because new technology tends to adapt existing technology. The other assumptions of the model remain the same, which can be expressed in the following relations.

$$\begin{aligned}
q_{it} &= \alpha_{it} + x_{it}\beta_t + \varphi_{it}\theta - \delta_i + \varepsilon_{it} \\
\alpha_{it} &= \alpha_{it-1} + \tau_i + \eta_{it} \quad ; \quad t > 0 \\
\beta_t &= \beta_{t-1} + \rho + \psi_t \quad ; \quad t > 0 \\
\alpha_{i0} &\sim N(\mu_\alpha, \sigma_\alpha) \\
\beta_0 &\sim N(\mu_\beta, \Sigma_\beta) \\
\varepsilon_{it} &\sim N(0, \sigma_\varepsilon) \\
\delta_i &\sim N^+(0, \sigma_\delta) \\
\eta_{it} &\sim N(0, \sigma_\eta) \\
\psi_t &\sim N(0, \Sigma_\psi)
\end{aligned}$$

Once again, due to the random-walk process in productivity and technology parameters, exact evaluation of the likelihood function is difficult but is possible by using the Kalman filter; the general estimation procedure in a Bayesian context is described in Sarris (1973). Once prior distributions are assigned to each parameter, the model can be estimated using Gibbs sampling. A full description of the model, including prior assumptions, can be found in Section 9.1.3 in the Appendix.

Results

This section presents results for each of the three models detailed in Section 6. Each of these models was estimated using Gibbs sampling, a Bayesian estimation technique; unlike classical statistical methods which produce a point estimate for each parameter, Bayesian methods like Gibbs sampling produce a distribution for each parameter that is dependent on prior assumptions, the data, and the structure of the model. Consequentially, I present statistics describing the distribution of each parameter. The distributions of parameters, especially of productivity, have relatively high variance; since many of these parameters are then exponentiated to get their economically-intuitive value, their distributions show significant skew. As a result, I present only estimated medians for each parameter as these will give a better view of the central tendency of these distributions. Parameter estimates for each model estimated are presented in Table 2.

Deterministic Trend in Productivity, Constant Technology

Median estimates from the model with a deterministic trend in productivity and static technology are presented in the Baseline column of Table 2, and median estimates of annual productivity for each firm are plotted in Figure 4. While this is the most basic model presented in this paper, it provides some general insight into productivity and growth in the industry. First, I estimate that average productivity growth was modest over the period from 1999 to 2014, with median estimates of 1.2% growth per year on average across all firms. There is also heterogeneity in productivity across firms; CN, CSX, NS, and UP show the highest levels of productivity, between 1.65 and 1.708 in 2014, while KCS has the lowest productivity at 1.52 in 2014. Mean productivity was estimated to be 1.661 in 2014. Estimates of efficiency range from 78.2% to 86.4%, with a mean of 81.4% across firms.

As discussed in Section 4, the increase in output that results from a input substitution and increased input usage is given by $F(X_{it})/F(X_{i0})$, which is represented by $\exp((X_{it} - X_{i0})\beta)$ in the empirical model. A plot of the proportional increase in output due to changing input quantities is given in Figure 5. CN, KCS, and the Soo Line all saw increases in input use over the sample, CSX, NS, and UP all decreased input use, and BNSF saw little change in output due to changing inputs.

Using a Bayesian estimation framework permits direct evaluation of the probability that a parameter of interest, like productivity growth, lies within a given range.

TABLE 9. Posterior Medians and Median Absolute Deviations

	Baseline	Constant Technology	Changing Technology
Productivity in 2014 $(\exp(\alpha))$			
BNSF	1.602 (2.345)	2.594 (0.622)	1.026 (1.513)
CN	1.708 (2.5)	4.962 (1.557)	1.004 (1.487)
CSX	1.682 (2.463)	4.077 (1.045)	0.979 (1.451)
KCS	1.52 (2.224)	2.95 (0.809)	0.873 (1.295)
NS	1.675 (2.453)	3.831 (1.058)	1.116 (1.655)
SOO	1.534 (2.245)	3.296 (1.023)	1.074 (1.592)
UP	1.65 (2.417)	3.82 (0.98)	0.891 (1.321)
Efficiency $(\exp(-\delta))$			
BNSF	0.814 (0.187)	0.938 (0.056)	0.612 (0.357)
CN	0.864 (0.138)	0.934 (0.043)	0.611 (0.363)
CSX	0.851 (0.16)	0.949 (0.04)	0.611 (0.351)
KCS	0.814 (0.191)	0.925 (0.061)	0.63 (0.357)
NS	0.864 (0.147)	0.954 (0.043)	0.618 (0.356)
SOO	0.782 (0.213)	0.943 (0.054)	0.623 (0.35)
UP	0.85 (0.159)	0.919 (0.064)	0.62 (0.356)
Productivity Trend (τ)	0.012 (0.005)		
BNSF		0.002 (0.008)	-0.01 (0.582)
CN		0.037 (0.016)	0.008 (0.669)
CSX		0.028 (0.009)	0.001 (0.432)
KCS		0.007 (0.007)	-0.008 (0.646)
NS		0.021 (0.007)	-0.003 (0.455)
SOO		0.018 (0.008)	-0.01 (0.58)
UP		0.021 (0.008)	-0.007 (0.597)
Input Parameters			
Locomotives	-0.249 (1.8)	-2.522 (1.216)	0.161 (7.774)
Cars	-0.728 (1.894)	-1.076 (1.57)	0.245 (7.647)
Road	1.965 (1.82)	0.216 (1.62)	0.383 (7.359)
Fuel	1.354 (1.795)	2.928 (1.274)	0.474 (6.53)
Labor	-0.916 (2.088)	1.185 (2.208)	0.476 (6.963)
$(\text{Locomotives})^2$	-0.21 (0.144)	-0.46 (0.126)	0.049 (3.475)
$(\text{Cars})^2$	0.586 (0.145)	0.341 (0.303)	-0.173 (4.978)
$(\text{Road})^2$	-0.173 (0.114)	-0.206 (0.12)	0.22 (3.228)
$(\text{Fuel})^2$	0.165 (0.127)	0.034 (0.153)	0.072 (3.379)
$(\text{Labor})^2$	0.727 (0.401)	-0.287 (0.259)	0.102 (4.505)
(Locomotives):(Cars)	-0.009 (0.207)	-0.009 (0.206)	-0.008 (5.846)
(Locomotives):(Road)	-0.147 (0.183)	-0.486 (0.21)	0.123 (4.963)
(Locomotives):(Fuel)	0.199 (0.204)	0.792 (0.193)	0.059 (4.82)
(Locomotives):(Labor)	0.065 (0.332)	-0.022 (0.399)	-0.277 (5.417)
(Cars):(Road)	0.082 (0.2)	0.051 (0.179)	0.126 (5.262)
(Cars):(Fuel)	0.285 (0.285)	-0.811 (0.216)	-0.016 (5.669)
(Cars):(Labor)	-1.049 (0.534)	0.61 (0.386)	0.157 (6.147)
(Road):(Fuel)	0.076 (0.196)	0.137 (0.217)	-0.277 (4.45)
(Road):(Labor)	0.105 (0.367)	0.348 (0.332)	-0.219 (5.076)
(Fuel):(Labor)	-0.765 (0.264)	-0.121 (0.21)	0.068 (5.884)
Network Characteristics			
Avg. Length of Haul	1.592 (1.972)	-0.99 (2.241)	0.047 (2.939)
Avg. Speed	0.685 (1.982)	-1.438 (1.194)	0.035 (2.926)
Miles of Road	-0.788 (1.585)	-0.541 (1.344)	0.115 (2.88)
% Unit	-3.558 (1.307)	-2.04 (1.077)	-0.046 (2.891)
% Bulk	2.178 (1.387)	1.544 (1.049)	0.019 (2.983)
$(\text{Avg. Length of Haul})^2$	-0.287 (0.205)	0.027 (0.164)	0.015 (1.66)
$(\text{Avg. Speed})^2$	-0.447 (0.316)	-0.187 (0.21)	0.022 (2.497)
$(\text{Miles of Road})^2$	0.061 (0.083)	0.059 (0.066)	-0.014 (0.899)
$(\%\ \text{Unit})^2$	-0.069 (0.122)	0.004 (0.089)	-0.053 (2.268)
$(\%\ \text{Bulk})^2$	-0.061 (0.147)	0.014 (0.108)	-0.095 (2.618)
(Avg. Length of Haul):(Avg. Speed)	0.675 (0.447)	0.51 (0.277)	-0.002 (2.535)
(Avg. Length of Haul):(Miles of Road)	0.136 (0.167)	0.049 (0.119)	0.06 (2.103)
(Avg. Length of Haul):(% Unit)	0.776 (0.278)	0.508 (0.219)	0.021 (2.273)
(Avg. Length of Haul):(% Bulk)	-0.189 (0.215)	-0.267 (0.141)	-0.023 (2.351)
(Avg. Speed):(Miles of Road)	-0.338 (0.177)	-0.197 (0.09)	0.012 (1.908)
(Avg. Speed):(% Unit)	-0.113 (0.33)	-0.076 (0.148)	-0.048 (2.424)
(Avg. Speed):(% Bulk)	-0.462 (0.344)	-0.479 (0.152)	0.034 (2.672)
(Miles of Road):(% Unit)	-0.095 (0.1)	-0.047 (0.078)	0.035 (1.51)
(Miles of Road):(% Bulk)	0.026 (0.097)	0.178 (0.06)	-0.054 (1.654)
(% Unit):(% Bulk)	-0.014 (0.17)	0.12 (0.12)	0.012 (2.641)

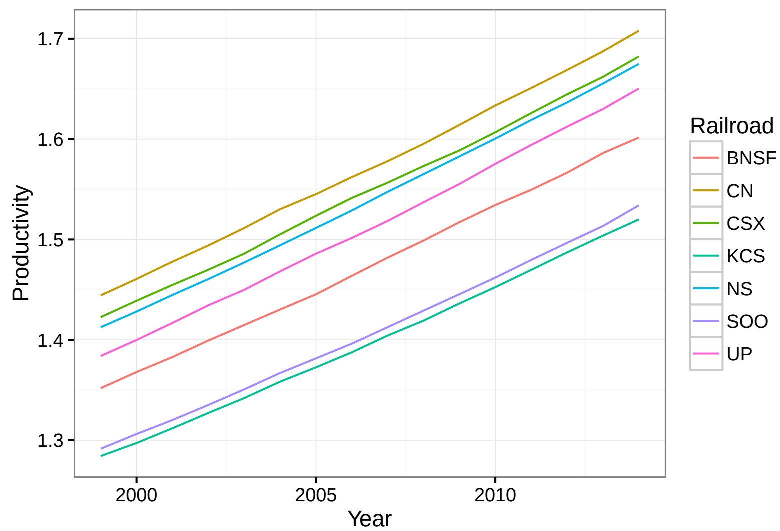

FIGURE 10. Railroad Productivity as a Deterministic Trend

Specifically, the probability that a parameter ϖ lies within a set S is

$$\Pr(\varpi \in S) = \int I(\varpi \in S) p(\varpi|D) d\varpi, \tag{3.19}$$

where $I(\cdot)$ is an indicator function and $p(\varpi|D)$ is the posterior distribution of ϖ conditional on the data D. I estimate the probability that average productivity increased (i.e., $\Pr(\tau > 1)$) is 99.314%.

In light of these results, one can safely conclude that there has been growth in average productivity since total consolidation in 1999. This is reassuring, the flexibility granted to firms by partial deregulation seems to have set the groundwork for continued long-term growth and sustained viability. However, while the industry appears to be growing on average, the performance of individual firms is not clear. To investigate how each railroad has progressed, I turn to my second model.

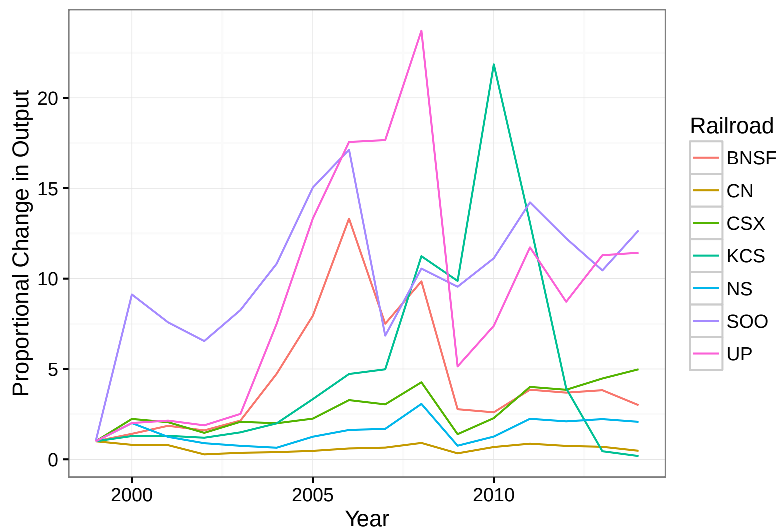

FIGURE 11. Increase in Output From Change in Inputs

Random Walk in Productivity, Constant Technology

Median estimates from the random walk in productivity model are presented in the Constant Technology column in Table 2. There are several modest differences from the previous model's results. First, estimates of productivity and efficiency are lower for this model; mean productivity is estimated to be 3.684 in 2014 and mean efficiency was estimated to be 92.9%. I estimate that *effective productivity* (i.e., the product of productivity and efficiency) was 3.385 in 2014. Trends in productivity show marked heterogeneity across firms, and mean productivity growth is estimated to be 1.96% per year, similar to the 1.2% growth estimated by the previous model.

