JOURNAL of the TRANSPORTATION RESEARCH FORUM

Volume 49, Number 3 Fall 2010

Transportation Research Forum
NDSU Dept. 2880
PO Box 6050
Fargo, ND 58108-6050
www.trforum.org

Journal of the Transportation Research Forum

The Transportation Research Forum, founded in 1958, is an independent, nonprofit organization of transportation professionals who conduct, use, and benefit from research. Its purpose is to provide an impartial meeting ground for carriers, shippers, government officials, consultants, university researchers, suppliers, and others seeking exchange of information and ideas related to both passenger and freight transportation. More information on the Transportation Research Forum can be found on the Web at www.trforum.org.

Disclaimer:

The facts, opinions, and conclusions set forth in the articles contained herein are those of the authors and quotations should be so attributed. They do not necessarily represent the views and opinions of the Transportation Research Forum (TRF), nor can TRF assume any responsibility for the accuracy or validity of any of the information contained herein.

Subscriptions:

The Journal of the Transportation Research Forum (JTRF) is distributed to members of the Transportation Research Forum and subscribers.

Currently, the JTRF is published three times per year. Annual subscriptions are available to the general public for $150 and must be purchased for a calendar year.

An electronic version (PDF format) is available for $50 per Journal or single Journal articles for $10 each.

Subscriptions and/or membership applications may be obtained from the following office or online at www.trforum.org

ISSN 1046-1469

Published and Distributed by
Transportation Research Forum
NDSU Dept. 2880
P.O. Box 6050
Fargo, ND 58108-6050
P: (701) 231-7766 • F: (701) 231-1945

or apply online at www.trforum.org

Journal of the Transportation Research Forum
Volume 49, Number 3 • Fall 2010

Table of Contents

On the cover: Many changes have occurred in the freight railroad industry since deregulation. Marvin Prater and co-authors examine the sufficiency of rail freight competition in rural areas and the effects of intramodal competition on rail rates in "Rail Competition Changes Since the Staggers Act."

BOOK REVIEWS

A Message from the JTRF Co-General Editors

You will notice this issue contains nine high quality papers instead of the seven papers found in the previous several issues. The Fall 2010 issue of *JTRF* contains the usual wide variety of contemporary transportation topics that is the distinguishing characteristic of *JTRF*. Topics in this issue include the following:

- Design standards of traffic signals
- Improving the performance of metropolitan transportation systems
- Impact of deregulation on productivity of the Canadian for-hire trucking industry
- Benefits of passing lanes on rural two-lane highways
- Tailgating issues and potential remedies
- Railroad operational benefits in the United States
- Aircraft leasing and airline debt burdens
- Rail competition changes since the Staggers Act
- Transatlantic port issues

In "Evaluation of Design Standards of Four-Hour Traffic Signal Warrant," Xuesong Zhu, Albert Gan, and David Shen examine the minimum standard under which a traffic signal installation is appropriate. The current standard (i.e., four-hour volume signal warrant) is based on a set of critical vehicle volumes for different lane combinations of major and minor streets. The authors use simulation to evaluate the critical volumes used in the current standard. The authors found significant differences in average time delay for minor street traffic under different volume conditions, lane configurations, turning volume percentages, heavy vehicle percentages, and the number of major street lanes. They conclude that the findings provide evidence of the need to revise the current standard.

Patrick DeCorla-Sonza assesses a strategy to alleviate congestion on urban highways in "Creating a Financially Feasible High-Performance Metropolitan Transportation System." The author evaluates a strategy to reduce recurring congestion on urban highways by adding dynamic capacity during peak periods, using shoulders as travel lanes, along with variable peak period user charges levied on all lanes to manage demand and pay for capacity improvements and multimodal investments. DeCorla-Souza presents an analysis of the traffic delay, fuel consumption, CO2 emissions, and cost and revenue impacts of the strategy. The analysis indicated that the strategy would reduce delay, fuel consumption, and CO2 emissions. The author also discusses various technical and public acceptance issues of the strategy and how these issues might be addressed.

In "Productivity and Deregulation in the Canadian For-Hire Trucking Industry, 1985-1996," James Nolan, Pamela Ritchie, and John Rowcroft measure the extent to which innovation affected productivity in the Canadian for-hire trucking industry during deregulation. The authors use a unique data set spanning the trucking deregulation era in Canada. The data are used to calculate non-parametric Malmquist productivity indices which are decomposed into the change in technical efficiency between time periods (i.e., the convergence of technology among the firms in the sample over time) and technology change (i.e., changes in the overall level of technology used by firms in the industry). The authors found innovation increased in the early stages of deregulation, but declined over time. The authors attributed this to firms adjusting to the new regulatory environment by emulating the existing technology rather than continuing to pursue innovation.

Ahmed Al-Kaisy and Zachary Freedman present an empirical investigation of the operational benefits of passing lanes in "Empirical Examination of Passing Lane Operational Benefits on Rural

Two-Lane Highways." The authors state that knowledge of the operational benefits and effective length of a passing lane would result in the selection of the most appropriate length and frequency of passing lanes on longer stretches of two-lane highways. Al-Kaisy and Freedman designated three performance measures of passing lanes, which were (a) the percentage of vehicles with short headways in the traffic stream, (b) the number of follower vehicles in a directional traffic stream per mile, and (c) average travel speed as a percentage of free flow speed. The authors collected field data from two rural sites in Montana. They concluded that operational benefits right after the passing lane ranged between 33% and 42% at one of the study sites and 12% to 19% at the other site. They also found that operational benefits persist for a remarkable distance beyond the end of the passing lane.

In "Studying the Tailgating Issues and Exploring Potential Treatment," Miao Song and Jyh-Hone Wang conduct a human factors study regarding the frequency of tailgating and possible treatments to reduce it. The authors performed a vehicle headway analysis to assess the tailgating situation on Rhode Island (RI) highways. Next, they surveyed drivers' opinions regarding the causes and effects of tailgating, their experiences and perceptions of tailgating behavior, and their preferences for possible tailgating treatments. Vehicle headways were collected from highway surveillance videos and serious tailgating problems were identified on RI highways. Song and Wang found that the results of the survey confirmed the vehicle headway analysis that most RI drivers maintained insufficient vehicle headways. In terms of tailgating treatments, they concluded that most of the survey subjects preferred a system consisting of equally spaced, white horizontal bars marked on pavement, and overhead graphic-aided message signs.

Denver Tolliver, John Bitzan, and Doug Benson examine the effects of increased car weights, train size, and distance on railroad operation performance in "Railroad Operational Performance in the United States." Using revenue ton-miles per train hour as the measure of operational performance, the authors use 20 years of Class I railroad data from R-1 reports to estimate a statistical model of performance. Explanatory variables include average length of haul, average cars per freight train, net load per freight car, average way train trips per carload, train miles per running track mile, locomotives per yard track mile, miles of road, time, and vectors of commodity and railroad indicator variables. The authors found that length of haul, net load, and train size contributed significantly to operational performance gains from 1987 to 2006. However, train miles per track mile rose 89% in the same period, resulting in rail line and yard congestion.

In "Aircraft Leasing and Its Effect on Air Carriers Debt Burdens: A Comparison Over the Past Several Decades," Richard Gritta and Ellen Lippman demonstrate the switch in the characteristics of aircraft leasing before and after 1976, and quantify the effects of aircraft leases on air carrier debt burdens. The paper updates two previous studies by the authors to 2008. The authors note that prior to 1976, many airlines employed a financial lease to finance aircraft. The lease had a major advantage over buying the aircraft, which was "off-balance sheet financing," whereby the obligations under the lease appeared in the footnotes to the balance sheet. After 1976, the regulations regarding lease accounting changed to eliminate off-balance sheet financing. The authors found that in response the airlines began to lease more of their aircraft by structuring leases as short term operating leases, which aren't reported on carrier balance sheets. The authors capitalized the operating leases and found that all the airlines debt ratios appear more risky. However, they found that capitalization of operating leases doesn't affect the relative riskiness of airlines relative to one another.

Marvin Prater and co-authors examine the sufficiency of rail freight competition in rural areas and the effects of intramodal competition on rail rates in "Rail Competition Changes Since the Staggers Act." The paper begins with a review of the importance of rail transportation for U.S. agricultural producers. Using inverse Herfindahl-Hirschman Indices (HHI), the authors examine the nature of competition faced by railroads. Prater and co-authors found that rail competition for grains and oilseeds generally decreased in the 1985-2007 period, even though rail competition has increased in some Crop Reporting Districts (CRDs). Also, revenue to variable cost ratios (R/VC) increased in most CRDs, and the ratios were related to the number of railroads competing in the CRD. The authors analyzed rail competition relative to the revenue per ton, revenue per ton-mile, and R/VC ratios associated with the level of competition for six states with the least rail-to-rail competition, and distant from water transportation, with those for four states having more rail-to-rail competition and close proximity to water transportation. The results were mixed.