Productivity for each firm over time is presented in Figures 6 and 7 in two ways. Figure 6 shows the expectation of firm productivity conditional on firm trends and information in the year 1999; this is identical to the deterministic part of productivity. Figure 7 shows estimated productivity, which includes both the trend as well as the random walk in productivity. The inclusion of the random walk is important because it

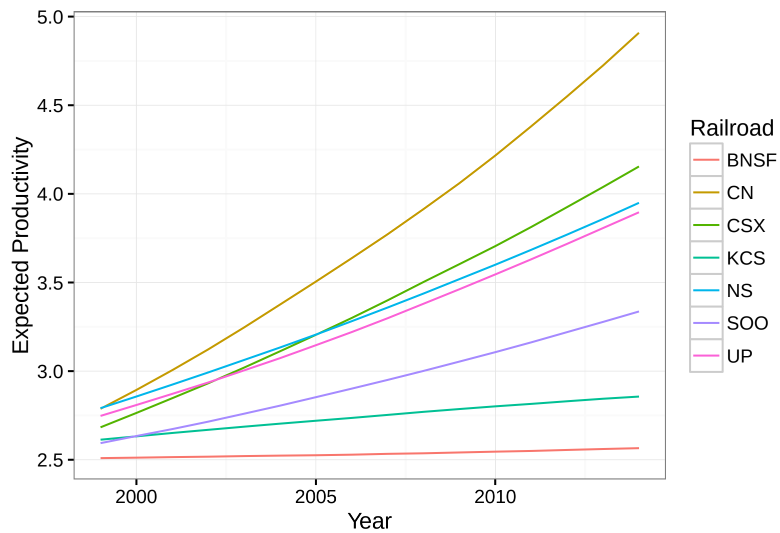

FIGURE 12. Expected Railroad Productivity as a Random Walk With Drift

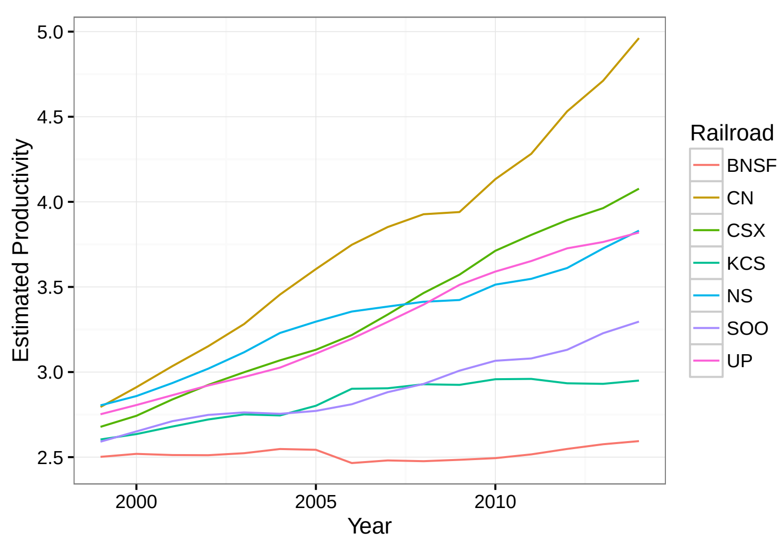

FIGURE 13. Estimated Railroad Productivity as a Random Walk With Drift

reflects actual year-to-year variation in productivity that cannot be captured by a simple deterministic trend. There are many factors that could potentially affect productivity, and it would be neither technically feasible nor even possible given data restrictions to include them all into the model. By assuming productivity follows a random walk with trend, variation in productivity can be captured flexibly.[17]

One can quickly see that actual productivity differs from its expected value. As an example, NS was expected to have productivity growth of 2.1% per year over the course of the sample; in actuality, NS received two negative shocks to productivity in 2008 and 2009, which resulted in lower estimated productivity growth of 1.889% per year. This indicates that there are other factors influencing productivity that cannot be described by changes in input use and a simple trend. As explained previously, it can be difficult to attribute an exact cause to these fluctuations.

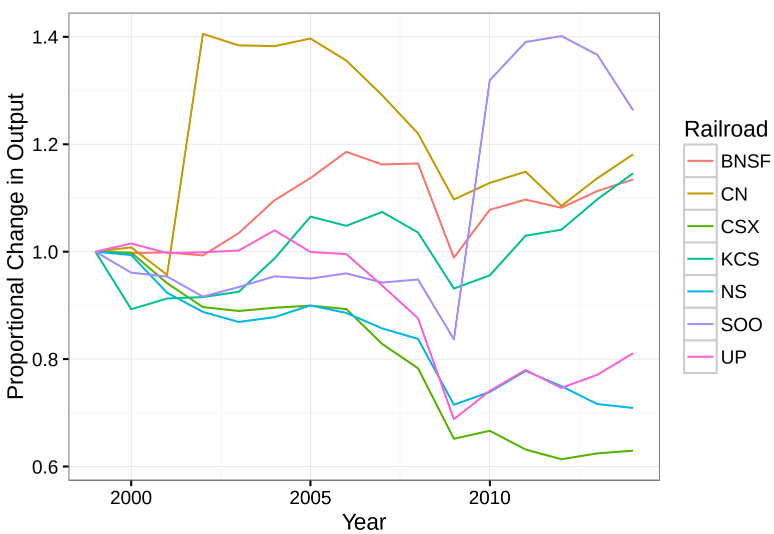

FIGURE 14. Increase in Output From Change in Inputs

[17]Of course, this assumption comes at a cost. While variation in productivity can be identified, the exact source of this variation cannot. As a result, one can only use institutional knowledge to posit why firms see fluctuations in productivity; in order to empirically identify what factors drive changes in productivity, those factors must be explicitly included in the model.

Estimates of the effect of changing input use on output are shown in Figure 8. Similar to the previous model, I find that BNSF, CN, KCS, and the Soo Line all increased the quantity of inputs used while CSX, NS, and UP saw decreases in input quantities. As noted earlier, CN and Soo Line also saw significant increases in productivity, indicating both matched increased demand with a combination of neutral factors and increased inputs.

I again estimate the probability that firms experience positive average growth in productivity as well as the median value of annualized average growth between 1999 and 2014, and the results are given in Table 3. Estimates suggest that each firm saw increases in productivity, with median estimates between 0.235% to 3.551% per annum, and probability of productivity growth between 65.7% and 100%. Further, I estimate the probability that *all* firms experience positive growth in productivity is 63.424%.

TABLE 10. Average Productivity Growth

Firm	Annual Productivity Growth	Probability of Increase
BNSF	0.235% (0.006)	65.7%
CN	3.551% (0.012)	100%
CSX	2.474% (0.007)	99.908%
KCS	0.816% (0.006)	94.102%
NS	1.889% (0.005)	99.992%
SOO	1.657% (0.006)	99.926%
UP	1.869% (0.006)	99.996%

These results offer optimistic outcomes for some firms and a more modest outlook for others. Similar to Tretheway et al. (1997), which found that CN showed high productivity growth through 1991, I estimate that CN has the highest productivity in 2014 as well as the highest growth rate over the sample. As noted by previous research, CN continued to operate on several less-profitable lines and had yet to fully take advantage of economies of density through the 1990s; as CN continued to improve on those frontiers, its productivity increased.

On the other hand, the remainder of the railroads exhibit modest growth in productivity. Some of these firms may no longer find it feasible to abandon lines and make improvements to economies of density and instead must turn towards innovating their production technology to realize higher growth. This model ignores the possibility of technological change; to investigate whether firms have been able to increase productivity through innovation, I turn to my third and final model.

Random Walk in Productivity and Technology

Median estimates from the changing technology model are presented in the Changing Technology column in Table 2, and median estimated productivity for each firm is plotted in Figure 9. As explained in Section 4, there are a few key values of interest in the analysis of productivity in the light of changing technology. First, the productivity growth estimated by this model does not include changes in productivity due to innovation but instead reflects growth from factors other than innovation. Both CN and CSX show growth in productivity due to non-innovative factors at 0.702% and 0.016% per year, respectively, indicating those firms may still be able to abandon lines and improve density to increase productivity. All other firms experience stagnating or even declining productivity due to non-innovative factors at rates between 0.283% and 0.901% per year, demonstrating that those other methods of increasing productivity are no longer feasible for these railroads.

The growth in productivity due to technological change relative to the base year is given by $X_{i0}(\beta_t - \beta_0)$.[18] A plot of the estimated change in productivity due to innovations for each railroad is presented in Figure 10. Each railroad experienced positive average growth in productivity due to innovations over the course of the sample. BNSF, KCS,

[18]Productivity growth due to innovation was given as $F_1(X_0)/F_0(X_0)$. Note that this is an analogous measure because $\ln(F_1(X_0)/F_0(X_0)) = f_1(X_0) - f_0(X_0) = X_0\beta_1 - X_0\beta_0 = X_0(\beta_1 - \beta_0)$.

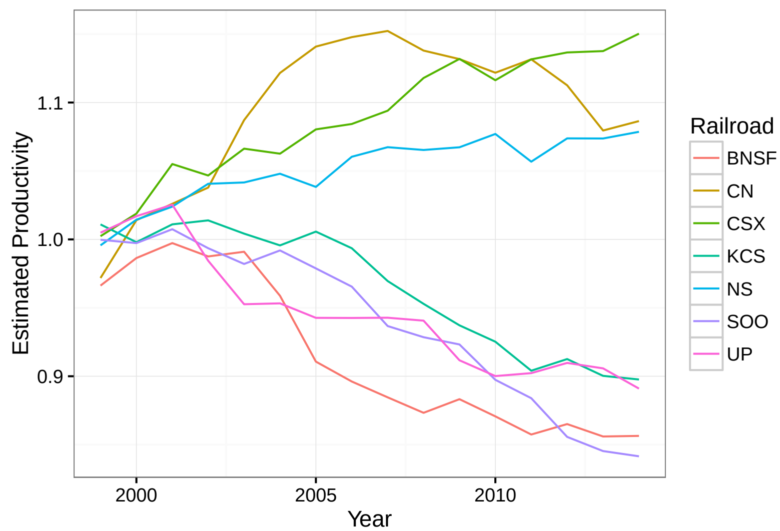

FIGURE 15. Estimated Railroad Productivity Accounting For Changing Technology

Soo Line, and UP saw the largest productivity gains due to chnages in technology. On the other hand, CSX, CN, and NS found more modest increases, indicating those firms have relied more heavily on other methods to increase their productivity. While I have a small sample of firms, I find that the large firms (i.e., BNSF and UP) are benefactors of changing technology; this correlates with the finding of Rose and Joskow (1990) where large firms are more likely to adopt innovations early due to risk preferences. A plot of technology-inclusive productivity, the sum of productivity due to innovation and that due to non-innovative factors, is presented in Figure 11. Overall, BNSF, CN, and Soo Line have shown the highest growth in total productivity. KCS and UP experienced modest gains in total productivity over the course of the sample, and all CSX and NS experienced relatively little productivity growth.

Median estimates of average annual technology-inclusive productivity growth and the probability that each firm experienced positive growth in productivity are given in Table 4. Productivity growth estimates are much more modest than in previous models; CN saw the largest expected increase at 3.065% per annum, while NS experienced small decreases

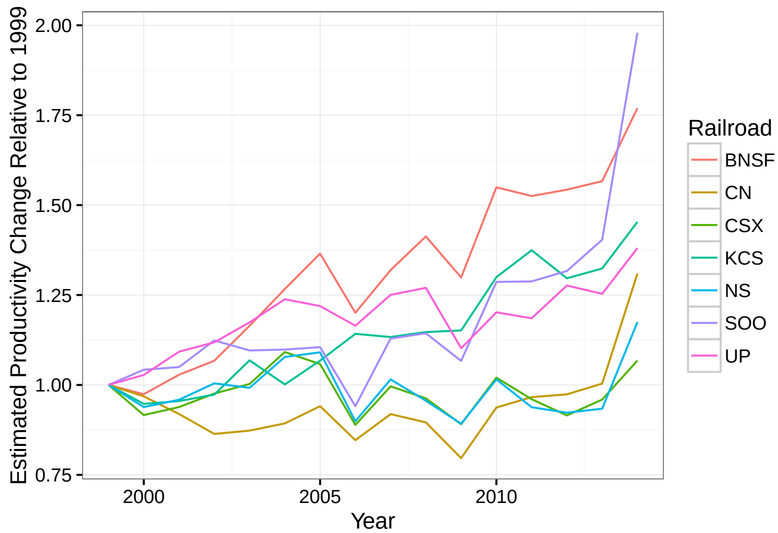

FIGURE 16. Estimated Change in Productivity Due to Innovations

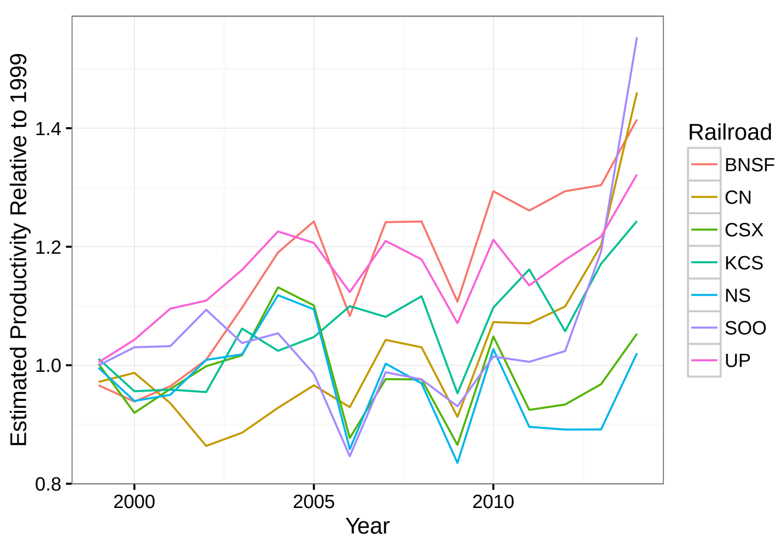

FIGURE 17. Estimated Productivity Including Effects of Innovation

in productivity of 0.007% annually. I estimate the probability that *all* firms experienced an increase in productivity between 1999 and 2014 is only 4.688%, indicating it is very likely that at least one firm saw a decrease in productivity over the sample period.

TABLE 11. Average Productivity Growth

Firm	Annual Productivity Growth	Probability of Increase
BNSF	2.272% (0.367)	52.324%
CN	3.065% (0.666)	51.792%
CSX	0.037% (0.414)	50.036%
KCS	1.248% (0.468)	51.156%
NS	-0.007% (0.395)	49.992%
SOO	2.79% (0.389)	53.064%
UP	1.623% (0.316)	52.028%

The change in output due to input substitution and increased input usage is shown Figure 12. Most firms didn't increased or even decreased the amount of inputs they used over the sample, with the exception of CN. While CN saw decreases in productivity due to technological change, especially before 2010, it dramatically increased its input use during that time. This provides evidence that CN relied on increasing its input use rather than increasing productivity through innovation, especially before 2010.

The preceding analysis examines how innovations affect the productivity of railroads' original plan of production. To determine whether firms made changes that took advantage of changing technology, namely by allocating more resources towards more productive factors, one can examine how changing technology affects new production plans. In Section 4, I described the comparison of two measures: Technology's benefit to the original production plan, $F_t(X_{t-1})/F_{t-1}(X_{t-1})$, and the benefit of innovation to the new production plan, $F_t(X_t)/F_{t-1}(X_t)$. Since $F_t(X_s)$ is expressible as $\exp(X_s\beta_t)$ in this model, these two ratios can be calculated as $\exp(X_{t-1}\beta_t - X_{t-1}\beta_{t-1})$ and $\exp(X_t\beta_t - X_t\beta_{t-1})$, respectively. Recall that if the latter is greater than the former, then the firm allocated resources towards factors that innovation made more productive; otherwise,

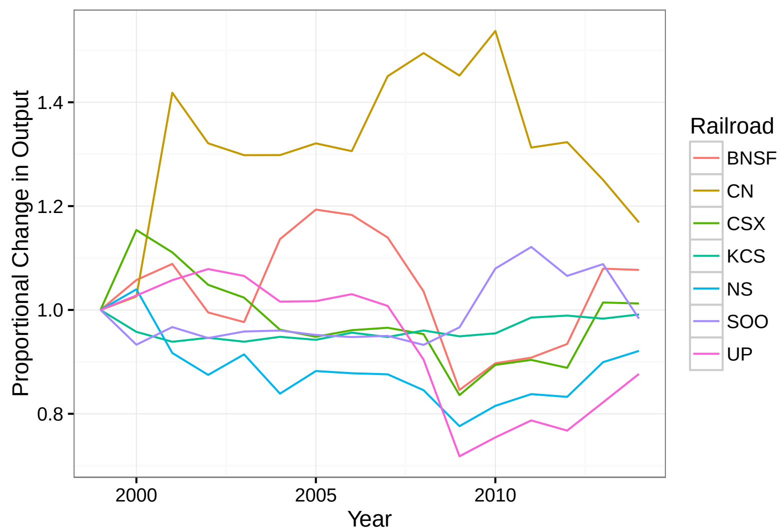

FIGURE 18. Increase in Output From Change in Inputs

it made changes that clashed with technology change and must have found it cheaper to increase productivity through means other than input substitution. To exhibit these results, I calculate distributions of both values for each firm and year. Bayesian estimation methods make it possible to evaluate the probability that one of these values exceeds the other; I plot the probability that each firm substitutes towards more productive inputs[19] in each year in Figure 10. Every railroad has near 50% probability of substituting towards more productive inputs and no railroad appears to have consistent behavior. This shows that firms are largely not able to anticipate changes in the relative productivity of inputs and make allocative changes in response.