In “Transatlantic Port Issues,” Mary Brooks and Larissa van der Lugt examined difference between Northwest Europe and Eastern North America with respect to port commercial activities, port policy, port hinterland access and competition, port governance, and port security. The authors note that the transatlantic trade route share of the trade with Europe has declined. They explore the issues the ports on the trade route have experienced in the last decade and examine the structure and strategy of ports as they seek to deal with the changing environment on all routes, not just the transatlantic. The authors conclude that the two main issues that are increasingly important for transatlantic ports are ports congestion in relation to hinterland accessibility and port security. Port congestion appears to be a bigger issue in Northwest Europe than in North America.

Michael W. Babcock
Co-General Editor

Kofi Oberg
Co-General Editor

Evaluation of Design Standards of Four-Hour-Volume Traffic Signal Warrant

by Xuesong Zhu, Albert Gan, and David Shen

Traffic signal warrants set the minimum conditions under which a traffic signal installation may be appropriate. The four-hour volume signal warrant in the current Manual on Uniform Traffic Control Devices (MUTCD) (FHWA 2009) is applied based on a set of critical vehicular volumes for different lane combinations of major and minor streets. This paper describes an effort to apply microscopic simulation to evaluate the critical volumes used in the four-hour warrant. The results show significant differences in average control delay for minor street traffic under different volume combinations, lane configurations, turning volume percentages, heavy vehicle percentages, and the number of major street lanes (four versus six lanes), most of which are not currently considered in the four-hour warrant. This finding provides some evidence of the need to possibly revise the critical design values of the current four-hour volume warrant.

INTRODUCTION

Traffic signal warrants provide the threshold conditions under which a traffic signal installation may be appropriate. The current version of the *Manual on Uniform Traffic Control Devices* (MUTCD) (FHWA 2009), which serves as the national standard for the application of traffic control devices, includes the following eight warrants: (1) eight-hour vehicular volume, (2) four-hour vehicular volume, (3) peak-hour volume, (4) pedestrian volume, (5) school crossing, (6) coordinated signal system, (7) crash experience, and (8) roadway network. Collectively, these warrants are designed to consider different conditions under which signal control is found to be a justifiable measure for the safer and more efficient operation of an intersection.

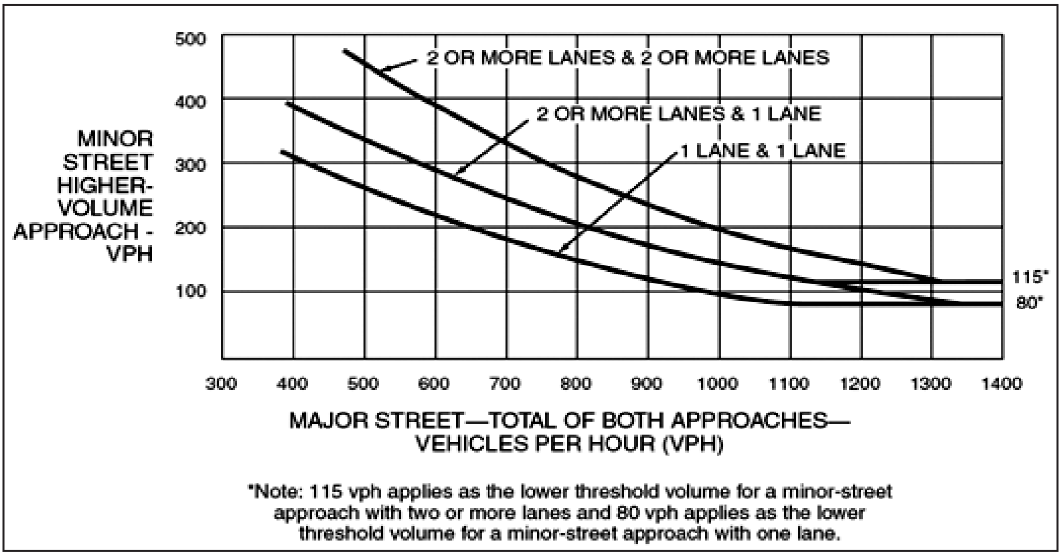
Figure 1 shows the curves representing the critical vehicular volumes for four different lane combinations of major and minor streets for low-speed (i.e., less than 40 mph for major street approach) roadways. The figure shows one curve fewer than the number of lane configurations, as two of the four configurations share the same curve. A similar set of standards with lower critical vehicular volumes is used for higher-speed (above 40 mph) roadways or for communities with a population below 10,000. The four-hour warrant is met when, in each of any four hours of an average day, the plotted point representing vehicles per hour on the major street (including the total of both approaches) and the corresponding vehicles per hour on the higher volume minor street approach fall above the critical curve for the corresponding lane configuration at the location under study.

LITERATURE REVIEW

Minimum vehicular volumes were generally used in warranting installation of a traffic signal. The origin of the four-hour warrant is found to be based on a large number of curves developed as a part of the Cooperative Research Panels Highway (NCHRP) project. These curves obtained consensus agreement among practicing traffic engineers before being included into MUTCD. Texas developed a graphic form of the four-hour warrant. This warrant, adopted by MUTCD in 1985 and continuously modified ever since, evolves into the current four-hour warrant.

Despite the widespread use, there are some concerns regarding this vehicular volume warrant. Sampson (1999) believed that warrants based on fixed vehicular volumes were inflexible and not often accepted by political decision makers. TTI (2001) recommended a reduction factor to be

Figure 1: Four-Hour Warrant for Normal Conditions



applied when vehicular volume based warrants are used to warrant the installation of a traffic signal at an intersection with pedestrian trip generators (e.g., activities center, pedestrian transportation facilities).

Researchers have been using simulation of intersection volume to assess the warranting of a traffic signal. Kell (1963) used a computer simulation to compare total delay between a pre-timed signalized intersection and a two-way stop controlled intersection. Park et al. (2000) employed Monte Carlo simulation to generate the probability that a particular traffic signal warrant can be met at an intersection. Recent advances in traffic simulation software offer the opportunity to effectively evaluate the appropriateness of critical volume standards set forth in the four-hour volume warrant more than two decades ago.

An important objective that the standards should strive to achieve is the reasonableness of the traffic conditions experienced by minor street traffic under different combinations of volumes and lane configurations. Critical volumes that result in inconsistent or unreasonable driver experiences for minor street traffic could possibly lead to either unwarranted or overlooked intersections, causing intersections with a lesser need for signalization to be selected for signal installation over intersections that have a greater need for these measures. This paper presents a study to evaluate the performance of the critical curves currently used in the four-hour warrant for low-speed conditions, often referred to as "normal conditions." Specifically, the study includes the following five evaluations:

1. The performance consistency along each critical curve and among the curves.
2. The appropriateness of sharing a critical curve for: (1) two or more lanes major and one-lane minor; and (2) one-lane major and two or more lanes minor.
3. The appropriateness of sharing the same critical curve for two and three lanes.
4. The impact of different turning volume percentages on minor street vehicles.
5. The impact of different heavy vehicle percentages on minor street vehicles.

The findings of the paper should provide academic researchers preliminary evidence of the need to revisit the current standards in the four-hour warrant. Practicing traffic engineers may use the findings to further investigate the appropriateness of applying four-hour warrant in their justification of signal installation.

METHODOLOGY OVERVIEW

The general methodology consists of the following four steps:

1. Select the appropriate measure of effectiveness (MOE) that can best measure the quality of operation for minor street traffic.
2. Identify the simulation model input data, which are vehicular volumes combinations obtained from the critical curves in the four-hour warrant.
3. Develop the simulation models for different volume combinations, turning traffic percentages, and heavy vehicle percentages under different lane configurations.
4. Analyze the simulation results and evaluate the critical volume curves based on the selected MOE.

MEASURE OF EFFECTIVENESS

Potential MOEs for measuring the operation of an unsignalized intersection can include delay, stop, fuel consumption, and average speed. It is reasonable to assume that favorable MOE values will lead to a better safety experience at intersections. In this study, control delay in seconds per vehicle (s/veh) for the minor street approach is selected as the MOE. Control delay is the component of delay that results from the type of control at the intersection. It is the difference between the travel time that would have occurred in the absence of the intersection control and the travel time that results because of the presence of the intersection control. Control delay includes initial deceleration delay, queue move-up time, stopped delay, and final acceleration delay. Control delay is the measure used in the Highway Capacity Manual (TRB 2000) to determine the level of service, and is commonly used as the measure to gauge driver satisfaction at both signalized and unsignalized intersections (Al-Omari and Benekohal 1999, Chandra et al. 2009).

MODEL INPUT DATA

Table 1 provides the critical volume combinations obtained from the three curves in Figure 1. These data are stratified by the proceeding five lane configurations (per direction as shown in Figure 2):

1. Major street has one lane and minor street has one lane.
2. Major street has one lane and minor street has two lanes.
3. Major street has two lanes and minor street has one lane.
4. Major street has two lanes and minor street has two lanes.
5. Major street has three lanes and minor street has one lane.

Figure 2: Five Lane Configurations Represented in Four-hour Warrant Curves

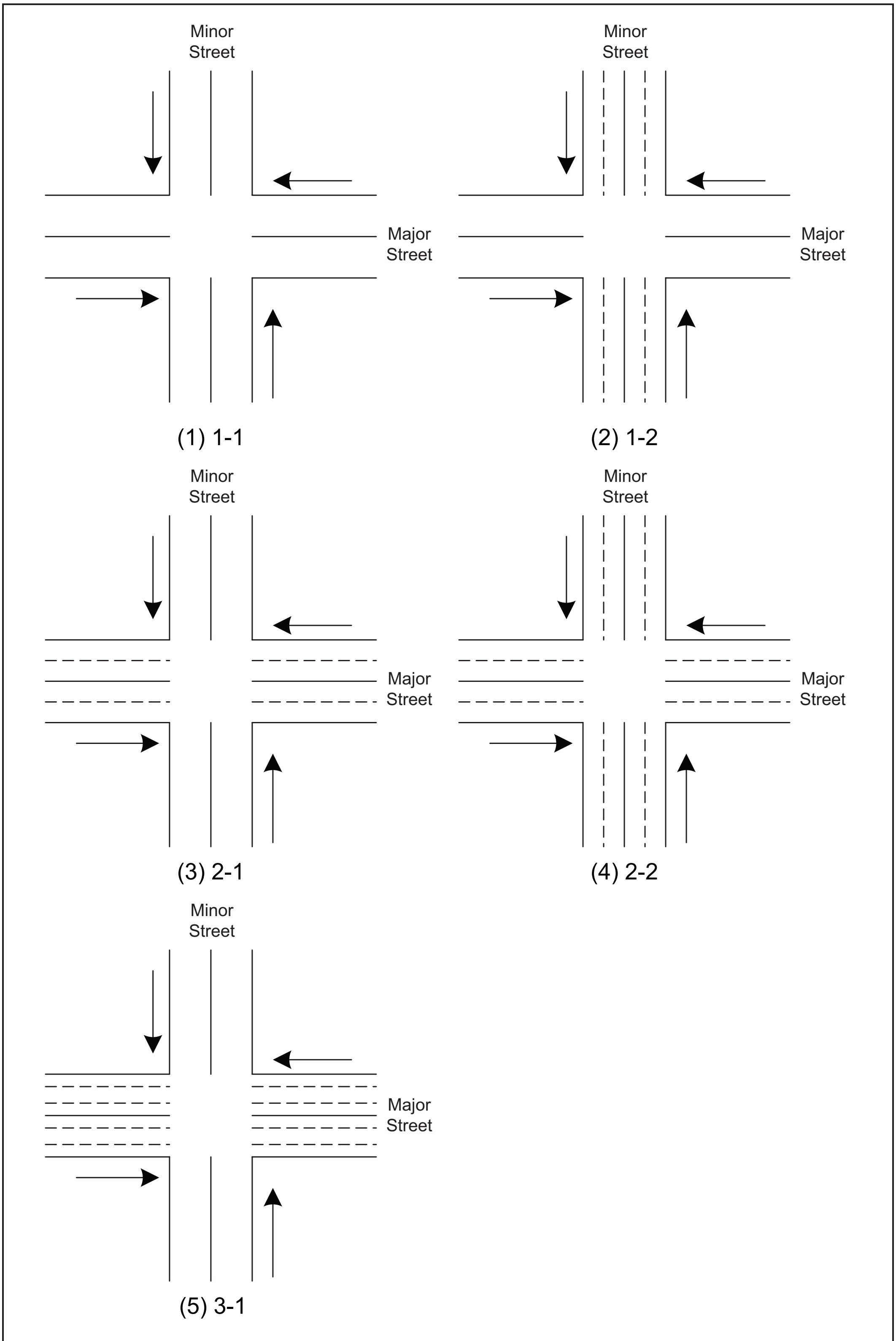


Table 1: Critical Traffic Volumes (vph) for "Normal Conditions" (from Critical Curves)

Point from Critical Curves[a]	Traffic Volumes on Major and Minor Streets (vph)									
	1 Lane Major and 1 Lane Minor (1-1)		1 Lane Major and 2 Lanes Minor (1-2)[b]		2 Lanes Major and 1 Lane Minor (2-1)[b]		2 Lanes Major and 2 Lanes Minor (2-2)		3 Lanes Major and 1 Lane Minor (3-1)[b]	
	Major	Minor	Major	Minor	Major	Minor	Major	Minor	Major	Minor
1	500	260	500	340	500	340	500	450	500	340
2	600	220	600	290	600	290	600	390	600	290
3	700	180	700	250	700	250	700	330	700	250
4	800	150	800	210	800	210	800	280	800	210
5	900	125	900	180	900	180	900	240	900	180
6	1000	95	1000	150	1000	150	1000	195	1000	150
7	1100	80	1100	130	1100	130	1100	170	1100	130
8	1200	80	1200	115	1200	100	1200	150	1200	100
9	1300	80	1300	115	1300	90	1300	115	1300	90
10	1400	80	1400	115	1400	80	1400	115	1400	80

Notes:

[a] Points selected from Critical Curves shown in Figure 1.

[b] The volume combinations for these lane configurations were obtained from the same curve, thus the volumes are the same except for (2-1), in which 115 vph on Minor Street is used for points 8, 9 and 10.

For ease of reference, these five lane configurations are labeled as 1-1, 1-2, 2-1, 2-2, and 3-1, respectively. For example, 1-1 represents the combination of one-lane major street and one-lane minor street. It is noted here that the current standard has the three-lane configurations, 1-2, 2-1, and 3-1, sharing the same curve.

Each point on the curve represents a combination of major street and minor street volumes. The major street volume is the sum of both approaches and the minor street volume represents the higher volume of the two minor street approaches. In the simulation, the volumes for each of the four approaches must be specified. For evaluation purposes, the two major approach volumes are assumed to be equal and the lower minor street approach volume is assumed to be equal to 80% of the higher minor street approach volume. Additionally, to evaluate the influence of the turning traffic percentages, the following three turning percentages, expressed as left-turn %, through %, right-turn %, were used: (10%, 80%, 10%), (20%, 60%, 20%), and (40%, 20%, 40%).