Bayesian Model Selection

An advantage of using a Bayesian estimation framework is that it allows for the direct computation of model probabilities, which can be used for model selection and

[19]That is, I calculate the probability that $F_t(X_t)/F_{t-1}(X_t)$ is greater than $F_t(X_{t-1})/F_{t-1}(X_{t-1})$.

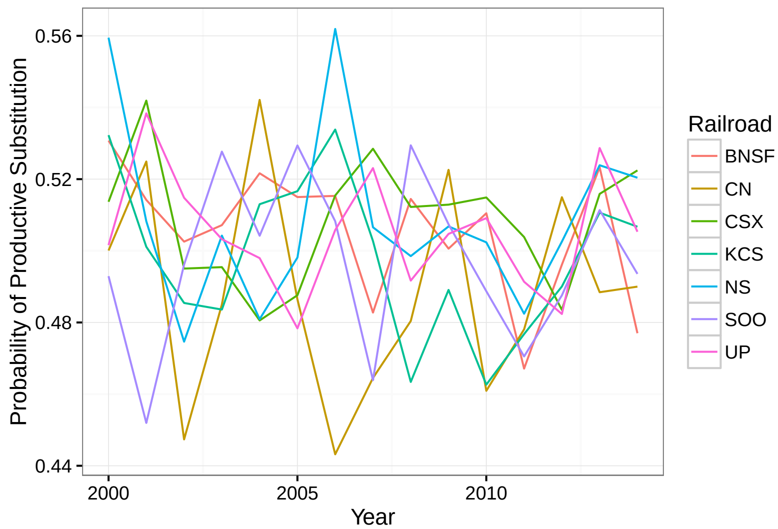

FIGURE 19. Probability of Substituting Towards More Productive Input

averaging of results. Letting each of the previously discussed models be M_1, M_2, and M_3, respectively, the probability of model M_i being the correct model is

$$\Pr(M_k|D) = \frac{\Pr(D|M_k)\Pr(M_k)}{\Pr(D)} = \frac{\Pr(D|M_k)\Pr(M_k)}{\sum_j \Pr(D|M_j)\Pr(M_j)}, \tag{3.20}$$

where $\Pr(D|M_k)$ is the marginal likelihood of the data D for model k and $\Pr(M_k)$ is the prior probability of model k, chosen by the researcher. Direct evaluation of the marginal likelihood is difficult in general, but can be computed using the methods described in Chib and Jeliazkov (2001).

I assume a uniform prior probability over the above three models, and posterior model probabilities are given in Table 3. The model that allows both productivity and technology to follow a random walk with drift has the highest probability of being the true model. The effects of changing technology have an important effect on productivity changes, as can be seen from the relatively low probability of the model that only allows a random walk in productivity and not in technological parameters.

TABLE 12. Posterior Model Probabilities

Model	Prior Probability	Posterior Probability
Deterministic trend	1/3	0.24311
Random walk in productivity	1/3	0.08665
Random walk in productivity and technology	1/3	0.67024

TABLE 13. Average Productivity Growth

Firm	Annual Productivity Growth	Probability of Increase
BNSF	1.817% (0.247)	64.907%
CN	2.636% (0.449)	67.522%
CSX	0.513% (0.279)	66.337%
KCS	1.181% (0.315)	66.585%
NS	0.433% (0.266)	66.315%
SOO	2.287% (0.262)	68.368%
UP	1.524% (0.213)	67.68%

Finally, I calculate median productivity for each firm over time, median productivity growth over the course of the sample, and the probability that firms experienced increases in productivity between 1999 and 2014 using Bayesian model averaging, which calculates parameters of interest for each model and weights them by their respective model probabilities. Specifically, for a statistic of interest ϖ, the average of ϖ over the models is

$$E[\varpi|D] = \sum_k \varpi_k \Pr(M_k|D),$$

where ϖ_k is the estimated value of ϖ in model k. Estimated productivity growth and the probability that each firm experienced growth in productivity over the course of the sample are given in Table 4, and figure 11 plots productivity averaged over the models for each firm over time.

Model average results show that CN experienced an average annual productivity growth of 2.636% per year between 1999 and 2014, the greatest of all firms in the sample; CSX and NS showed the least growth, at 0.513% and 0.433%, respectively, while the

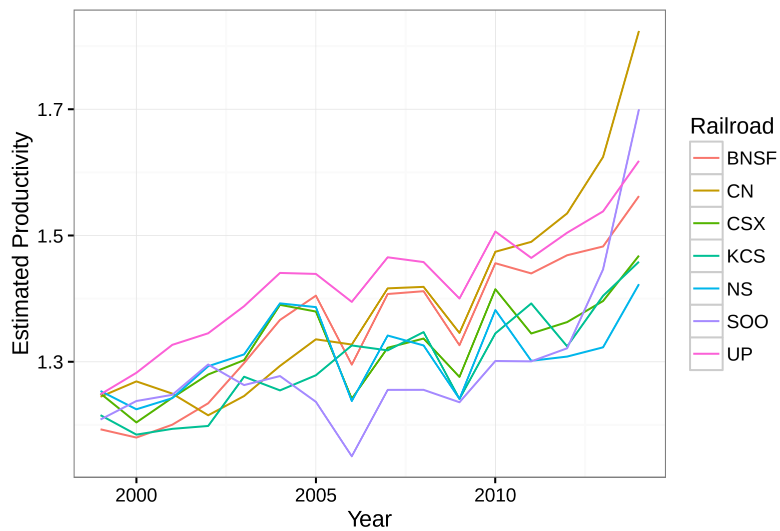

FIGURE 20. Model Average of Productivity Over Time

remainder of the firms saw productivity growth between 1.181% and 2.287% annually. Each firm had a relatively high probability of experiencing positive growth, between 64.907% and 68.368%. Finally, I calculate that the probability that *all* firms experienced positive growth 32.782%, meaning it is more likely than not that at least one firm saw a decrease in productivity between 1999 and 2014.

Conclusion

The level and growth of productivity offer important insight into the functioning of an industry, from its economies of scale to important factors of growth and even its long-term viability. The railroad industry has changed dramatically since partial deregulation 1980, which occurred largely because of worries about regulation and its effect on productivity growth. Many studies have examined the growth in productivity immediately following the industry's partial deregulation, but few have looked at how productivity has changed since the industry became more stable in 1999. Following the massive changes that occurred through the 1980s and early 1990s, it is unlikely that firms will be able to continue pursuing broad changes like line abandonment to increase productivity. Instead, they need to turn towards improving technology and substituting inputs towards more productive factors to increase their productivity.

Unfortunately, existing models of productivity fail to account for technological change. In this paper, I develop a model that flexibly accounts for changes in productivity and technology and use it to decompose changes in productivity into those caused by innovation and those caused by other factors. To my knowledge, no published research exists that separately identifies productivity growth due to technological change and that due to broad non-innovative changes. Further, this model allows productivity and technology to evolve flexibly over time and can produce estimates of the level of productivity and its growth, which can inform key values in regulation. Finally, I discover a metric that determines whether firms have allocated additional resources towards factors that innovation makes more productive in order to realize further productivity gains. I apply my model to the railroad industry to investigate the recent change in productivity and to determine whether it is being driven by technological change or factors other than innovation.

I find that each Class I railroad has likely seen productivity growth since 1999, but the driving forces behind this growth differ. BNSF, KCS, Soo Line, and UP have seen large increases in its productivity due to technological change, by as much as 60% between 1999 and 2014. On the other hand, CN, KCS, and NS saw much slower growth induced by changing technology. Instead, these railroads relied on other methods to increase productivity, such as continuing to abandon unprofitable lines. The probability that firms substituted inputs towards factors that technological change makes more productive is about 50%, with no discernible pattern across firms or over time, indicating that railroads are not able to anticipate innovations or simply aren't adjusting inputs to take advantage of technological change. Finally, I perform Bayesian model selection and find that the model that allows for flexibility in both productivity and technology has the highest probability of being the true model. Using Bayesian model averaging, I find that each firm experienced modest growth in productivity between 1999 and 2014, with median estimates ranging from 0.433% to 2.636% per year. This chapter investigates productivity broadly, so to more clearly identify inefficiency in production and its relationship to regulation and competition, I turn to my fourth chapter, "Competitive Pressures and Inefficiency in Allocation."

CHAPTER IV

COMPETITIVE PRESSURES AND INEFFICIENCY IN ALLOCATION

Abstract

There is a wealth of literature that points to inefficiencies in production. Inefficiencies can arise in the production of outputs from overutilization of inputs in the production process (technical inefficiency) or from errors in optimization that misalign factor prices and optimal input decisions (allocative inefficiency). In examinations of inefficiency, many studies use an inflexible production technology that fails to account for differences in the technology of firms, which has the potential to bias estimates of allocative inefficiency. In this study, I develop a model that flexibly accounts for differences in the production process across firms. I use the model to derive optimal input quantities for each firm and compare them to observed quantities to obtain estimates of allocative inefficiency. Using Bayesian model selection, I find that incorporating flexibility in production is appropriate and necessary for obtaining unbiased estimates of allocative inefficiency. Next, I find that firms generally overcapitalize in the rate-regulated rail industry, providing evidence of the Averch-Johnson effect. Finally, I allow allocative errors to be correlated with variables describing competitive pressures and find that greater market power is associated with less allocative inefficiency, providing evidence against X-inefficiencies.

Introduction

Beginning with the seminal work of Leibenstein (1966), researchers have recognized that the inefficiency of firms can emanate from a lack of competitive pressure that allows firms to depart from normal activities to reduce costs. There is an abundance of research investigating firms' inefficiency in production, which can arise both from technical inefficiencies in the application of inputs and allocative inefficiencies where there are errors in the optimal allocation of inputs. While technical inefficiencies emerge from the production process, allocative inefficiencies are caused by mistakes in the decision-making process of firms. There are many reasons why inefficiency in allocation exists, from imperfect observation of factor prices to stickiness in the amounts of inputs used, or even that firms may find reducing allocative errors more costly than the errors themselves. In this paper I reexamine these issues by developing and estimating a model that encompasses both forms of inefficiencies and also the effect of competition on inefficiency. Unlike much of the previous research, I develop a model that describes production flexibly, accounts for differences in the production technology across firms, and evaluates the role of competitive pressures.

Allocative inefficiency has been studied extensively, but I introduce a model that offers some important extensions. I remain agnostic of the exact causes of allocative inefficiency, but I focus on measuring allocative errors and determining their relationship with firm characteristics. First, existing models that estimate allocative inefficiency fail to account for differences in the production process between firms. As noted by Tsionas (2002), accounting for differences in production is necessary to obtain unbiased estimates of technical inefficiency, but is presumably also crucial to estimate allocative inefficiency. As an example, if some firm uses a given input more productively than another, then it should use relatively more of that input to maximize profit. Thus, assuming that all firms share the production technology can lead to biases in estimates of allocative errors. Additionally, I allow total factor productivity to follow a random-walk with drift, which is a flexible form that is able to capture idiosyncratic shocks in productivity that could otherwise lead to bias in estimates of inefficiency. Using this extended framework, I can test whether allowing firm production to vary is appropriate using Bayesian model selection. I also look for evidence of overcapitalization, which would support the Averch-Johnson hypothesis that firms in rate-regulated industries over-invest in capital so that total profit allowed by regulation is greater.

Finally, firms may face different incentives to reduce allocative inefficiencies. If some firm is able to attract higher quality inputs, the cost of a mistake in allocation will be higher, providing a greater incentive to minimize such errors. It is also possible that so called "X-inefficiencies" contribute to allocative inefficiencies. As first noted in Leibenstein (1966), firms with large market power may not face sufficient incentives to minimize costs by, for example, reducing errors in the allocation of inputs. I introduce a model that allows allocative inefficiencies to be correlated with competitive pressures to test for the effects of input quality and X-inefficiencies on the allocation process.

I apply my model to the railroad industry. The industry is ideal for this analysis because of its history of regulation and the geographically distinct nature of many firms. First, partial deregulation of the industry dramatically reduced costs and improved firm viability but also resulted in the consolidation of the industry into just seven firms, causing concerns of excess market power. Second, there is little geographic overlap across railroad networks, which only exacerbates the problem of market power but also means firms can face different input supply markets and therefore different input quality. While there has been research examining allocative inefficiency in the railroad industry (most notably by Kumbhakar (1988)), I am not aware of any published research that has used a flexible form for the production function, allowed for variability in production across firms, or investigated the relationship between allocation errors and firm characteristics.

I first find that controlling for differences in the production technology over firms is appropriate, and failure to do so will result in biased estimates of allocative inefficiency. Next, I find strong evidence of overcapitalization in the rail industry and no significant evidence of undercapitalization. Since the industry is rate-regulated, this finding suggests the Averch-Johnson effect may hold for railroads. Finally, I find that as measures of market power increase, allocative errors decrease or don't change at all. This finding appears to refute the existence of X-inefficiencies in the rail industry since we would expect to see allocative errors increase with an increase in market power if X-inefficiencies were present.

I begin by reviewing the history of regulation, productivity, and inefficiency in the railroad industry. I continue by reviewing the literature relevant to the estimation of allocative inefficiency and its study in the railroad industry. I then develop my theoretical model of production. The data sources are described, and empirical models are presented. I then review and interpret results and follow with a conclusion.

Institutional Background

The railroad industry has faced some form of federal regulation since the passage of the Interstate Commerce Act (ICA) of 1887. While there was significant competition between railroads on a large scale before regulation, firms tended to operate more as monopolies on a local scale (Brown, 2013). Railroads readily engaged in price discrimination based on shipper's access to competing modes and willingness to pay. Firms would also often act as cartels, and many shippers were affected by excessive rates (MacDonald and Cavalluzzo, 1996). The ICA established the Interstate Commerce

Commission, which was granted the ability to set maximum rates, oversee mergers, and ensure availability of services (Hilton, 1966). Regulation caused shippers to see more reasonable rates and encouraged healthy competition in the industry (MacDonald and Cavalluzzo, 1996).

The development and improvement of competing modes of transportation and changes to the types of goods being shipped caused existing regulation to impede efficiency as the 20th century progressed. Not only were new modes of transportation like truck and air able to transport goods faster and serve more customers, but efficiency of barges had improved and plastics, which are much less dense than other bulk materials, began to constitute a greater part of all shipments (Wilson, 1994). As a result, railroad profits fell, and many raised concerns over the inefficiency of regulation and working with a bloated regulatory agency (Lahner, 1975). The industry was partially deregulated with the passage of the 4R Act of 1976 and the Staggers Act of 1980. Deregulation granted firms much more flexibility over the rates they set, allowed for contracts to be negotiated, granted the regulatory agency less oversight over mergers, and allowed railroads to more easily abandon lines (Johnson and Thomas, 1983).