SIMULATION MODELING

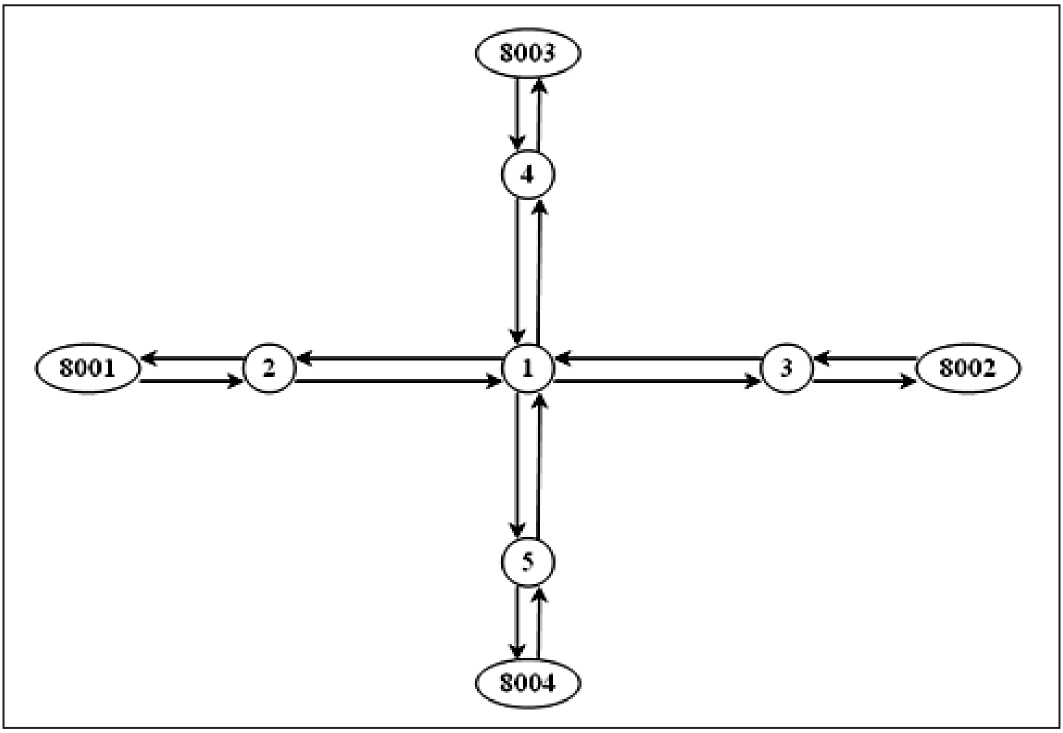
Simulation Tool

CORSIM (CORridor SIMulation) was used to simulate the different traffic scenarios and to obtain the corresponding MOE in this study. CORSIM is a microscopic simulation model and is capable of modeling traffic and traffic control conditions on both surface streets and freeway networks (ITT Systems and Sciences Corporation 1998). A microscopic simulation model is a definition relative to its macroscopic counterpart. A microscopic simulation model details the movement of individual vehicles, and traffic information associated with individual vehicles can be analyzed. First developed by the U.S. Department of Transportation in the 1970s, CORSIM has become a commonly used and accepted simulation model in the United States. CORSIM was designed to represent traffic flow on a roadway system using these commonly accepted driver and vehicle behavior models: (1) NETSIM, a microscopic stochastic simulation model for street networks, and (2) FRESIM, a microscopic stochastic simulation model for freeway networks. Furthermore, an important element of TSIS (Traffic Software Integrated System) is the TRAFVU (TRAF Visualization Utility) output processor, which allows the user to graphically view network and traffic operations.

Network Coding

CORSIM is based on a link-node network model. In this model, roadway segments are represented by links and intersections are represented by nodes. Figure 3 shows the link-node diagram of the base network used in this study. This base network was modified for different lane configurations, volume combinations, left- and right-turn percentages, and heavy vehicle percentages to simulate different geometric and traffic scenarios.

Figure 3: Link-Node Diagram of Base Network



Simulation Runs

After the CORSIM input file for the base network was coded, simulation runs for different scenarios were performed by modifying the base input file. Table 1 shows a total of 50 volume combinations for the three four-hour warrant curves. To perform the necessary analysis, three turning traffic volume percentages were then taken into account to yield a total of 150 simulation scenarios. In these 150 scenarios, the heavy vehicle percentages are assumed to be 5%. To evaluate the effect of different heavy vehicle percentages on minor street delay, another 60 scenarios based on the more common 2-1 lane configuration were created. These 60 scenarios are stratified by three different turning percentages and two different heavy vehicle percentages (10% and 20%). Therefore, a total of 210 scenarios were created.

To simulate “normal” conditions (i.e., major street approach speeds of up to 40 mph), the free-flow speed on the major street is set at 35 mph. Due to the stochastic nature of simulation models, simulation runs based on different random numbers may produce significantly different results. Five to 10 simulation replications have been commonly used in traffic simulation studies. In this study, 10 replications were performed for each scenario, which results in a total of 2,100 simulation runs. The traffic delays from the 10 replications were then averaged to obtain the final average for each scenario.

Simulation Outputs

The simulated average control delays for the 150 scenarios (5% heavy vehicle percentage) are given in Table 2. The results are stratified by the five different lane configurations, three turning volume percentages, and 10 volume combinations. Table 3 gives the simulated average control delays for another 60 scenarios stratified by different left- and right-turning percentages for two additional heavy vehicle percentages (10% and 20%).

Table 2: Average Minor Street Control Delays (seconds per vehicle, s/veh) for 150 Scenarios at 5% Heavy Vehicles

Point from Critical Curves	Average Minor Street Control Delays (s/veh)														
	1 Lane Major and 1 Lane Minor (1-1)			1 Lane Major and 2 Lanes Minor (1-2)			2 Lanes Major and 1 Lane Minor (2-1)			2 Lanes Major and 2 Lanes Minor (2-2)			3 Lanes Major and 1 Lane Minor (3-1)		
	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
1	5.0	5.5	6.0	4.6	5.2	5.5	8.2	9.2	10.0	5.8	6.5	7.6	8.5	9.8	10.9
2	5.6	6.0	6.7	5.2	6.4	6.9	8.3	9.6	11.5	6.2	6.9	8.2	9.1	11.9	12.7
3	6.0	6.6	7.6	5.4	5.8	7.2	9.0	10.9	13.7	7.1	7.8	8.8	11.2	13.8	16.2
4	6.5	8.2	9.9	6.1	6.9	8.2	9.6	13.4	18.6	7.6	9.2	11.3	12.2	17.5	20.2
5	7.1	9.3	11.0	6.5	7.3	8.8	10.6	16.7	22.8	8.2	10.5	13.2	12.8	19.1	23.9
6	8.4	10.0	11.6	7.0	6.9	9.5	11.8	18.8	24.6	10.2	11.5	14.3	15.6	21.5	26.1
7	9.8	11.8	15.2	7.3	8.7	12.1	14.2	21.3	26.8	12.3	13.5	17.6	18.5	25.4	30.5
8	10.5	12.8	15.7	9.2	11.2	12.9	17.5	25.1	31.5	14.8	16.8	19.0	21.0	30.9	35.8
9	11.9	14.4	17.6	9.9	11.4	14.2	19.5	26.7	35.1	16.5	19.2	24.4	24.8	34.5	42.1
10	13.0	18.2	22.1	12.2	13.0	16.8	22.9	43.5	48.5	21.5	28.5	31.0	28.0	55.6	65.1

Note: a, b and c represent scenarios with different turning volume percentage: a - (10%, 80%, 10%), b - (20%, 60%, 20%), c - (40%, 20%, 40%).

Table 3: Average Minor Street Control Delays (s/veh) for 60 Scenarios (2-1 Configuration)

Point from Critical Curve (2-1)	Average Minor Street Control Delays (s/veh)					
	10% Heavy Vehicle			20% Heavy Vehicle		
	a	b	c	a	b	c
1	7.7	8.0	8.8	7.5	7.0	8.5
2	7.2	9.4	10.9	7.3	8.7	10.1
3	7.9	10.0	12.6	7.5	9.0	11.0
4	9.1	12.7	16.5	8.3	10.6	14.6
5	9.7	15.3	20.2	8.8	14.0	16.8
6	10.8	16.2	21.5	10.0	14.8	18.6
7	12.6	18.3	23.1	11.9	16.1	20.4
8	15.8	21.5	27.3	14.2	18.8	24.7
9	17.7	24.5	30.3	16.0	22.0	25.6
10	20.7	36.3	42.2	19.0	32.0	35.9

Note: a, b and c represent scenarios with different turning volume percentage: a - (10%, 80%, 10%), b - (20%, 60%, 20%), c - (40%, 20%, 40%).

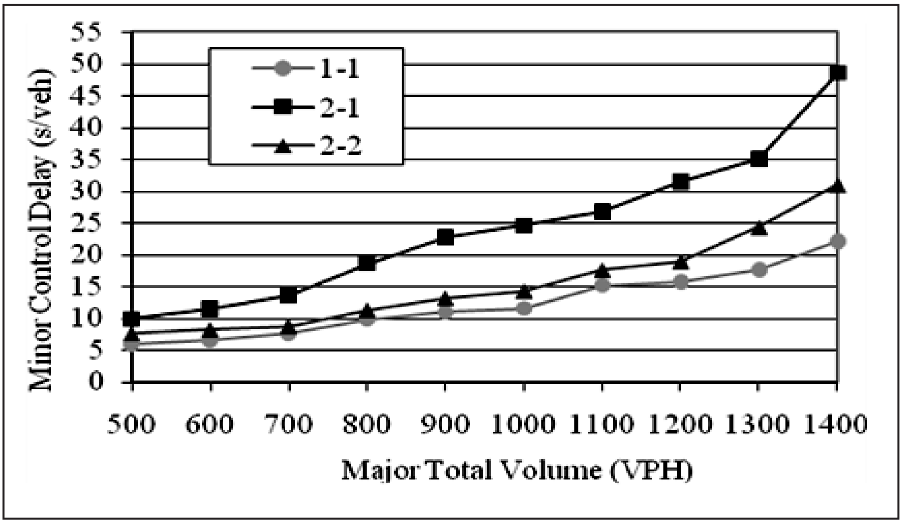
EVALUATION

Delay Consistency

Figure 4 plots the average control delay for the 1-1, 2-1, and 2-2 lane configurations for the case of (40%, 20%, 40%) turning percentages. It can be seen from the figure that minor street traffic in the case of the 2-1 lane configuration experienced a significantly higher delay than those for the 1-1 and 2-2 lane configurations, especially at higher major street volumes. When major street traffic is at 1,100 vph or higher, the delay experienced by minor street traffic for the 2-1 lane configuration is about 100% higher than that of the 1-1 lane configuration, and 60% higher than that of the 2-2 configuration. For example, when major street volume is at 1400 vph, the average delay for minor street traffic (at 80 vph, see Table 2) under the 2-1 configuration is about 48 s/veh. However, for the 2-2 lane configuration, the average delay for minor street traffic (at 115 vph, see Table 2) drops to 31 s/veh. There is no apparent reason for minor street traffic to experience 60% more delay simply because the minor street approach has one lane rather than two lanes. This suggests a possible need to adjust the two critical volumes for minor street traffic (e.g., 80 vph for one lane and 115 vph for two lanes at major street traffic = 1400 vph) in order to make the two critical volumes more consistent.

Figure 4 also shows that the delay is lengthened by major street traffic volumes at an increasing rate. While it can be argued that when major street volumes are higher, minor street vehicles are expected to tolerate a higher delay, whether the magnitude of the increase is reasonable should be subjected to further scrutiny. For example, based on the current curve in four-hour warrant (Figure 1), for the 2-1 lane configuration, an unsignalized intersection with minor street traffic experiencing as low as 10 sec/veh (at 500 vph major street traffic) has the same chance of being signalized as another intersection with minor street traffic experiencing as high as 48 sec/veh (at 1,400 vph major street traffic). This large difference suggests that the current standard is more likely to warrant an intersection with low major street volume, one that may not require signalization, rather than warrant one under high major street traffic, which evinces a greater need.

Figure 4: Delay Comparisons for Different Volume Combinations and Lane Configurations at (40%, 20%, 40%) Turning Percentages

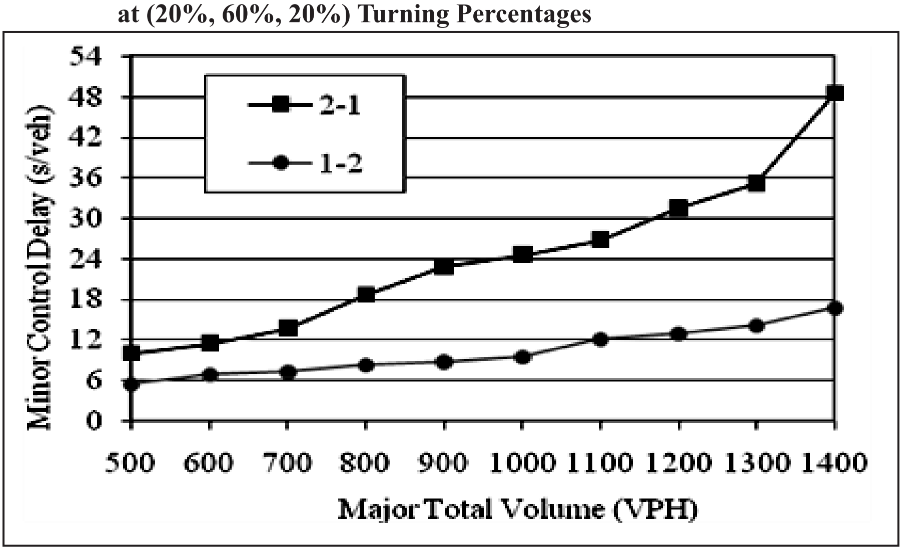


Sharing the Same Curve for 2-1 and 1-2 Lane Configurations

Figure 5 plots the control delays for the 2-1 and 1-2 lane configurations with (20%, 60%, 20%) turning percentages. Since these two lane configurations share the same critical curve, the volume combinations used in the simulation are also the same.

It can be seen from the figure that the delay for the 2-1 lane configuration is significantly higher than the 1-2 counterpart. This suggests that a separate critical curve should perhaps be used for each lane configuration, especially at high major street volume, when the two configurations experienced more than 100% difference in delay.

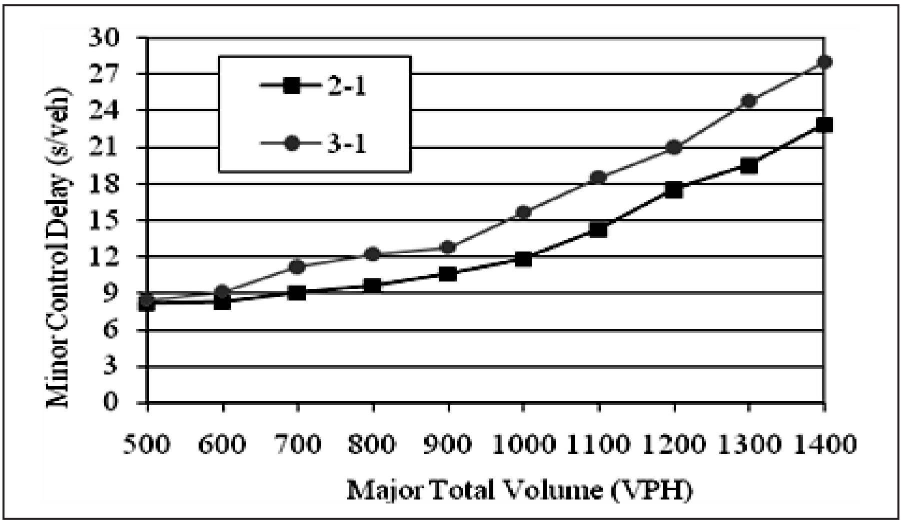
Figure 5: Delay Comparisons for 1-2 and 2-1 Lane Configurations at (20%, 60%, 20%) Turning Percentages



Sharing the Same Curve for Two Lanes and Three or More Lanes

The current four-hour warrant uses the same curve for both two lanes and more than two lanes. For evaluation purposes, a configuration with three major street lanes (i.e., 3-1) was created. Figure 6 shows the delay comparisons for the 3-1 and 2-1 lane configurations based on the (10%, 80%, 10%) turning percentages. As expected, the average control delay for the three-lane case is consistently higher than the two-lane counterpart, as it is obvious that crossing a six-lane arterial is much harder than crossing one with four lanes. In recent years, mounting traffic congestion has led to a significant increase in the number of these six-lane arterials in urban areas. It is thus important to consider the impact of two additional lanes on minor street traffic, and possibly create a separate set of critical volumes for each lane configuration.

Figure 6: Delay Comparisons Between 3-1 and 2-1 Lane Configurations at (10%, 80%, 10%) Turning Percentages



Turning Volume Percentages

Depending on local conditions, an unsignalized intersection can experience a wide range of turning movement percentages, which can have a considerable impact on minor street operation. The current standard does not consider these turning percentages. Thus, theoretically, an intersection with 100% right turns is treated the same as one with 100% left turns, even though the latter is likely to experience a much higher delay while the former may only need to increase right-turn opportunities, such as by adding a right-turn bay rather than signalization.

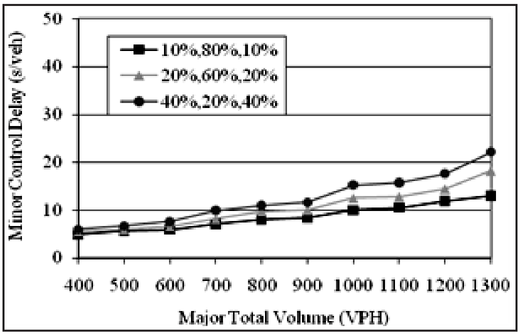
Figures 7 (a) to (d) compares the control delays for three turning percentages ranging from low to high left- and right-turn vehicles for the 1-1, 1-2, 2-1, and 2-2 lane configurations, respectively. As expected, the higher the left- and right-turn percentages, the higher the average delay experienced by the minor street vehicles. The difference also becomes increasingly larger with higher major street volumes. These observations are consistent for all four of the lane configurations. The results suggest a possible need to consider turning percentages in the four-hour warrant.