Deregulation certainly resulted in efficiency improvements. First, since railroads were operating many unprofitable routes and had significant excess capacity, there were enormous gains to be realized in economies of density (Keeler, 1974). After deregulation, railroads began abandoning lines to realize benefits of increased density, and the total network controlled by Class I railroads fell from 164,822 miles in 1980 to just 95,391 miles in 2013 (United States Surface Transportation Board, 2015). The abolishment of minimum rate regulation had a more humble effect; railroads were not able to use their greater flexibility in pricing to attract as much traffic from competing modes as previously thought, and the effect of rate deregulation on prices was negligible (Boyer, 1987).

Railroads also aggressively pursued mergers and acquisitions following partial deregulation. The number of Class I railroads fell from 40 in 1980 to 7 by 1999, where it remains today. The merging of firms has two contradictory effects on efficiency. On the one hand, larger firms are able to take advantage of economies of scale, which are prevalent in the railroad industry. On the other hand, mergers reduce the level of competition in the market, potentially putting upward pressure on prices and reducing the incentive to minimize costs. However, between 1986 and 2001, consumer surplus rose by about 30% in U.S. rail freight markets, indicating that the benefits of mergers appear to have outweighed the cost, at least initially (Ivaldi and McCullough, 2005). Overall, cost savings amounted to up to 40% by 1989, though they have leveled off more recently (Wilson, 1997).

The more recent effects of reduced competition on the railroad industry have been less studied. While many have looked towards excessive rates and market power,[1] few have examined efficiency losses. As firms experience lower levels of competition, they not only lose the incentive to keep prices low, but also have less incentive to minimize costs (Leibenstein, 1966). Further, I am not aware of any published research that has examined the effect of excess market power on errors in the allocation of inputs. While precise allocation reduces production costs and may be important for competitive firms, the allocation process is costly and railroads that face lower levels of competition may not have the incentive to invest in accurate allocation.

Literature Review

The inefficiency of firms has been studied in a number of contexts. Inefficiency has been separated into many components including technical and allocative, and various methods have been used to empirically estimate these values. While the inefficiency

[1] For examples of studies that examine railroad pricing after deregulation, see Bitzan and Wilson (2007), McFarland (1987), and MacDonald and Cavalluzzo (1996).

of many industries has been studied, railroads have long been a focal point due to the structure and importance of the industry and its history of regulation. In this section, I first review the relevant literature surrounding the study of inefficiency, its various types, and methods used for its estimation. I then cover studies of inefficiency in the railroad industry, how they relate to regulation, and areas that have been less examined.

Sources of Inefficiency

Researchers have long been interested in measuring inefficiency, defined as the realized deviation from maximum possible output. As the study of inefficiency grew, new sources were discovered and quantified, ranging from technical inefficiencies in production to suboptimal allocation of inputs to absence of incentives to minimize costs when there is a lack of competition. While inefficiency has been discussed, measured, and studied using a variety of methods, most current empirical work relies on stochastic frontier models. In this section I review the study of inefficiency and its various sources and how it has been empirically measured.

Aigner et al. (1977) began the empirical study of inefficiency with the introduction of the stochastic frontier model. In contrast with traditional production models like Solow (1957) that assume any empirical errors are driven by differences in productivity, stochastic frontier models separate errors into differences in productivity, inefficiency, and measurement error by invoking a structural assumption on inefficiency. Since it is defined as deviation from maximum possible output, inefficiency inherently one-sided; that is, inefficiency must be non-negative and cannot exceed 100%. By assuming it follows a one-sided distribution such as log-normal or half-normal, it is possible to separately identify inefficiency from the two-sided measurement error. By further imposing structure on productivity, specifically that it is constant across firms and time, the authors were able to separately estimate productivity, inefficiency, and measurement error.

While standard stochastic frontier models can estimate inefficiency, they are not able to assign a cause to it. There have been many extensions made to the stochastic frontier model to ascertain these causes. Perhaps most notably, Schmidt and Lovell (1979) introduces the concept of allocative inefficiency, which occurs when firms make errors in the allocation of inputs. There could be several causes for these errors. On the one hand, firms may not accurately observe prices, leading them to use a suboptimal bundle of inputs; on the other hand, firms may also incorrectly predict needed input quantities or may face rigidities that make selecting the optimal input allocation difficult or impossible. The authors extend the standard stochastic frontier framework to model input demand via the first-order condition for the firm's profit maximization problem. By comparing the optimal and actual input allocations, it is possible to find systematic over- or under-allocation towards specific inputs and calculate the profit loss from this inefficiency. Further, the underlying stochastic frontier model provides another measure of inefficiency, which the authors call technical inefficiency, otherwise known as inefficiency in production. The authors analyze privately-owned steam-electric generating plants and find evidence of overuse of capital goods, providing evidence of the Averch-Johnson effect, where firms in heavily regulated industries, especially those that focus on the price-to-capital ratio, will tend to overcapitalize to increase profit potential.

Still other researchers have focused on precise causes of inefficiency. Oum and Zhang (1995) examine the U.S. telephone industry and how the introduction of new firms and competition affects allocative inefficiencies. The authors found that when a new firm enters a market, the increased competition reduces incumbent firms' over-allocation towards capital inputs caused by the Averch-Johnson effect. Kumbhakar et al.

(1991) investigates causes of both technical and allocative inefficiency in U.S. dairy farms using a Cobb-Douglas production function. The authors relate inefficiency to various characteristics of farms including their size and education of the owner. Higher levels of education reduced both technical and allocative inefficiency, while larger farms had lower technical inefficiency but similar allocative inefficiency compared to small- and medium-sized farms.

In studying efficiency, it is also important to consider variation in the production process across firms. Tsionas (2002) used a stochastic frontier framework with random coefficients to study technical inefficiency. The author found that differences in production were responsible for most of the apparent variation in inefficiency. When the author assumed production was constant across firms, inefficiency was high and variable; when the author allowed for variation in the production process, firms were found to be highly efficient with far lower differences in efficiency across firms. While it certainly appears important to consider variability in firm production, I am not aware of any research that considers how those differences might bias estimates of allocative inefficiency.

Inefficiency of Railroads

The inefficiency of railroads has long been a focal point of research and has driven discussion of the regulation and partial deregulation of the industry. Prior to the partial deregulation of the industry, many worried that existing regulation impeded efficiency. Railroads expressed concerns over their own viability in the current regulatory environment and how efficiency was affected not only by firm's limited ability to set rates, merge with others, or abandon unprofitable routes but also by how cumbersome it was to work with regulatory agencies (Lahner, 1975). Caves et al. (1981a) compared the performance U.S. railroads with their less-regulated Canadian counterparts and found that while U.S. railroads faced higher demand and generally better economic conditions, Canadian firms experienced greater productivity growth, with excessive regulation being a major cause. Had U.S. railroads grown at the same rate as their Canadian counterparts, industry costs would have been up to 41% or $4 billion lower in 1974. Many believed that specifically allowing railroads greater flexibility over rate setting would result in large efficiency gains and improvements in viability by allowing them to more easily compete with other modes of transportation, especially highway trucking (Harbeson, 1969).[2] Still others found that deregulation would have a substantial effect on efficiency, but not because of rate flexibility but rather the ability to merge and abandon unprofitable lines (Boyer, 1987). Finally, others worried that deregulation would result in insufficient competition in the industry, leading firms to have little incentive to either reduce prices or minimize costs, potentially causing net social welfare losses (Johnson and Thomas, 1983).

Once the industry was partially deregulated in 1976 with the passage of the 4R Act and in 1980 with the passage of the Staggers Act, the effects on the industry could be more clearly seen. First, flexibility in rate setting appears to have improved the profits and viability of firms but may result in welfare losses (Boyer, 1987). Levin (1981a) found that in combination with other effects of deregulation, such as the ability to more easily merge and abandon lines, competition would fall, especially considering the geographically distinct nature of railroad networks. It was possible for regulators to promote competition in the industry by, for example, more carefully scrutinizing parallel over end-to-end mergers. However, more attention was paid to rate regulation which had the ability to limit welfare losses from excessive rates, but did little to maintain competition and provide an incentive to minimize costs (Levin, 1981b). Surprisingly, mergers had little overall

[2]For more reading on the idea that rate regulation was the greatest source of inefficiency among railroads, see Meyer et al. (1959), Friedlaender (1969), and Moore (1975).

effect on efficiency, indicating that gains earned from exercising economies of scale and scope may have been lost to the lack of competition (Chapin and Schmidt, 1999).

Few have examined the allocative inefficiency of railroads, especially recently. Notably, Kumbhakar (1988) measures the allocative errors in the industry prior to deregulation and found that between 1951 and 1975, allocative inefficiency rose dramatically for the industry as a whole, from 12.03% to 20.4% of costs, clearly indicating the need for deregulation of the industry. While many have examined the effects of falling competitive forces on prices, especially on a local level[3], I am not aware of any published research that examines the recent inefficiency of railroads (either technical or allocative) and how they relate to competitive pressures and input quality.

Conceptual Framework

I begin by assuming production is Hicks-neutral with an inefficiency component:

$$Q_{it} = A_{it} F_i(X_{it}; \Phi_{it}) \Delta_{it}. \tag{4.1}$$

In this equation, Q_{it} is the output of firm i in year t, A_{it} represents firm i's total factor productivity in year t, F_i is a transformation function that is constant across time but specific to firm i, X_{it} measures inputs, Φ_{it} measures firm and network characteristics, and Δ_{it} is a number between zero and one that quantifies efficiency. It is important to include inefficiency into the specification of the production function to draw a distinction between a firm's production potential and its actual level of output (Aigner et al., 1977). In order to maintain minimal assumptions about the shape of the production function, I opt to

[3]For more reading, see Bitzan and Wilson (2007), McFarland (1987), MacDonald and Cavalluzzo (1996), and Burton and Wilson (2006).

approximate it using a first-order log Taylor approximation:

$$q_{it} \approx \alpha_{it} - \delta_{it} + \sum_j \frac{\partial \ln F_i}{\partial \ln X^j} x_{it}^j + \sum_j \frac{\partial \ln F_i}{\partial \ln \Phi^j} \varphi_{it}^j. \tag{4.2}$$

Lower-case variables are log-transformed versions of upper case variables, and superscripts index vectors of variables. As an exception, $\Delta_{it} = \exp(-\delta_{it})$, and the inefficiency term δ_{it} is restricted to be positive to ensure that the efficiency term Δ_{it} is between zero and one. Letting partial derivatives with respect to inputs be called β_i and those with respect to network characteristics be called θ_i, Equation (4.2) can then be rewritten in log-linear form as

$$q_{it} \approx \alpha_{it} - \delta_{it} + \sum_j \beta_i^j x_{it}^j + \sum_j \theta_i^j \varphi_{it}^j. \tag{4.3}$$

Finally, I gather inputs x_{it}^j into x_{it} and characteristics θ_i^j into θ_i and label errors accrued in approximation and measurement ε_{it}, so log production can be expressed exactly as

$$q_{it} = \alpha_{it} - \delta_{it} + x_{it}\beta_i + \varphi_{it}\theta_i + \varepsilon_{it}. \tag{4.4}$$

As is, equation (4.4) assumes exact cost minimization, and as a result, the inefficiency term δ in only captures technical inefficiencies that are the result of "an equiproportianate overutilization of all inputs" (Schmidt and Lovell, 1979). Since this research is also concerned with allocative inefficiencies that result from using suboptimal proportions of inputs, I also model the firm's choice of inputs to separate technical from allocative inefficiency. Assuming firms are price-takers, profit can be written as

$$\pi(X_{it}) = P_{it}Q_{it} - \sum_j W_{it}^j X_{it}^j, \tag{4.5}$$

where P_{it} is the price firm i receives for its output in year t, W_{it}^j is the price of input j for firm i in year t, and output Q_{it} depends on inputs X_{it}. The first-order condition for profit maximization[4] with respect to input k is given by

$$\frac{\partial \pi}{\partial X_{it}^k} = P_{it}\frac{\partial Q_{it}}{\partial X_{it}^k} - W_{it}^k = 0, \tag{4.6}$$

which implies that

$$\frac{\partial Q_{it}}{\partial X_{it}^k} = \frac{W_{it}^k}{P_{it}}. \tag{4.7}$$

Now, since Q_{it} had a Cobb-Douglas form, this condition has a closed-form solution. First notice that

$$\frac{\partial Q_{it}}{\partial X_{it}^k}\frac{X_{it}^k}{Q_{it}} = \frac{\partial \ln Q_{it}}{\partial \ln X_{it}^k} = \beta_i^k. \tag{4.8}$$

Thus, using the condition from equation (4.6),

$$\frac{W_{it}^k}{P_{it}} = \frac{\partial Q_{it}}{\partial X_{it}^k} = \beta_i^k \frac{Q_{it}}{X_{it}^k}. \tag{4.9}$$

Rewritten,

$$X_{it}^k = \beta_i^k \frac{P_{it}Q_{it}}{W_{it}^k}. \tag{4.10}$$

Then, denoting the capital input X_{it}^1, the optimal ratio of capital to input k is

$$\frac{X_{it}^1}{X_{it}^k} = \frac{\beta_i^1 W_{it}^k}{\beta_i^k W_{it}^1}, \tag{4.11}$$

or in logarithms,

$$x_{it}^1 - x_{it}^k = \log\left(\frac{\beta_i^1}{\beta_i^k}\right) - \log\left(\frac{W_{it}^1}{W_{it}^k}\right). \tag{4.12}$$

[4] I assume production is concave so that there is a solution to the firm's problem, and I assume firms use positive amounts of each input, which is verified empirically.

Now, I assume that firms experience allocative inefficiency, i.e., deviation from optimal input use. I measure these deviations with errors η_{it}^k so that the observed ratio of capital to input k is

$$x_{it}^1 - x_{it}^k = \log\left(\frac{\beta_i^1}{\beta_i^k}\right) - \log\left(\frac{W_{it}^1}{W_{it}^k}\right) + \eta_{it}^k. \tag{4.13}$$

I also assume that there are systematic allocative errors, so that $\eta_{it}^k \sim N(\mu_i^k, \sigma_i^k)$. Here, μ_i^k represents the average amount firm i overcapitalizes relative to input k.

Data

To conduct my analysis, I use data from R1 financial forms filed with the United States Surface Transportation Board (STB). These forms are filed annually by each Class I railroad and describe financial and operating statistics. The data cover the time period from 1999 to 2014. This period was chosen because all Class I mergers occurred before 1999 and railroad operations have been relatively stable since, providing a sample over which differences in productions technologies across firms can be more easily identified. The Class I railroads in this sample are Burlington Northern Santa Fe (BNSF), the Canadian National Railway (CN), CSX Transportation (CSX), the Kansas City Southern Railway (KCS), the Norfolk Southern Railway (NS), the Soo Line Railroad (SOO)[5], and the Union Pacific Railroad (UP). Descriptive statistics are presented in Table 4.1.