Heavy Vehicle Percentages

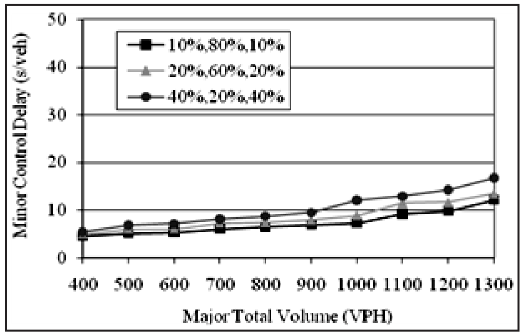
Figures 8 (a) to (c) compare the control delays for three heavy vehicle percentages ranging from low to high for three different left- and right-turn vehicle percentages, respectively. The same percentages are assumed for both minor and major streets. The results show that, as expected, the

Figure 7: Delay Comparison - Different Turning Traffic Percentages

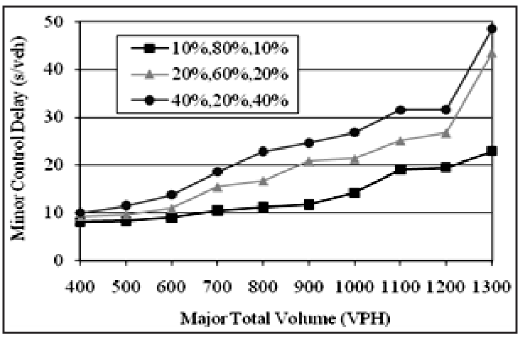
(a) 1-1



(b) 1-2



(c) 2-1



(d) 2-2

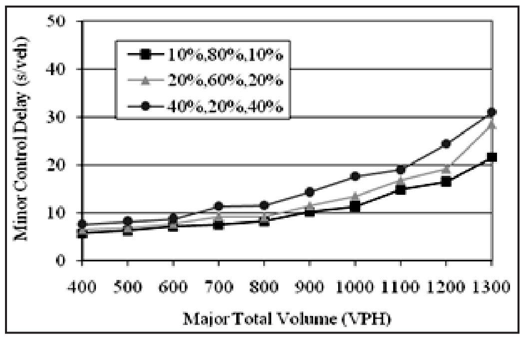
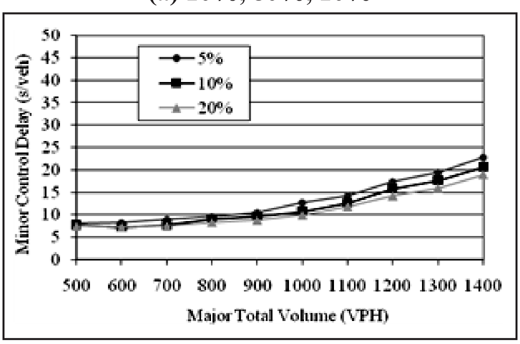
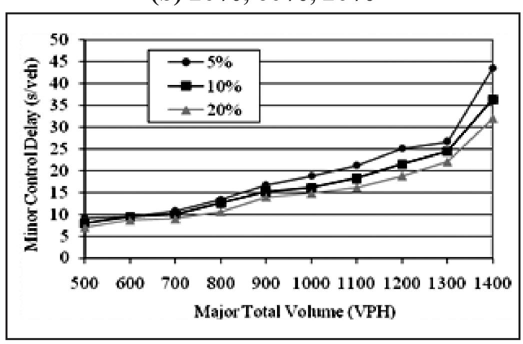


Figure 8: Delay Comparison - Different Heavy Vehicle Percentages

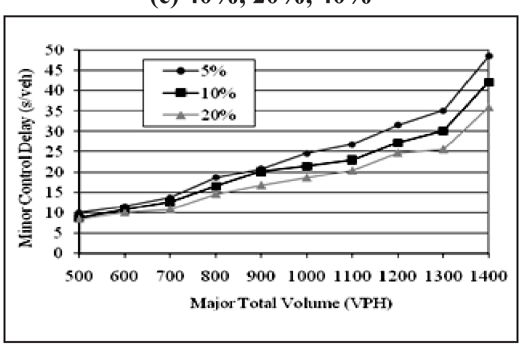
(a) 10%, 80%, 10%



(b) 20%, 60%, 20%



(c) 40%, 20%, 40%



average delay for minor street traffic increases with increasing heavy vehicle percentage, especially at higher turning volumes. It should be recognized that the difference in delay in this case is not as large simply because the percentages evaluated are not significantly different (15% between the lowest and highest percentage) to reflect the more typical range of heavy vehicles on local streets. In other words, it should not be interpreted as heavy vehicles having less of an impact than, for example, turning percentages.

CONCLUSIONS AND RECOMMENDATIONS

Today's advances in traffic simulation software provide opportunities for the reevaluation and revision of the decades old traffic standards currently in effect. This paper has presented results from a simulation evaluation of the critical volume curves of the four-hour signal warrant based on the average control delay experienced by minor street traffic. In this study, CORSIM was used as the evaluation tool to simulate a two-way stop-controlled intersection for five different lane configurations under various combinations of major and minor street volumes. The results show that:

1. The delays are significantly inconsistent along the different volume combinations of major and minor streets and among the different lane configurations evaluated.
2. There is a major difference in the delay experienced by minor street traffic under the two lane configurations (2-1 and 1-2) that currently share the same critical curves, suggesting a possible need to create a separate critical curve for each.
3. Whether the major street has four or six lanes has a major impact on minor street delay, suggesting a possible need to have separate critical curves for four-lane and six-lane arterials, especially in light of the increasing number of six-lane arterials in urban areas.
4. The percentages of left- and right-turn vehicles, which are not considered in the current four-hour warrant, have a significant impact on minor street delay.
5. The percentages of heavy vehicles, which are not currently considered in any of the existing warrants, have a significant impact on traffic operations at unsignalized intersections.

The above findings are based on the current standards for low-speed major streets (i.e., speeds of up to 40 mph for major street approach, which are normal conditions defined in MUCTD). The impacts for high-speed major streets (i.e., speeds of above 40 mph for major street approach) can be similarly evaluated. The findings of this study provide some preliminary evidence of the need to revisit the current standards in the four-hour warrant, and to perform further research to possibly refine them in order to better identify intersections that merit signalization. It is understood that various factors (e.g., empirical values, judgments) went into the establishment of the MUTCD'S critical vehicular volumes, which were not simply developed based on numerical analysis. It is also recognized that the signal warrants should generally be kept as simple as possible, however, standards that can lead to poor decisions should be reviewed and modified as needed. This is especially true for those factors that can be incorporated without significantly complicating the application, such as adjusting for more appropriate volume combinations and having separate critical curves for the 1-2 and 2-1 lane configurations, both of which can potentially achieve a more consistent traffic experience for motorists.

References

Al-Omari, B.H. and R.F. Benekohal. "Hybrid Delay Models for Unsaturated Two-Way Stop Controlled Intersections." *Journal of Transportation Engineering* ASCE 125(4), (1999): 291–296.

Chandra, S., A. Agrawal, and A. Rajamma. "Microscopic Analysis of Service Delay at Uncontrolled Intersections in Mixed Traffic Conditions." *Journal of Transportation Engineering* ASCE 135(6), (2009): 323–329.

ITT Systems and Sciences Corporation. *CORSIM User's Manual, Version 1.04*. Federal Highway Administration (FHWA), U.S. Department of Transportation, Washington, D.C., 1998.

Federal Highway Administration (FHWA), U.S. Department of Transportation. *Manual on Uniform Traffic Control Devices*. Washington, D.C., 2009.

Kell, J.H. "Intersection Delay Obtained by Simulating Traffic on a Computer." *Highway Research Record* 15, (1963): 73–97.

Park, B., M.N. Rouphail, and E.J. Hummer. "Probabilistic Approach to Implementing Traffic Signal Warrants." *Journal of Transportation Engineering* ASCE 126(4), (2000): 332–342.

Sampson, J.D. "Queue-based Traffic Signal Warrants." *Institute of Transportation Engineer ITE Journal* 69(4), (1999): 39.

Texas Transportation Institute (TTI). *Revising the Traffic Signal Warrants to Better Accommodate Pedestrians and Cyclists*. College Station, Texas, 2001.

Transportation Research Board (TRB). *Highway Capacity Manual*. Washington, D.C., 2000.

Xuesong Zhu *obtained his bachelor's and master's degrees in civil engineering from Tongji University, China, and his Ph.D. degree in transportation engineering from Florida International University. He is currently working as a transportation engineer with AECOM USA.*

Albert Gan *is an associate professor with the department of civil and environmental engineering at Florida International University. His areas of research have included traffic engineering, intelligent transportation systems (ITS), transit planning, demand modeling, and highway safety.*

David Shen *is a professor with the department of civil and environmental engineering at Florida International University. Dr. Shen's research interests center on guideway transit and transportation facility planning and design with special emphasis on intermodal guideway public transit systems.*

Creating a Financially Feasible High-Performance Metropolitan Transportation System

by Patrick DeCorla-Souza

This paper assesses a strategy to alleviate recurring congestion on metropolitan highway systems by adding "dynamic" capacity during peak periods, using shoulders as travel lanes, along with variable peak-period user charges levied on all lanes, to manage demand and pay for the capacity improvements and complementary multimodal investments. It presents an analysis of the traffic, delay, fuel consumption, CO2 emissions, and cost and revenue impacts. The paper then discusses various technical and public acceptance issues with regard to the concept, and how these issues might be addressed.

INTRODUCTION

Background

Major metropolitan areas in the U.S. are facing increasing levels of highway congestion, but it is extremely expensive to widen highways in urban areas, and environmental and social constraints limit the ability to expand the highway footprint. The normal cost of construction to add a lane on an urban freeway is estimated at about $13.4 million per lane mile by the Federal Highway Administration (2007). Weekday peak period use of a lane addition amounts to about 10,000 vehicles per weekday. As shown in Table 1, the cost per lane mile translates to a cost of about $7.00 for a 20-mile trip made on an added urban freeway lane during peak periods.

On the other hand, fuel taxes generated from a 20-mile trip amount to a total of only about 40 cents, calculated based on fuel taxes averaging 40 cents per gallon and an average vehicle fuel efficiency rating of 20 mpg. The gap between user-based taxes and cost for a new lane is even higher for high cost urban freeways (see Table 1). Moreover, if new highway capacity is provided for use free of charge, it will continue to encourage low-density land use patterns and increased auto dependence, leading to a return to congestion.

This paper assesses a new financially feasible approach to address metropolitan mobility and access issues that involves congestion pricing with complementary multimodal strategies. Congestion pricing involves the imposition of variable tolls to manage demand during peak periods. By increasing the cost of travel perceived by commuters, it can reduce congestion and provide the revenue needed to make multimodal transportation infrastructure improvements.

Review of the Literature and State of the Practice

Congestion pricing has been discussed in the economics literature for over 150 years, beginning with the work of Dupuit (1844), Pigou (1920), and Knight (1924). During the 1960s, there was a revitalization of interest with the work of Walters (1961), the Smeed Report (Ministry of Transport 1964), and Vickrey (1968).

This early work led to the earliest implementation of congestion pricing in the form of an Area Licensing Scheme in the Singapore Central Business District (CBD) in 1975. This was followed by implementation of cordon-based pricing or "toll rings" in several Norwegian cities in the late 1980s and 1990s. In the 2000s electronic toll collection (ETC) facilitated the implementation of congestion pricing in the form of a charge to drive within Central London. Stockholm has implemented cordon-based charges to drive into and out of the central city. Over the past 15 years, congestion pricing

Table 1: Costs for Highway Construction in Major Urban Areas

	Normal Cost	High Cost
Construction cost per lane mile*	$13,400,000	$55,900,000
Daily traffic volume in peak periods	10,000	10,000
Construction cost per vehicle per mile	$1,340	$5,590
Construction cost for 20-mile round trip	$26,800	$111,800
Annualized construction cost for 20-mile round trip**	$1,742	$7,267
Cost for 20-mile round trip per working day	$6.97	$29.07
Gas tax paid for a 20-mile round trip	$0.40	$0.40

* Source: FHWA 2007 (2006 dollars)
** Annualization factor 0.065 assuming a 5.25% discount rate and 30 years

has also been implemented on some existing toll facilities in France and the United States, and on a few existing or new lanes on freeways in the United States, known as express toll lanes or High-Occupancy Toll (HOT) lanes. HOT lanes are High-Occupancy Vehicle (HOV) lanes which may be used by non-HOVs with payment of a toll. With such priced lanes a two-class service is offered, for which users have a choice: a standard or "free" untolled service that may incur delay, and a premium or "express" service that offers higher speed for the price of a toll.

As Lee (2009) has demonstrated, however, in many situations single-class *unpriced* service may be better than two-class service with one class priced. Lee points out that splitting a facility into two classes reduces the effective capacity and total vehicle throughput of a facility. He finds that the justification for more than one class of service requires that the preferences (value of time) among users be very heterogeneous, but peak period travelers are probably more homogeneous than average because they are employed and have higher incomes. He concludes that the only way a priced lane can be superior to a fully-priced single-class facility is if prices are charged on *both* classes of lane, standard and premium. Nevertheless, HOT and express toll lane projects with adjacent unpriced lanes continue to be implemented in the United States, and 62 more are in various stages of planning, development, and implementation (FHWA 2009). They are publicly acceptable because generally no one is made worse off, since no existing general purpose lanes are priced.

Although a fully-priced single-class facility might be economically superior, as Lee (2009) suggests, it has encountered obstacles to implementation due to public acceptance issues. This is understandable, because both the tolled and the tolled off (i.e., those who avoid using the road during the priced periods due to the new charges) are shown to be worse off under a constant value of time assumption (Hau 2005). Those who are tolled off are "forced" into less desirable modes, routes, or times of day. Those who are tolled are worse off on average because, in order to reduce demand, the toll must exceed the average value of time saved by the tolled as a result of the smoother flow of traffic. If differences in time valuation for tolled motorists are taken into account, only those with the highest values of time are made better off. Even those who had not been using the priced road before implementation of pricing (i.e., the "untolled," such as those using public transit or those driving on alternative routes) may be made worse off as a result of more crowded buses or roads made more congested by diverted traffic. The main beneficiary is the government, which benefits from the toll revenue.

Improving the Public Acceptance of Fully-Priced Facilities

The key to gaining public acceptability of any type of congestion pricing that involves tolling of existing free roads is the use of revenues. Small (1992) suggests that two-thirds of the revenue be given back to the motorist via commuter allowances, reduced road user taxes, and a reduction in general taxes. The other third would be used to pay for new road infrastructure, improvements in public transit, and in public services to business. Goodwin (1989) suggests that the revenue be arbitrarily divided such that a third would be used for new roadway infrastructure, a third would be allocated to improve public transport, and a third would be used to reduce the general tax burden or increase social spending.

This paper presents a direct approach for use of the revenues. It attempts to more directly "make whole" each of the three categories of travelers who are made worse off, i.e., the tolled, the tolled off, and the untolled. To benefit the tolled, it attempts to provide benefits commensurate with the value of tolls paid by providing significant travel time savings for them. This is accomplished by providing new highway capacity and advanced technologies that maintain the free flow of traffic and reduce the impact of incidents and accidents. To benefit the tolled off and the untolled, it enhances travel options, deploys new technologies for multimodal traveler information, and provides more effective traffic incident management and arterial traffic management.

The concept builds on emerging strategies implemented in Europe and now being explored by the transportation community in the United States as possible options for providing new highway capacity without the need for new rights-of-way or major reconstruction. However, the safety of these approaches has not yet been fully assessed in the U.S. context. Therefore, the concept proposed in this paper is not ready for immediate application. Rather, the intent is to engender discussion and further exploration through collaboration among the transportation planning, finance, safety, and operations communities to find workable strategies to advance the concept in the United States.

FLEXIBLE AND EFFICIENT EXPRESS (FEE) HIGHWAYS

Providing Benefits for the Tolled

In congested metropolitan areas, a new "dynamic" travel lane could be created on limited access highways by narrowing existing lanes and using shoulder space. The shoulder travel lane would be open for use by authorized vehicles with trained drivers in conjunction with active management of all lanes on the highway, using overhead lane controls and dynamic message signs to harmonize speeds and keep traffic flowing freely and safely (Figure 1). During peak periods, variable user fees would be used in conjunction with ramp metering to keep demand for use within the capacity of the reconfigured highway. Should an incident occur in any lane, surveillance cameras would automatically communicate with overhead lane controls, which would shut down the appropriate lane(s) in advance of the incident location.

User fee rates would be pre-scheduled according to time of day and location in order to manage demand based on observed traffic patterns. Dynamic pricing, which involves changing toll rates as often as every few minutes in response to real time traffic conditions, would not be used. This is because when an entire highway facility is priced, it is necessary to pre-schedule toll rates so that travelers can make rational travel choices before they begin their trips. To accomplish the same effect on flow as dynamic pricing, ramp metering would be used to fine tune traffic volume on the priced facilities by controlling access to the highway when demand is too close to capacity. This will ensure that traffic flow does not break down. Fees would be collected electronically at free-flow highway speeds using in-vehicle electronic tags. Vehicles without valid electronic tags would be identified and charged using license plate recognition technology. Buses and pre-registered vanpools would be issued special transponders that would exempt them from the fees.