The dependent variable in my analysis is aggregate revenue-ton-miles, defined as one ton of product which generates railroad revenue that is shipped one mile. I describe the output of revenue-ton-miles with input use and characteristics of the railroad's network. There are many inputs described in the R1 forms, and I use total numbers of cars and locomotives, total number of hours worked, amount of fuel consumed, and the book

[5]While Canadian Pacific Railway has owned the Soo Line Railroad since 1990, Soo changed in name to Canadian Pacific in the early 2000s; I will continue to refer to this railroad as SOO.

value of capital. I also control for several characteristics that differentiate production across firms. First, I include the traffic mix for each railroad, which includes both the types of goods being shipped and the different types of shipments that railroads provide (Tretheway et al., 1997). Empirically, I include the percentage of shipments that contain bulk goods[6] and the percentage of shipments that are on unit trains. I also control for the size of railroad networks, empirically represented by miles of road, defined as the total length of non-redundant track controlled by a firm. As noted by Bitzan and Wilson (2007), shipment distance is an important factor in production, so I control for the average length of haul. Finally, the quality of track may have an important effect on efficiency and maintenance costs. As discussed in Wilson (1997), average train speed should be positively correlated with track quality, so I also include it as a network characteristic.

In order to determine optimal input quantities, I use factor prices for each input. These prices are easy to calculate for labor and fuel because the R1 forms contain total costs and quantities for those factors. However, the economic cost of cars, locomotives, and track are not directly available. I infer that the opportunity cost of a capital input j is

$$\frac{\text{Annual Depreciation}_j + \text{ROI}_j}{X_j},$$

where X_j is the amount of input j used, $\text{ROI}_j = (\text{Investment}_j - \text{Accumulated Depreciation}_j) \times \text{Cost}^K$, and Cost^K is the cost of capital. The R1 forms contain information on investment, depreciation, and input quantities, and I use the Rail Cost Adjustment Factor (RCAF) published by the American Association of Railroads as the cost of capital. Finally, I use average output price, defined as total revenue divided by revenue-ton-miles, as my measure

[6] I define bulk products as belonging to one of the following categories: Metallic ores, nonmetallic minerals (not fuels), waste/scrap metals, clay/concrete/glass/stone, farm products.

TABLE 14. Descriptive Statistics

	BNSF	CN	CSX	KCS	NS	SOO	UP	Total
Output								
Revenue ton-miles (billions)	589.625	50.471	232.188	26.437	190.25	27.312	526.688	234.71
	(77.897)	(8.422)	(14.829)	(4.91)	(11.716)	(6.972)	(29.993)	(221.554)
Price	0.026	0.04	0.04	0.034	0.047	0.036	0.03	0.036
	(0.003)	(0.006)	(0.004)	(0.003)	(0.005)	(0.004)	(0.005)	(0.008)
Inputs								
Gallons of fuel (millions)	1309.606	102.26	547.779	63.283	482.172	54.176	1235.647	542.132
	(113.96)	(17.879)	(53.895)	(6.05)	(33.314)	(12.274)	(127.394)	(504.555)
Hours of labor (millions)	81.29	12.9	57.64	5.726	57.407	6.274	96.616	45.408
	(5.997)	(1.163)	(5.838)	(0.511)	(4.212)	(1.262)	(8.105)	(35.03)
Locomotives	6180.875	541	3889.562	545.75	3787.688	421.438	7913.688	3325.714
	(895.528)	(107.54)	(248.982)	(53.432)	(266.951)	(90.79)	(646.264)	(2814.122)
Cars	83837.5	24652.75	94733.75	12846.688	96078.125	14690.812	90103.75	59563.339
	(7253.763)	(6450.657)	(20162.264)	(1585.578)	(11392.32)	(1985.792)	(15344.713)	(38578.147)
Investment per mile of road (thousands of $)	32654175.23	8118768.226	19139711.582	2440727.381	19295732.892	1946351.693	37843500.091	17348423.871
	(4931781.558)	(2953618.184)	(4281948.858)	(711259.107)	(5111877.293)	(1211550.09)	(4713550.699)	(13695858.209)
Input Prices								
Fuel price	1.901	1.877	1.919	1.907	1.859	2.063	1.952	1.925
	(0.884)	(0.9)	(0.857)	(0.882)	(0.869)	(0.884)	(0.86)	(0.855)
Average wage	41.468	43.269	42.492	38.266	30.479	36.537	38.887	38.771
	(3.092)	(5.43)	(2.328)	(2.373)	(2.194)	(2.391)	(2.617)	(5.074)
Opportunity cost of locomotives	98703.232	80472.921	112343.754	44226.118	103682.365	58312.829	93820.205	84508.775
	(19954.892)	(35218.273)	(7479.84)	(32882.05)	(6127.92)	(22277.111)	(21263.086)	(32399.188)
Opportunity cost of cars	2390.587	3302.028	3520.771	1171.713	3203.51	2621.953	2817.87	2718.347
	(405.62)	(1167.264)	(495.391)	(767.3)	(329.293)	(1000.944)	(815.184)	(1047.773)
Opportunity cost of road invesment	0.128	0.126	0.126	0.112	0.12	0.124	0.126	0.123
	(0.017)	(0.016)	(0.01)	(0.013)	(0.01)	(0.014)	(0.017)	(0.015)
Network Characteristics								
Average length of haul	1046.849	288.825	533.266	363.193	464.402	440.834	921.168	579.791
	(67.523)	(22.837)	(40.975)	(43.623)	(21.087)	(38.505)	(24.854)	(271.546)
% unit shipments	0.476	0.17	0.333	0.414	0.252	0.277	0.415	0.334
	(0.031)	(0.054)	(0.022)	(0.095)	(0.02)	(0.046)	(0.021)	(0.111)
% bulk shipments	0.196	0.295	0.194	0.161	0.156	0.294	0.183	0.211
	(0.009)	(0.1)	(0.011)	(0.012)	(0.007)	(0.034)	(0.014)	(0.068)
Miles of road	32482	5847.812	21783.938	3096.312	20889.375	4123.125	32447.312	17238.554
	(424.674)	(1353.077)	(1038.289)	(163.18)	(693.435)	(1325.3)	(575.021)	(12034.503)
Herfindahl	0.526	0.095	0.441	0.05	0.362	0.052	0.474	0.286

of price. All nominal variables have been deflated using the GDP price deflator published by FRED with 2009 as the base year.

Empirical Models

Common Technology

From Section 4, recall that output is expressed as

$$q_{it} = \alpha_{it} - \delta_{it} + x_{it}\beta_i + \varphi_{it}\theta_i + \varepsilon_{it}. \tag{4.14}$$

First, it is not possible to separately identify α_{it}, δ_{it}, and ε_{it} without assuming additional structure. I first assume that inefficiency is constant across time, so that $\delta_{it} = \delta_i$ for all t. Additionally, I use the assumption that inefficiency is one-sided (specifically that $\delta_i \sim N^+(0, \sigma_\delta)$) to identify δ_i. I also assume that network characteristics affect firms in the same way, so that $\theta_i = \theta$. For the moment, I also assume firms share the same production process, so $\beta_i = \beta$. Finally, to flexibly capture changes in productivity over time, I assume it follows a random-walk with drift.[7] Specifically, I assume that

$$\alpha_{it} = \alpha_{it-1} + \tau_i + \nu_{it}, \tag{4.15}$$

where τ_i is the trend in firm i's productivity and ν_{it} is a normally distributed error term. With these assumptions, the production function becomes

$$q_{it} = \alpha_{it} - \delta_i + x_{it}\beta + \varphi_{it}\theta + \varepsilon_{it}. \tag{4.16}$$

[7]For more examples of the various uses of time-varying-parameter models, see Leybourne (1993), Mazzocchi (2003), and Del Negro and Otrok (2008).

I assume that $\nu_{it} \sim$ iid $N(0, \sigma_\nu)$, $\delta_i \sim$ iid $N^+(0, \sigma_\delta)$, and $\varepsilon_{it} \sim$ iid $N(0, \sigma_\varepsilon)$. For details on prior assumptions over the parameters, see the Appendix.

Next, recall that firm i's overcapitalization relative to input k in year t is given by

$$x_{it}^1 - x_{it}^k = \log\left(\frac{\beta^1}{\beta^k}\right) - \log\left(\frac{W_{it}^1}{W_{it}^k}\right) + \eta_{it}^k. \tag{4.17}$$

Once again, I assume that $\eta_{it}^k \sim N(\mu_\eta, \sigma_\eta)$, which captures systematic allocative inefficiencies.

Estimation of this model is similar to that of traditional stochastic frontier models. I use a Bayesian estimation framework to both mitigate problems of parameter instability and to properly express parameter uncertainty, both of which are issues that commonly appear when classical methods are used to estimate these types of models (van Den Broeck et al., 1994; Koop et al., 1995). Since the model includes equations describing both output and input decisions, estimation is conducted in stages. First, parameters in the output equation (including β) are drawn using Gibbs sampling. Then, conditional on a value for β, allocative errors can be computed and parameters describing mean and variance of allocative errors can be drawn via Gibbs sampling.

Of course, firms transform inputs to output differently from each other, and those differences can influence how allocative inefficiency is estimated. To investigate the impact of these differences, I turn to my second empirical model.

Variation in Production

Variability in the production process can clearly have an effect on estimates of allocative inefficiency. As an example, if one firm has an appreciably higher productivity of capital, it will optimally use less capital than other firms. Under the assumption of uniform production, this difference in productivity manifests as a bias towards overcapitalization.[8] To capture differences in production processes across firms and control for its effect on estimates of allocative inefficiency, I assume that β_i can vary by firm. This makes the output and inefficiency equations

$$q_{it} = \alpha_{it} - \delta_i + x_{it}\beta_i + \varphi_{it}\theta + \varepsilon_{it} \tag{4.18}$$

$$x_{it}^1 - x_{it}^k = \log\left(\frac{\beta_i^1}{\beta_i^k}\right) - \log\left(\frac{W_{it}^1}{W_{it}^k}\right) + \eta_{it}^k. \tag{4.19}$$

To separately identify each β_i, I assume they come from a common distribution; specifically, I assume that

$$\beta_i \sim N(\mu_\beta, \Sigma_\beta). \tag{4.20}$$

Estimation is again carried out in a Bayesian framework and follows two stages. The first draws samples of parameters in the output equation through Gibbs sampling, and the second draws values of parameters related to allocative inefficiency, also via Gibbs sampling. For more detail on the sampler and prior assumptions over the parameters, see the Appendix.

While this model describes the inefficiency of firms, it doesn't attribute any cause to those inefficiencies. To investigate the effect of firm characteristics on misallocation of inputs, I turn to my final model.

Relating Misallocation and Firm Characteristics

In this model, I draw connections between allocative errors and firm characteristics, especially those related to competition and input quality. That is, rather than simply

[8] This can also be seen in equation (4.13); if firm i's β^1 increases, allocative inefficiency η_i^k must be decrease to maintain the equality.

estimating the amount of allocative inefficiency as in equation (4.17), I assume that

$$x_{it}^1 - x_{it}^k = \log\left(\frac{\beta_i^1}{\beta_i^k}\right) - \log\left(\frac{W_{it}^1}{W_{it}^k}\right) + C_{it}\gamma_k + \eta_{it}^k. \tag{4.21}$$

where C_{it} describes competitive pressures felt by firm i in year t and η_{it}^k is a normally distributed empirical error. Thus, γ_k captures the effect of competitive pressures on the allocative error in input k. If X-inefficiencies are present and variables C_{it} positively measure competition, we would expect those variables to increase the magnitude of misallocation. If the industry is generally overcapitalized so that $x_{it}^1 - x_{it}^k > 0$, then we would expect coefficients γ_k to be positive if X-inefficiencies exist.

I maintain the other assumptions of the model, and as a result, estimation follows a similar process. Conditional on having prior assumptions over the parameters of this model, sampling follows a two stage process where parameters can be drawn via Gibbs sampling.

Results

This section presents and discusses results of the three empirical models discussed in the previous section. While specific results will be discussed in the proceeding subsections, parameter estimates and estimates of allocative inefficiency for each of the models can be found in Tables 2 and 3, respectively. Further, estimates of returns to scale, which is equal to the sum of input elasticities, are given in Table 4. Recall that I use Bayesian methods to estimate these models, which yields a probability distribution for each parameter conditional on the data and prior assumptions over the parameters. The estimates presented in the following section are the means of the distributions, and estimates in parentheses represent standard deviations.

TABLE 15. Parameter Estimates

	Common Production	Random Production	Competitive Variables
Technical inefficiency			
BNSF	0.279	0.317	0.338
	(0.348)	(0.412)	(0.469)
CN	0.305	0.324	0.3
	(0.464)	(0.411)	(0.416)
CSX	0.293	0.34	0.324
	(0.417)	(0.457)	(0.46)
KCS	0.296	0.325	0.304
	(0.363)	(0.491)	(0.387)
NS	0.299	0.351	0.315
	(0.38)	(0.518)	(0.402)
SOO	0.308	0.326	0.312
	(0.432)	(0.448)	(0.43)
UP	0.303	0.34	0.32
	(0.437)	(0.509)	(0.434)
Input productivities			
Capital	0.17	0.195	0.13
	(0.053)	(0.083)	(0.024)
Cars	0.09	0.151	0.157
	(0.073)	(0.076)	(0.099)
Locomotives	0.372	0.464	0.484
	(0.075)	(0.231)	(0.165)
Fuel	0.346	0.236	0.257
	(0.106)	(0.098)	(0.032)
Labor	0.051	0.085	0.097
	(0.042)	(0.042)	(0.032)
Network characteristics			
Average length of haul	0.538	0.532	0.506
	(0.122)	(0.113)	(0.116)
Miles of road	-0.277	-0.271	-0.276
	(0.126)	(0.131)	(0.123)
Percent unit	0.049	0.049	0.032
	(0.049)	(0.042)	(0.044)
Percent bulk	-0.07	-0.047	-0.082
	(0.079)	(0.076)	(0.076)
N	112	112	112
Model probability	0.2911	0.51947	0.18943

TABLE 16. Allocative Inefficiency Relative to Capital

	Common Production	Random Production	Competitive Variables		Common Production	Random Production	Competitive Variables
BNSF				*NS*			
Cars	2.016	2.498	2.541	Cars	0.673	0.173	0.862
	(1.134)	(1.026)	(0.983)		(0.496)	(0.794)	(0.683)
Locomotives	1.473	2.665	2.522	Locomotives	0.891	0.282	0.952
	(1.134)	(0.994)	(0.902)		(0.491)	(0.783)	(0.7)
Fuel	0.965	1.686	2.537	Fuel	0.945	0.536	0.88
	(1.139)	(1.102)	(0.951)		(0.498)	(0.725)	(0.683)
Labor	1.972	2.472	2.521	Labor	0.617	-0.19	0.951
	(1.143)	(0.99)	(0.906)		(0.497)	(0.837)	(0.7)
CN				*SOO*			
Cars	0.965	1.79	2.534	Cars	0.937	0.081	0.854
	(1.138)	(0.886)	(0.928)		(0.489)	(0.793)	(0.684)
Locomotives	0.724	1.009	2.521	Locomotives	0.414	0.722	1.22
	(1.162)	(0.929)	(0.906)		(1.025)	(1.233)	(1.031)
Fuel	1.947	2.233	2.539	Fuel	1.619	2.289	2.089
	(1.14)	(0.954)	(0.963)		(1.024)	(1.056)	(1.039)
Labor	1.439	1.517	2.315	Labor	0.169	0.801	1.392
	(0.472)	(0.663)	(0.591)		(1.021)	(0.965)	(1.026)
CSX				*UP*			
Cars	2.497	1.902	2.378	Cars	1.041	1.558	2.18
	(0.459)	(0.963)	(0.588)		(1.023)	(0.997)	(1.045)
Locomotives	1.735	2.129	2.327	Locomotives	0.215	0.907	1.552
	(0.445)	(0.585)	(0.579)		(1.022)	(0.998)	(1.024)
Fuel	1.708	1.539	2.385	Fuel	0.685	0.823	2.177
	(0.463)	(0.619)	(0.596)		(1.028)	(1.086)	(1.044)
Labor	1.839	2.417	2.339	Labor	0.351	0.677	1.327
	(0.454)	(0.569)	(0.572)		(1.023)	(1.014)	(1.027)
KCS							
Cars	1.548	1.431	2.385				
	(0.459)	(0.626)	(0.595)				
Locomotives	1.584	1.75	2.322				
	(0.456)	(0.592)	(0.583)				
Fuel	0.911	0.157	0.842				
	(0.489)	(0.836)	(0.686)				
Labor	1.202	1.255	0.941				
	(0.498)	(0.861)	(0.696)				

TABLE 17. Returns to Scale

	Common Production	Random Production	Competitive Variables
BNSF	1.03	1.158	1.192
	(0.118)	(0.185)	(0.141)
CN	1.03	0.873	0.872
	(0.118)	(0.171)	(0.141)
CSX	1.03	1.242	1.199
	(0.118)	(0.191)	(0.164)
KCS	1.03	1.083	1.066
	(0.118)	(0.184)	(0.143)
NS	1.03	1.263	1.243
	(0.118)	(0.178)	(0.164)
SOO	1.03	1.076	1.176
	(0.118)	(0.153)	(0.153)
UP	1.03	1.192	1.177
	(0.118)	(0.169)	(0.142)

Common Technology

Parameter estimates for this model are given in the "Common Production" column of Tables 2 and 3, respectively. Technical inefficiency, which reflects inefficiencies in transforming inputs into outputs, ranges between 0.279 and 0.308, with BNSF having the greatest technical efficiency. Specifically, these estimates indicate that BNSF produced $\exp(-0.279) \approx 75.7\%$ of what it could have if it used inputs to their maximum efficiency. Input productivities refer to the elasticity of output with respect to input variables, i.e., the percentage increase in output that results from a 1% increase in the use of a given input. Estimates relating to network characteristics are also elasticities.

Estimates of allocative inefficiency are presented as overcapitalization (i.e., investment in capital beyond what is profit-maximizing) with respect to each input and for each firm in Table 3. Estimates indicate that railroads systematically over-invest in capital relative to what is profit-maximizing. This finding provides evidence for the Averch-Johnson hypothesis, which predicts that firms in industries where regulators

TABLE 18. Pairwise Allocative Inefficiency

	Capital	Cars	Locomotives	Fuel	Labor
Capital		-1.438	-1.764	-0.883	-0.642
		(1.25)	(0.561)	(0.525)	(1.133)
Cars	1.438		-0.327	0.555	0.795
	(1.25)		(1.878)	(1.84)	(2.848)
Locomotives	1.764	0.327		0.882	1.122
	(0.561)	(1.878)		(0.59)	(1.599)
Fuel	0.883	-0.555	-0.882		0.24
	(0.525)	(1.84)	(0.59)		(1.56)
Labor	0.642	-0.795	-1.122	-0.24	
	(1.133)	(2.848)	(1.599)	(1.56)	

restrict revenue-to-cost ratios tend to overcapitalize to increase costs, thereby increasing the total profit potential.

It can also be useful to consider the overutilization of any input j to any other input k. This information is given in Table 5; the ijth entry in Table 5 measures the overutilization of input j with respect to input i on average over all firms.[9] First, as noted previously, firms tend to overcapitalize, and the excess is largest on average relative to quantities of cars and locomotives. Labor is overutilized on average with respect to every input except capital, with rates between 0.24 and 1.122, potentially providing evidence of union effects.

Since production (and therefore input elasticities) are assumed to be shared by firms in this model, estimates of returns to scale do not vary by firm. I find that estimate the mean of returns to scale is 1.03, indicating that a 1% increase in all inputs will increase output by 1.03% on average. Returns to scale near unity indicates that firms are near minimum efficient scale; that is, those firms operate where their average costs are minimized.

[9]It is important to note that these estimates are averaged over all firms, and that a specific firm might experience misallocation even when the industry doesn't on average.

Variability in Production

As mentioned in previous sections, estimates of allocative inefficiency will be biased inasmuch as the productivity of inputs varies across firms. This model investigates the relevancy of incorporating differences in production between firms and how that affects estimates of allocative errors.

First, I use Bayesian model selection to determine the importance of allowing the production technology to vary across firms. In a Bayesian framework, model selection considers a set of models, of which one is assumed to be correct, and attributes to each model the probability of being the correct model. The probability that a model M_k is the correct model conditional on the data D is given by

$$\Pr(M_k|D) = \frac{\Pr(D|M_k)\Pr(M_k)}{\Pr(D)} = \frac{\Pr(D|M_k)\Pr(M_k)}{\sum_j \Pr(D|M_j)\Pr(M_j)}, \tag{4.22}$$

where $\Pr(D|M_k)$ is the marginal likelihood of the data in model M_k and $\Pr(M_k)$ is the prior probability attributed to M_k. In general, it can be difficult to compute the marginal likelihood, but given I use a Gibbs sampler for my estimation, I use the methods in Chib and Jeliazkov (2001) for efficient calculation of model probabilities. Model probabilities are given in the second to last column of Table 2.

The only difference between this model and the Common Technology model is flexibility in describing firms' production processes. I find that the model which allows production to vary over firms is approximately 1.785 times more likely to be correct than the model which assumes input productivities are constant over firms. This clearly indicates that it is necessary to incorporate differences in the production process of firms when modeling allocative inefficiency in rail markets.

The bias that results from restrictive assumptions on firm production is clear when comparing results between the common and variable production models. To more clearly see these effects, I have included estimates of input productivities for each firm in Table 6, which can be compared to allocative inefficiency in the "Random Production" column of Table 3. There are many cases to draw from, but as an example, the productivity of fuel is higher for CN (0.304) than for any other firm (between 0.203 and 0.243). In the model that assumed a common production technology, CN was assumed to have the same productivity with respect to locomotives as other firms, resulting in a modest estimate of overcapitalization with respect to locomotives. Intuitively, after accounting for the difference in production, we should infer that since CN has high fuel productivity, it should use even more fuel than we previously estimated, resulting in a higher estimate of overcapitalization. Indeed, after accounting for variation in production processes, the estimate of CN's overcapitalization with respect to fuel increases from 1.947 to 2.233, indicative of the bias present in the common technology model.

TABLE 19. Input Productivity Estimates

	BNSF	CN	CSX	KCS	NS	SOO	UP
Capital	0.215	0.14	0.177	0.206	0.161	0.223	0.226
	(0.215)	(0.223)	(0.149)	(0.179)	(0.552)	(0.233)	(0.093)
Cars	0.149	0.18	0.162	0.149	0.157	0.127	0.126
	(0.14)	(0.226)	(0.157)	(0.571)	(0.215)	(0.218)	(0.097)
Locomotives	0.498	0.179	0.571	0.388	0.62	0.434	0.552
	(0.177)	(0.149)	(0.127)	(0.388)	(0.304)	(0.203)	(0.091)
Fuel	0.215	0.304	0.239	0.243	0.233	0.218	0.203
	(0.206)	(0.18)	(0.126)	(0.62)	(0.239)	(0.082)	(0.074)
Labor	0.082	0.07	0.093	0.097	0.091	0.074	0.085
	(0.161)	(0.162)	(0.498)	(0.434)	(0.243)	(0.07)	(0.085)

While specific estimates, especially those relating to allocative inefficiency, have changed from the previous model, general takeaways remain the same. Estimates of parameters and allocative errors can be found in the "Random Production" column of

Tables 2 and 3, respectively, and a pairwise comparison of allocative inefficiencies across inputs is given in Table 7. Firms are generally overcapitalizing with inefficiencies highest with respect to cars and locomotives, again providing evidence of the Averch-Johnson effect. Further, firms tend underutilize cars and overutilize fuel with respect to every input apart from capital. Firms also use more labor than what is profit-maximizing relative to both cars and locomotives, possibly indicative of persistent rigidities in labor allocation due to labor unions. As in the previous model, average industry overcapitalization with respect to fuel and labor is not significantly different from zero, but there is significant overcapitalization on average with respect to both cars and locomotives.

TABLE 20. Pairwise Allocative Inefficiency

	Capital	Cars	Locomotives	Fuel	Labor
Capital		-2.05	-1.812	-0.328	-1.111
		(1.125)	(0.75)	(0.912)	(1.19)
Cars	2.05		0.238	1.723	0.94
	(1.125)		(1.829)	(2.098)	(2.682)
Locomotives	1.812	-0.238		1.484	0.701
	(0.75)	(1.829)		(1.395)	(1.979)
Fuel	0.328	-1.723	-1.484		-0.783
	(0.912)	(2.098)	(1.395)		(2.248)
Labor	1.111	-0.94	-0.701	0.783	
	(1.19)	(2.682)	(1.979)	(2.248)	

Estimates of returns to scale vary across firms in this model since each firm is permitted to have a different production technology. Still, returns to scale are near unity, indicating production near minimum efficient scale. I find that CN has the lowest returns to scale at 0.873, meaning it experiences increasing average costs, while the remainder of the firms have returns to scale slightly above one, allowing them to experience increasing returns to scale.

Relating Misallocation and Firm Characteristics

The final empirical model relates firm characteristics to allocative errors. Since the structure of this model remains otherwise unchanged from the previous model, we would expect estimates of parameters and returns to scale to be similar. As seen in the "Competitive Variables" column in Tables 2 and 4, all estimates not related to allocative inefficiency change little from the previous model. Further, as seen in average allocative inefficiency in Table 8, firms still tend to overcapitalize, with the excess being greatest on average relative to cars, yet again providing evidence supporting the Averch-Johnson hypothesis.

TABLE 21. Pairwise Allocative Inefficiency

	Capital	Cars	Locomotives	Fuel	Labor
Capital		-2.518	-2.392	-0.963	-2.281
		(0.913)	(0.606)	(0.706)	(1.052)
Cars	2.518		0.126	1.555	0.237
	(0.913)		(1.2)	(1.331)	(1.939)
Locomotives	2.392	-0.126		1.429	0.111
	(0.606)	(1.2)		(0.865)	(1.473)
Fuel	0.963	-1.555	-1.429		-1.318
	(0.706)	(1.331)	(0.865)		(1.604)
Labor	2.281	-0.237	-0.111	1.318	
	(1.052)	(1.939)	(1.473)	(1.604)	

I allow allocative errors for each input to be correlated with the Herfindahl index, a measure of market share, to test for the existence of potential X-inefficiencies. The intuition is larger firms that face less competition may not have sufficient incentives to minimize costs. This effect could be further amplified by the Averch-Johnson effect in rate-regulated industries because by increasing costs, those firms could realize greater total profits. The relationship between the Herfindahl index and allocative inefficiency is shown in Table 9. The Herfindahl index does not have a significant effect on overcapitalization with respect to any variable except labor, for which it reduces overcapitalization. An increase in a firm's market share only increases overcapitalization on average relative to cars, but this effect is insignificant. Overall, these findings do not lend evidence to the X-inefficiency hypothesis; in fact, I find that larger firms tend to have lower levels of misallocation.

TABLE 22. Effects of Firm Characteristics on Allocative Errors

	Cars	Locomotives	Fuel	Labor
Herfindahl	0.043	-0.148	-0.23	-2.016
	(1.044)	(0.683)	(0.469)	(0.686)

Conclusion

The inefficiency of firms has been dissected and studied in many ways and in many contexts. Fundamentally, inefficiencies may occur despite firms' best efforts, as in the case of imperfect observation of prices, or because of incentives firms face due to regulation or the state of competition, as in the case of Averch-Johnson effects and X-inefficiencies. First, to analyze sources of inefficiency, it is crucial that the empirical framework can provide consistent estimates of inefficiency. Previous research shown the importance of controlling for differences in production over firms in obtaining unbiased estimates of technical inefficiency, but no published research has examined its effect on estimating errors in allocation. I develop models that allow production to vary flexibly over firms and estimate both technical and allocative inefficiency. I use these models to test whether incorporating flexibility in production across firms is appropriate and important in obtaining unbiased estimates of allocative errors. I go on to look for evidence of overcapitalization and Averch-Johnson effects as well as whether increased market power increases inefficiency, which could provide evidence for X-inefficiencies.

I first examine whether it is appropriate and even necessary to control for differences in production over firms when estimating allocative inefficiency. Using Bayesian model selection, I compare two models, one where the production technology is shared among firms and one where it is allowed to vary, and I find the model with differences in production better describes the data and is far more likely to be the correct model. Further, estimates of allocative inefficiency are clearly and evidently biased when it is assumed firms share the production technology. Thus, it is not only appropriate to incorporate flexibility in production, but it is also crucial to obtaining unbiased estimates of allocative errors.

I also look for evidence of the Averch-Johnson effect, which states that firms in rate-regulated industries tend to over-invest in capital to increase total profit allowed by regulators. I find no evidence of firms undercapitalizing with respect to any other input and in most cases find significant evidence of overcapitalization. This finding is present in each of my three models, providing strong evidence that firms over-invest in capital in the rate-regulated rail industry, as predicted by the Averch-Johnson hypothesis.

Finally, I look for evidence of X-inefficiencies, which can arise if a firm does not have sufficient incentives to minimize their costs. To test this hypothesis, I allow allocative inefficiencies to be correlated with the Herfindahl index, a measure of a firm's market power. I find that allocative errors decrease or don't change at all as market power increases. This finding appears to refute the existence of X-inefficiencies in the rail industry.

CHAPTER V

CONCLUSION

Considering its long history of regulation and its critical role in transportation, the railroad industry provides an interesting context to study efficiency, productivity, and competition and how each are affected by regulation. In these essays, I examine railroads after the partial deregulation of the industry and investigate recent progress in efficiency as well as potential negative effects that result from decreases in competition between firms.

I first analyze markups and scale elasticities in the industry and find that prices are significantly greater than marginal costs and that production is generally near minimum efficient scale. I then examine productivity changes in the industry and separate them into those caused by innovation and those caused by other factors such as mergers or the abandonment of unprofitable lines. I find that most firms have not seen productivity increases since 1999, and the sources of productivity growth vary depending on the ability of firms to pursue actions like line abandonment. My final chapter investigates errors in the allocation of inputs, methods of obtaining unbiased estimates of those errors, and how they are related to competitive pressures, with the understanding that competition provides the incentive for firms to both keep prices low as well as minimize costs. I find that incorporating flexibility into production across firms is vital for securing consistent estimates of allocative inefficiency and that increased market share actually decreases allocative errors, providing evidence against the X-inefficiency hypothesis in the rail industry.

These studies provide a descriptive and granular insight into key functionings of the railroads. As the industry continues to grow and serve as a part of our nation's critical transportation infrastructure, concern of issues regarding pricing, productivity, and efficiency will persist, especially in light of how the industry continues to be regulated. Further, this research expounds on how the industry has progressed since its last consolidation in 1999. While it appears that markups remain in excess of marginal costs and there is evidence of significant overcapitalization in the rail industry, firms have seen growth in productivity, turning towards technological change when other channels of increasing productivity have been exhausted, and increased market power does not appear to increase errors in allocation. While the prospects for the continued operation and success of railroads looks optimistic, attention to consumer outcomes and regulation's role in assuring those outcomes will be needed.

APPENDIX A

MARKUPS AND SCALE ELASTICITIES FOR DIFFERENTIATED RAIL NETWORKS

Bayesian Flexible Trend Model

This section presents the Flexible Trend version of the Bayesian model in its entirety. First, the data in this model are the measure of output $\widehat{q}_{it}$, capital input use $\widehat{x}_{it}^{K}$, variable input use $\widehat{x}_{it}^{V}$, network characteristics φ_{it}, and instruments Z_{it}. The instrumental variables approach to estimating this model has two stages; in the first stage I use instruments and exogenous variables to predict the endogenous variables $\widehat{x}_{it}^{V}$ and $\widehat{x}_{it}^{K}$:

$$\widehat{x}_{it}^{V} \sim N\left(F_i^V + Z_{it}\alpha_V + \varphi_{it}\beta_V, \sigma_V^2\right) \qquad \text{for each } i, t$$

$$\widehat{x}_{it}^{K} \sim N\left(F_i^K + Z_{it}\alpha_K + \varphi_{it}\beta_K, \sigma_K^2\right) \qquad \text{for each } i, t.$$

Here, F_i represents firm fixed effects, α and β are parameter vectors, and σ^2 is the measure of uncertainty in these regressions. I then construct the fitted values from this first stage, which are given by

$$\widetilde{x}_{it}^{V} = F_i^V + Z_{it}\alpha_V + \varphi_{it}\beta_V$$

$$\widetilde{x}_{it}^{K} = F_i^K + Z_{it}\alpha_K + \varphi_{it}\beta_K.$$

The second stage of the model is then given by

$$\widehat{q}_{it} \sim N\left(F_i + \mu_{it}\widetilde{x}_{it}^{V} + \eta_{it}\widetilde{x}_{it}^{K} + \varphi_{it}\beta, \sigma^2\right).$$

Now, to obtain the posterior density of μ_{it} and η_{it}, there needs to be some assumption on the distributions from which those parameters are drawn. Specifically, I assume that

$$[\mu_{1t} - 1, ..., \mu_{Ft} - 1, \eta_{1t}, ..., \eta_{Ft}]' \sim \ln MVN \left([\mu_t, ..., \mu_t, \eta_t, ..., \eta_t]', \Sigma_{2F} \right),$$

where Σ_{2F} is a covariance matrix and $[\mu_t, \eta_t]'$ is assumed to be independently and identically distributed each year as

$$\begin{bmatrix} \mu_t \\ \eta_t \end{bmatrix} \sim MVN \left(\begin{bmatrix} \varpi_1 \\ \varpi_2 \end{bmatrix}, \chi \right),$$

where ϖ_1 and ϖ_2 are hyperparameters and χ is a 2 by 2 covariance matrix. This concludes the likelihood and random parameter portion of the model, and prior distributions for each parameter are given below. This priors are intended to be diffuse in order to limit the effect of prior assumptions on posterior results.

$$F_i^V, F_i^K, F_i, \alpha_V, \alpha_K, \beta_V, \beta_K, \beta \sim \text{iid } N(0, 25)$$

$$\sigma_V, \sigma_K, \sigma \sim \text{iid Gamma}(0.5, 0.5)$$

$$[\varpi_1, \varpi_2]' \sim MVN \left([0, 0]', I_2 \right)$$

$$\chi^{-1} \sim \text{Wishart}(I_2, 2)$$

$$\Sigma_{2F}^{-1} \sim \text{Wishart}(I_{2F}, 2F)$$

Bayesian Linear Trend Model and Results

Rather than allowing μ_t and η_t to drift flexibly across time as in the Flexible Trend model, the Linear Trend version of the Bayesian model assumes that $\mu_t = \mu + \delta_\mu t$ and

$\eta_t = \eta + \delta_\eta t$, where δ_μ and δ_η are drift parameters. Thus, each μ_{it} will have a mode of $(\exp(\delta_\mu))^t \exp(\mu) + 1$ and each η_{it} will have a mode of $(\exp(\delta_\eta))^t \exp(\eta)$. My priors for these parameters are

$$\delta_\mu, \delta_\eta \sim N(0, 1)$$

$$[\mu, \eta]' \sim MVN\left([0, 0]', I_2\right).$$

All other assumptions and priors remain the same as in the Flexible Trend model.

I present results analogous to those given for the Flexible Trend version of this model. Markup and scale elasticity means and quantiles are given for each firm in Table 8. Results from the Linear Trend model are very similar to those from the Flexible Trend model; however, 2012 markup estimates tend to be higher and 2012 scale estimates tend to be lower in the Linear Trend model. Additionally, elasticities of network characteristics are similar to previous results as well. Finally, given that mean and median estimates of $\exp(\delta_\mu)$ and $\exp(\delta_\eta)$ are greater than one, I observe some evidence that markups and scale elasticities have been drifting upward over the course of the sample at a rate of 4% per year for markups and 0.8% per year for scales.

I additionally present densities for markups and scales for each firm in 2012 in Figure 5 and median markup and scale estimates for each year in Figure 6. Markup and scale estimates show similar patterns over time in both the Linear and Flexible Trend models; for example, BNSF experienced a sustained increase in its scale elasticity that began in 2007, as observed in the Flexible Trend model. There is also a clearly identifiable upward trend in markups and some positive trend in scales in this model. Given that I don't expect markups and scales to strictly adhere to a trend, I prefer the Flexible model to the Linear Trend model; however, the similarity of results between these models provides evidence for the robustness of these models.

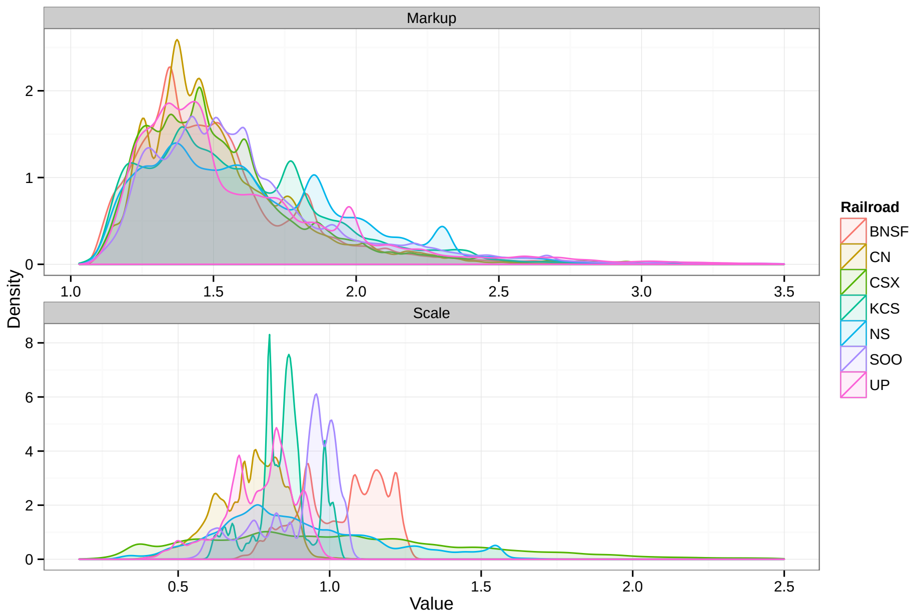

FIGURE 21. Markups and Scales in 2012

FIGURE 22. Markups and Scales Over Time

TABLE 23. Trend Model Estimation

	Mean	*Quantiles:* 5%	25%	50%	75%	95%
2012 Markups						
BNSF	1.529	1.168	1.32	1.456	1.636	2.134
CN	1.517	1.211	1.345	1.457	1.638	2.045
CSX	1.542	1.194	1.329	1.469	1.651	2.159
KCS	1.596	1.177	1.35	1.528	1.772	2.256
NS	1.646	1.184	1.36	1.573	1.862	2.31
SOO	1.627	1.218	1.377	1.535	1.739	2.372
UP	1.591	1.201	1.331	1.465	1.746	2.377
2012 Scales						
BNSF	1.043	0.804	0.927	1.074	1.159	1.225
CN	0.736	0.532	0.655	0.752	0.82	0.891
CSX	1.051	0.39	0.705	0.985	1.317	1.913
KCS	0.85	0.657	0.801	0.856	0.893	1.005
NS	0.9	0.515	0.713	0.854	1.061	1.456
SOO	0.899	0.627	0.822	0.947	0.996	1.044
UP	0.769	0.534	0.697	0.789	0.851	0.931
$\exp(\delta_\mu)$	1.04	0.907	0.97	1.023	1.099	1.223
$\exp(\delta_\eta)$	1.008	0.987	0.999	1.008	1.017	1.03
Average length of haul	0.001	-0.005	-0.002	0.001	0.004	0.007
Percent unit	-0.162	-0.342	-0.235	-0.165	-0.097	0.047
Percent bulk	0.091	-0.069	0.038	0.096	0.149	0.233
Network size	0.107	-0.094	0.008	0.063	0.206	0.395

APPENDIX B

DECOMPOSING CHANGES IN PRODUCTIVITY USING BAYESIAN METHODS

Model and Sampling Specifications

Deterministic Trend in Productivity, Constant Technology

As stated in the Empirical Models section, this model assumes productivity follows a deterministic trend that is shared across all firms. The production technology is common across firms and is constant through time. The model can be expressed in the following relations.

$$
\begin{aligned}
q_{it} &= \alpha_{it} + x_{it}\beta + \varphi_{it}\theta - \delta_i + \varepsilon_{it} \\
\alpha_{it} &= \alpha_i + \tau t \\
\alpha_i &\sim N(\mu_\alpha, \sigma_\alpha) \\
\varepsilon_{it} &\sim N(0, \sigma_\varepsilon) \\
\delta_i &\sim N^+(0, \sigma_\delta)
\end{aligned}
$$

I use Gibbs sampling to draw inference on this model and estimate the posterior distribution of the parameters conditional on the data. This distribution is the likelihood of the data conditional on the parameters and the following prior assumptions over the model parameters:

$$
\begin{aligned}
\beta, \theta, \mu_\alpha &\sim N(0, 5) \\
\tau &\sim N(0, 1) \\
\sigma_\alpha, \sigma_\varepsilon, \sigma_\delta &\sim \Gamma(1.5, 1)
\end{aligned}
$$

These assumptions were chosen to be diffuse with respect to their real world values. For example, input elasticities are rarely estimated to be greater than five,[1] which is just one standard deviation of the prior distribution.

Conditional on a value for δ_i, this is a linear random-effects model, which can be estimated via Gibbs sampling. To draw values of δ_i conditional on other parameters, first notice that

$$\begin{aligned} & p(\{\delta_i\}|\{\alpha_i\}, \mu_\alpha, \sigma_\alpha, \tau, \beta, \theta, \sigma_\varepsilon, \sigma_\delta; q, x, \varphi) \\ & \propto p(q|\{\alpha_i\}, \tau, \beta, \theta, \sigma_\varepsilon, \{\delta_i\}; x, \varphi) \times p(\{\delta_i\}|\sigma_\delta) \\ & \propto \prod_i p(q_i|\alpha_i, \tau, \beta, \theta, \sigma_\varepsilon, \delta_i; x_i, \varphi_i) \times p(\delta_i|\sigma_\delta). \end{aligned} \tag{B.3}$$

Thus, each δ_i can be drawn in its own independent block. Then, the conditional distribution of δ_i is

$$\begin{aligned} & p(\delta_i|\alpha_i, \tau, \beta, \theta, \sigma_\varepsilon, \delta_i; q_i, x_i, \varphi_i) \\ & \propto p(q_i|\alpha_i, \tau, \beta, \theta, \sigma_\varepsilon, \delta_i; x_i, \varphi_i) \times p(\delta_i|\sigma_\delta). \end{aligned} \tag{B.4}$$

Next,

$$\begin{aligned} & p(q_i|\alpha_i, \tau, \beta, \theta, \sigma_\varepsilon, \delta_i; x_i, \varphi_i) \\ & \propto \exp\left(-\frac{1}{2\sigma_\varepsilon^2}\sum_t (q_{it} - (\alpha_i + \tau t + x_{it}\beta + \varphi_{it}\theta - \delta_i))^2\right) \\ & \propto \exp\left(-\frac{1}{\sigma_\varepsilon^2}\sum_t (q_{it} - (\alpha_i + \tau t + x_{it}\beta + \varphi_{it}\theta))\delta_i - \frac{1}{2\sigma_\varepsilon^2}\sum_t \delta_i^2\right) \\ & = \exp\left(\delta_i\left(-\frac{1}{\sigma_\varepsilon^2}\sum_t (q_{it} - (\alpha_i + \tau t + x_{it}\beta + \varphi_{it}\theta))\right) + \delta_i^2\left(-\frac{T}{2\sigma_\varepsilon^2}\right)\right). \end{aligned} \tag{B.5}$$

Further, since $\delta_i|\sigma_\delta^2 \sim N^+(0, \sigma_\delta)$, the normalizing constant of this half-normal distribution does not depend on δ_i. Thus,

[1]As an example, Solow (1957) estimated the elasticity of capital to be 0.353.

$$p(\delta_i|\sigma_\delta) \propto \exp\left(-\frac{1}{2\sigma_\delta^2}\delta_i^2\right); \qquad \delta_i \geq 0. \tag{B.6}$$

So,

$$\begin{aligned} &p(\delta_i|\alpha_i, \tau, \beta, \theta, \sigma_\varepsilon, \delta_i; q_i, x_i, \varphi_i) \\ \propto \exp &\left(\delta_i\left(-\frac{1}{\sigma_\varepsilon^2}\sum_t(q_{it} - (\alpha_i + \tau t + x_{it}\beta + \varphi_{it}\theta))\right) + \delta_i^2\left(-\frac{T}{2\sigma_\varepsilon^2} - \frac{1}{2\sigma_\delta^2}\right)\right). \end{aligned} \tag{B.7}$$

This expression can then be factored so that

$$p(\delta_i|\alpha_i, \tau, \beta, \theta, \sigma_\varepsilon, \delta_i; q_i, x_i, \varphi_i) \propto \exp(-\frac{1}{2s^2}(\delta_i - m)^2),$$

where

$$\begin{aligned} m &= -\frac{\sigma_\delta^2}{T\sigma_\delta^2 + \sigma_\varepsilon^2}\sum_t(q_{it} - (\alpha_i + \tau t + x_{it}\beta + \varphi_{it}\theta)) \\ s^2 &= \frac{\sigma_\delta^2\sigma_\varepsilon^2}{T\sigma_\delta^2 + \sigma_\varepsilon^2}. \end{aligned} \tag{B.8}$$

This is the kernel of a normal distribution with mean m and standard deviation s; thus, $\delta_i|\alpha_i, \tau, \beta, \theta, \sigma_\varepsilon, \delta_i; q_i, x_i, \varphi_i \sim N^+(m, s)$, so this block can be sampled via rejection sampling or by directly sampling from a truncated normal distribution.

The posterior distribution of the parameters was estimated using 10,000 warmup iterations to achieve convergence of the Markov chain and 100,000 iterations to sample the posterior distribution. Convergence was checked by examining trace plots and autocorrelation factors. Prior distributions were also varied to ensure prior assumptions weren't driving results.

Random Walk in Productivity, Constant Technology

As discussed in the Empirical Models section, this model allows productivity to follow a more flexible process, a random walk with drift. Each firm is allowed to have its own trend in its productivity process. The production technology is still assumed to be constant across firms and time. The model can be expressed in the following relations:

$$
\begin{aligned}
q_{it} &= \alpha_{it} + x_{it}\beta + \varphi_{it}\theta - \delta_i + \varepsilon_{it} \\
\alpha_{it} &= \alpha_{it-1} + \tau_i + \eta_{it} \quad ; \quad t > 0 \\
\alpha_{i0} &\sim N(\mu_\alpha, \sigma_\alpha) \\
\varepsilon_{it} &\sim N(0, \sigma_\varepsilon) \\
\delta_i &\sim N^+(0, \sigma_\delta) \\
\eta_{it} &\sim N(0, \sigma_\eta)
\end{aligned}
$$

I once again use a Gibbs sampler to draw values from the posterior distribution of the parameters conditional on the data. This is complicated by the random walk process in productivity, but the procedure is outlined in Sarris (1973). Samples of inefficiency terms δ_i conditional on other parameters are taken from a half-normal distribution as described in Section 9.1.1. The posterior distribution is also dependent on prior assumptions, which are given below.

$$
\begin{aligned}
\beta, \theta, \mu_\alpha &\sim N(0, 5) \\
\tau_i &\sim N(0, 1) \\
\sigma_\alpha, \sigma_\varepsilon, \sigma_\delta, \sigma_\eta &\sim \Gamma(1.5, 1)
\end{aligned}
$$

The posterior distribution of the parameters was estimated using 10,000 warmup iterations and 100,000 sampling iterations. Convergence was checked using previously described methods, and various prior distributions were tested.

Random Walk in Productivity and Technology

This model allows each firm's productivity as well as the parameters describing the production technology to follow a random walk with drift. The production technology is assumed to be shared across firms, but is allowed to follow a flexible process over time. The model is expressed in the following relations:

$$
\begin{aligned}
q_{it} &= \alpha_{it} + x_{it}\beta_t + \varphi_{it}\theta - \delta_i + \varepsilon_{it} \\
\alpha_{it} &= \alpha_{it-1} + \tau_i + \eta_{it} \quad ; \quad t > 0 \\
\beta_t &= \beta_{t-1} + \rho + \psi_t \quad ; \quad t > 0 \\
\alpha_{i0} &\sim N(\mu_\alpha, \sigma_\alpha) \\
\beta_0 &\sim N(\mu_\beta, \Sigma_\beta) \\
\varepsilon_{it} &\sim N(0, \sigma_\varepsilon) \\
\delta_i &\sim N^+(0, \sigma_\delta) \\
\eta_{it} &\sim N(0, \sigma_\eta) \\
\psi_t &\sim N(0, \Sigma_\psi)
\end{aligned}
$$

I assume that Σ_ψ and Σ_β are diagonal and label the kth diagonal element of each $\Sigma_\psi^{kk^2}$ and $\Sigma_\beta^{kk^2}$. Once again, I draw inefficiency conditional on other parameters using the method described in Section 9.1.1. I use a Gibbs sampler with 5,000 warmup iterations and 50,000 sampling iterations to draw inference on the parameters. Prior assumptions over the parameters are given below.

$$
\begin{aligned}
\theta, \mu_\alpha, \mu_\beta &\sim N(0, 5) \\
\tau_i &\sim N(0, 1) \\
\rho &\sim N(0, I) \\
\sigma_\alpha, \Sigma_\beta^{kk}, \sigma_\varepsilon, \sigma_\delta, \sigma_\eta, \Sigma_\psi^{kk} &\sim \Gamma(1.5, 1)
\end{aligned}
$$

REFERENCES CITED

AAR (2013). Association of American railroads railroad facts.

Aigner, D., Lovell, C. K., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6:21–37.

Barnekov, C. C. and Kleit, A. N. (1990). The efficiency effects of railroad deregulation in the United States. *International Journal of Transport Economics*, 17(1):21–36.

Bereskin, C. G. (2009). Railroad economies of scale, scope, and density revisited. *Journal of the Transportation Research Forum*, 48(2):23–38.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Bitzan, J. D. (2000). Railroad cost conditions - implications for policy. *Prepared for the Federal Railroad Administration*.

Bitzan, J. D. and Keeler, T. E. (2003). Productivity growth and some of its determinants in the deregulated us railroad industry. *Southern Economic Journal*, pages 232–253.

Bitzan, J. D. and Keeler, T. E. (2007). Economies of density and regulatory change in the U.S. railroad freight industry. *The Journal of Law and Economics*, 50(1):157–180.

Bitzan, J. D. and Wilson, W. W. (2007). Industry costs and consolidation: Efficiency gains and mergers in the us railroad industry. *Review of Industrial Organization*, 30(2):81–105.

Bowman, R. (2013). 'Captive' Rail Shippers Plead: Set Us Free. Forbes.

Boyer, K. D. (1987). The costs of price regulation: Lessons from railroad deregulation. *RAND Journal of Economics*, 18(3):408–416.

Brown, J. H. (2013). The 'railroad problem' and the interstate commerce act. *Review of Industrial Organization*, 43:7–19.

Burton, M. and Wilson, W. W. (2006). Network pricing: Service differentials, scale economies, and vertical exclusion in railroad markets. *Journal of Transport Economics and Policy*, 40(2):255–277.

Burton, M. L. (1993). Railroad deregulation, carrier behavior, and shipper response: A disaggregated analysis. *Journal of Regulatory Economics*.

Casavant, K., Jessup, E., Prater, M. E., Blanton, B., Bahizi, P., Nibarger, D., Hill, J., and Weingram, I. (2012). Rail rate and revenue changes since the Staggers act. *Journal of the Transportation Research Forum*, 50(1).

Caves, D. W. and Christensen, L. R. (1980). The relative efficiency of public and private firms in a competitive environment: The case of Canadian railroads. *Journal of Political Economy*, 88(5).

Caves, D. W., Christensen, L. R., and Swanson, J. A. (1981a). Economic performance in regulated and unregulated environments: A comparison of U.S. and canadian railroads. *The Quarterly Journal of Economics*, 96(4):559–581.

Caves, D. W., Christensen, L. R., and Swanson, J. A. (1981b). Productivity growth, scale economies, and capacity utilization in U.S. railroads, 1955-74. *The American Economic Review*, 71(5):994–1002.

Caves, D. W., Christensen, L. R., and Tretheway, M. W. (1984). Economies of density versus economies of scale: Why trunk and local service airline costs differ. *The RAND Journal of Economics*, 15(4):471–489.

Chapin, A. and Schmidt, S. (1999). Do mergers improve efficiency? Evidence from deregulated rail freight. *Journal of Transport Economics and Policy*, 33(2):147–162.

Chib, S. and Jeliazkov, I. (2001). Marginal likelihood from the Metropolis–Hastings output. *Journal of the American Statistical Association*, 96(453):270–281.

Christensen, L. R., Jorgenson, D. W., and Lau, L. J. (1973). Transcendental logarithmic production frontiers. *The Review of Economics and Statistics*, 55(1):28–45.

Cooley, T. M. (1884). Popular and legal views of traffic pooling. *Railway Review*, 24.

Corts, K. S. (1999). Conduct parameters and the measurement of market power. *Journal of Econometrics*, 88:227–250.

Del Negro, M. and Otrok, C. (2008). Dynamic factor models with time-varying parameters: measuring changes in international business cycles. *FRB of New York Staff Report*, (326).

Eisenbeis, R. A., Ferrier, G. D., and Kwan, S. H. (1999). The informativeness of stochastic frontier and programming frontier efficiency scores: Cost efficiency and other measures of bank holding company performance. *Federal Reserve Bank of Atlanta Working Paper Series*, 99(23).

Friedlaender, A. E. (1969). *The Dilemma of Freight Transport Regulation*. Brookings Institute, Washington.

Friedlaender, A. F. and Spady, R. H. (1981). *Freight Transportation Regulation*. The MIT Press.

Friesner, D., Mittelhammer, R., Rosenmane, R., et al. (2006). Inferring the latent incidence of inefficiency from dea estimates and bayesian priors. Technical report.

Hall, R. E. (1988). The relation between price and marginal cost in U.S. industry. *Journal of Political Economy*, 96(5):921–947.

Harbeson, R. W. (1969). Toward better resource allocation in transport. *The Journal of Law and Economics*, 12(2):321–338.

Henrickson, K. E. and Wilson, W. W. (2014). *The Business of Transportation: Modes and Markets*, chapter 9, pages 176–198. Praeger.

Hilton, G. W. (1966). The consistency of the interstate commerce act. *The Journal of Law & Economics*, 9:87–113.

Ivaldi, M. and McCullough, G. (2005). Welfare tradeoffs in U.S. rail mergers.

Johnson, B. and Thomas, S. D. (1983). The Staggers rail act of 1980: Deregulation gone awry. *West Virginia Law Review*, 85:725–738.

Jondrow, J., Lovell, C. K., Materov, I. S., and Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19(2-3):233–238.

Jorgenson, D. W. and Griliches, Z. (1967). The explanation of productivity change. *The Review of Economic Studies*, 34(3):249–283.

Kadiyahi, V., Sudhir, K., and Rao, V. R. (2001). Structural analysis of competitive behavior: New empirical industrial organization methods in marketing. *International Journal of Research in Marketing*, 18:161–186.

Keeler, T. E. (1974). Railroad costs, returns to scale, and excess capacity. *The Review of Economics and Statistics*, 56(2):201–208.

Keeler, T. E. (1983). *Railroads, Freight, and Public Policy*. The Brookings Institution.

Klette, T. J. (1999). Market power, scale economies and productivity: Estimates from a panel of establishment data. *The Journal of Industrial Economics*, 47(4):451–476.

Koop, G., Steel, M. J., and Osiewalski, J. (1995). Posterior analysis of stochastic frontier models using gibbs sampling. *Computational Statistics*, 10(4):353–373.

Kumbhakar, S. C. (1988). On the estimation of technical and allocative inefficiency using stochastic frontier ffunction: The case of U.S. Class I railroads. *International Economic Review*, 29(4):727–743.

Kumbhakar, S. C., Ghosh, S., and McGuckin, J. T. (1991). A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms. *Journal of Business and Economic Statistics*, 9(3):279–286.

Lahner, W. F. (1975). Regulation and the plight of the railroad industry. *National Tax Association*, 68:225–230.

Laurits R. Christensen Associates, Inc. (2009). A study of competition in the U.S. freight railroad industry and analysiso of proposals that might enhance competition. *Surface Transportation Board*.

Leibenstein, H. (1966). Allocative efficiency vs. "x-efficiency". *The American Economic Review*, 56(3):392–415.

Levin, R. C. (1981a). Railroad rates, profitability, and welfare under deregulation. *The Bell Journal of Economics*, 12(1):1–26.

Levin, R. C. (1981b). Railroad regulation, deregulation, and workable competition. *The American Economic Review*, 71(2):394–398.

Leybourne, S. J. (1993). Empirical performance of the aids model: Constant-coefficient versus time-varying-coefficient approaches. *Applied Economics*, 25:453–463.

Lim, S. H. and Lovell, C. K. (2009). Profit and productivity of U.S. Class I railroads. *Managerial and Decision Economics*, 30(7):423–442.

MacDonald, J. M. and Cavalluzzo, L. C. (1996). Railroad deregulation: Pricing reforms, shipper rsponses, and the effects on labor. *Industrial and Labor Relations Review*, 50(1):80–91.

Mazzocchi, M. (2003). Time-varying coefficient in the almost ideal demand system: An empirical appraisal. *European Review of Agricultural Economics*.

McFarland, H. (1987). Did railroad deregulation lead to monopoly pricing? *The Journal of Business*, 60(3):385–400.

Meyer, J. R., Peck, M. J., Stenason, J., and Zwick, C. (1959). *The Economics of Competition in the Transportation Industries*. Harvard University Press, Cambridge, Massachusetts.

Miller, G. H. (1954). Origins of the Iowa Granger law. *The Mississippi Valley Historical Review*, pages 657–680.

Moore, T. G. (1975). *Promoting Competition in the Transport Industries*, chapter Deregulting Surface Freight Transportation. Brookings Institute, Washington.

Oum, T. H. and Yu, C. (1994). Economic efficiency of railways and implications for public policy. *Journal of Transportation Economics*, pages 121–138.

Oum, T. H. and Zhang, Y. (1995). Competition and allocative efficiency: The case of the U.S. telephone industry. *The Review of Economics and Statistics*, 77(1):82–96.

Park, B. U. and Simar, L. (1994). Efficient semiparametric estimation in a stochastic frontier model. *Journal of the American Statistical Association*, 89(427):929–936.

Peltzman, S. and Winston, C. (2000). *Deregulation of Network Industries*. The Brookings Institution.

Roberts, M. J. (1983). Railroad maximum rate and discrimination control. *Transportation Journal*, 22(3):23–33.

Rose, N. L. and Joskow, P. L. (1990). The diffusion of new technologies: Evidence from the electric utility industry. *The RAND Journal of Economics*, 21(3):354–373.

Sarris, A. H. (1973). Kalman filter models: A Bayesian approach to estimation of time-varying regression coefficients. *Annals of Economic and Social Measurement*, 2(4):501–523.

Schmalensee, R. (2012). "On a level with dentists?" Reflections on the evolution of industrial organization. *Review of Industrial Organization*, 41:157–179.

Schmalensee, R. L., Boyer, K. D., Ellig, J., nez, J. A. G.-I., Goodchild, A. V., Wilson, W. W., and Wolak, F. A. (2015). Modernizing freight rail regulation. *Transportation Research Board Special Report 318*.

Schmidt, P. and Lovell, C. K. (1979). Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers. *Journal of Econometrics*, 9:343–366.

Schmidt, S. (2001). Market structure and market outcomes in deregulated rail freight markets. *International Journal of Industrial Organization*, 19(1):99–131.

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 39(3):312–320.

Tretheway, M. W., II, W. G. W., and Fok, A. K. (1997). The total factor productivity of the Canadian railways, 1956-91. *Journal of Transportation Economics and Policy*, 31(1):93–113.

Tsionas, E. G. (2002). Stochastic frontier models with random coefficients. *Journal of Applied Econometrics*, 17:127–147.

United States Surface Transportation Board (2015). Complete R-1 railroad annual reports. Technical report.

van Den Broeck, J., Koop, G., Osiewalski, J., and Steel, M. J. (1994). Stochastic frontier models: A bayesian perspective. *Journal of Econometrics*, 61(2):273–303.

Waters, W. G. (2007). Evolution of railroad economics. *Research in Transportation Economics*, 20:11–67.

Wilson, W. W. (1994). Market-specific effects of rail deregulation. *The Journal of Industrial Economics*, 42(1):1–22.

Wilson, W. W. (1997). Cost savings and productivity in the railroad industry. *Journal of Regulatory Economics*, 11:21–40.

Wilson, W. W. and Bitzan, J. D. (2003). Costing individual railroad movements. *Federal Railroad Administration*.

Winston, C., Corsi, T. M., Grimm, C. M., and Evans, C. A. (1990). *The Economic Effects of Surface Freight Deregulation*. The Brookings Institution.

Yan, J., Sun, X., and Liu, J. J. (2009). Assessing container operator efficiency with heterogenous and time-varying production frontiers. *Transportation Research Part B*, 43:172–185.