Highway performance could be guaranteed to users. When an incident shuts down a lane or two, leading to congestion delays, all fees could be suspended since users would not get the level of service promised. The potential loss of revenue would provide an additional incentive to the highway operator to clear incidents quickly, and would reinforce the "fee for service" nature of the new charges in the minds of motorists.

This concept is termed "Flexible and Efficient Express" or FEE highways by the author, in order to reflect the "fee for service" aspect of this concept. It would require little or no new rights-

Figure 1: Shoulder Travel Lane with Overhead Lane Controls in the Netherlands

of-way, could be financed by leveraging the user fee revenue stream, and could be implemented in a relatively short period of time due to limited or no expansion of the highway footprint. Revenues from user fees could be used to pay for operations and maintenance, and for capital costs for shoulder improvements, new emergency pull-off areas, tolling infrastructure, active traffic management systems, enhanced transit and ridesharing services to provide viable alternatives to driving alone, and traffic flow improvements on alternative free routes.

In addition to alleviating recurring congestion, FEE highways could increase person throughput as well as vehicle throughput. Vehicle throughput would increase because an extra lane (i.e., the dynamic shoulder travel lane) would be available for use when needed. Also, demand management with pricing and ramp metering would prevent traffic flow breakdowns and thus eliminate the loss of vehicle throughput that currently occurs on unmanaged highways due to breakdown of traffic flow at bottlenecks (DeCorla-Souza 2010). Person throughput would increase due to increased use of transit and ridesharing.

Limited access highways with standard lanes and shoulders could be re-striped as follows to create FEE highways:

- Left shoulder with reduced width, if feasible;
- 11-foot-wide lanes, i.e., the lanes would be reduced from the standard 12-foot width, similar to narrowing of lanes in construction zones and on other urban facilities which accommodate heavy vehicles, albeit at lower speeds;
- A 13-foot-wide dynamic shoulder lane on the extreme right. (The wider lane is needed to allow for adequate lateral clearance between vehicles in the lane and any structures on the highway's edge.)

Providing Benefits for the Tolled Off and the Untolled

Since costs for highway expansion would be significantly reduced by the FEE highway concept and users of *all* lanes would be charged tolls, a significant surplus of revenues is anticipated. (An analysis to confirm this is presented later in the paper). The FEE concept would use surplus revenues for the benefit of the tolled off and the untolled by investing in complementary strategies as discussed below.

Enhancement of travel choices. Bus Rapid Transit (BRT), express bus and bus priority systems (e.g., signal preemption) would create a "backbone" of fast and frequent transit service, supplemented with vanpooling, regular carpooling, and flexible carpooling. Flexible carpooling is similar to "slugging" as practiced in Northern Virginia, Houston, and the San Francisco Bay area, but with pre-certification of drivers and riders (Minett 2010a and 2010b). With flexible carpooling, potential carpoolers gather at designated locations and are picked up by a solo driver who seeks to qualify for use of HOV lanes or to avail himself or herself of HOV toll discounts or exemptions on priced facilities. This type of carpooling would supplement transit services and help reduce costs for operating peak period transit services while offering commuters an additional travel choice.

Flexible telecommuting. Incentives would be provided to employers to allow some of their employees to telecommute each morning until peak period tolls are no longer in effect and then drive in to work, or to leave work before peak tolls are in effect and then work from home.

Off-peak transit use incentives. If transit vehicles are too crowded in peak hours, some pressure could be relieved by off-peak or peak shoulder discounts or other incentives, such as the chance to win a lottery (Merugu, Prabhakar, and Rama 2009).

Technology. Pre-trip decisions on mode and time of travel could be made easier for the traveler by incorporating technologies for timely and reliable multimodal traveler information. Those who continue to drive could benefit from investment in technologies that provide traffic flow improvements, more effective traffic incident management, arterial and corridor management, and safety and mobility management in work zones.

Implementation

With so many different components and innovations, it will be difficult to implement the FEE highway concept all at once. Moreover, it is important to ensure that all strategies that are designed to provide better travel options are in place *before* actual congestion charging begins so that the public has time to become fully aware of the new options, and to ensure that they are all working effectively. Implementation of FEE highways might be pursued by corridor. The first step could be creation of shoulder travel lanes and use by transit and vanpool vehicles with trained drivers. This could be followed by opening up the lanes for use by carpools with three or more persons (with trained drivers) to provide a carpooling incentive. With these multimodal alternatives in place, the final step would be implementation of full pricing on all lanes along with multimodal traveler information systems and active traffic management.

Public-private partnerships (PPPs) could be used to increase the potential for innovation and on-time delivery. One type of PPP arrangement attempts to address the conflict between the private sector goal to maximize revenue and the public goal to maximize vehicle throughput. Under this method, termed "Concurrent Real and Shadow Tolling" (DeCorla-Souza 2006), the concessionaire would have freedom to set the toll rates needed to ensure free-flowing traffic during congested periods, but all "excess" charges above a benchmark flat toll rate (called a shadow toll) specified in the concession agreement would go the public authority. To maximize its shadow toll revenue, the private operator would need to maximize vehicle throughput, thus maximizing public goals.

IMPACTS

Traffic Volumes, Speeds and Travel Times

Table 2 presents results from a traffic impact analysis of the FEE highway concept for a 10-mile highway segment with six existing lanes (i.e., three per direction) reconfigured to four lanes per direction (i.e., total of eight lanes). For comparison, a "No Pricing" alternative with the same configuration was also analyzed, and the results are presented in Table 3.

The analysis was performed for a prototypical urban freeway that is severely congested, using procedures from the *Highway Capacity Manual* (Transportation Research Board 2000), i.e., the HCM, and the following inputs and analytical procedures:

- The existing urban freeway is assumed to have standard 12-foot-wide lanes and a standard 10-foot-wide right shoulder.
- Existing traffic demand is assumed to be at the maximum level that can be accommodated without breakdown of traffic flow (but at slow speeds), known as Level of Service E (LOS E). In reality, many segments of urban highways operate under breakdown flow conditions involving queuing. This is known as Level of Service F (LOS F). The LOS E traffic demand assumption provides a conservative estimate of benefits from pricing. For long-term analysis, a 5% (total) secular growth is assumed in order to account for population and economic growth. This throws the facility into breakdown flow conditions, i.e., LOS F.
- To be conservative in estimating benefits, it is assumed that there will be no change in travel mode. An increase in transit and carpool use would increase person throughput on the FEE highway, and consequently the benefits.
- For the No Pricing alternative, induced and diverted traffic resulting from the increased capacity is estimated using elasticity of travel demand relative to travel time, rather than elasticity relative to total generalized cost of travel, because the other components of generalized cost perceived by motorists (i.e., vehicle operating cost and accident risk) will not change significantly as a result of capacity expansion. A short-term elasticity of demand relative to travel time of -0.28 and a long-term elasticity of -0.57 is used based on prior studies (DeCorla-Souza and Cohen 1997). Induced traffic is estimated in the column titled "GP Lanes (Initial)" in Table 3. A total of 984 vehicles per hour (vph) is induced, amounting to a 14% increase above the base case traffic volume of 6,930 vph.
- It is assumed that a minimum speed of 55 mph is to be maintained on the FEE highway, and traffic volume on the highway is estimated on that basis. As indicated earlier, dynamic "real time" pricing is not needed because ramp metering would be used to hold at the entrance ramps any excess demand that threatens to reduce speeds below that threshold. Note that the FEE highway can accommodate more vehicles than in the base (six-lane) configuration due to the additional travel lane in each direction (see Table 2). However, the pricing mechanism curbs the higher level of induced travel seen in Table 3 when all lanes are unpriced. A small amount of induced traffic raises traffic volume to 7,069 vph, a 2% increase above the base case traffic volume of 6,930 vph.
- Free flow speeds for 12-foot lanes are estimated at 65 mph based on Exhibit 13-6 of the HCM.
- Free flow speeds for 11-foot lanes are based on Exhibit 23-4 of the HCM; these speeds are further reduced due to reduced right shoulder lateral clearance, based on Exhibit 23-5 of the HCM. The resulting free flow speed is estimated at 60.7 mph.
- Actual operating speeds are estimated based on the Bureau of Public Roads (BPR) formula, with parameters from Nakamura and Kockelman (2002). The BPR formula and parameters used are explained in the footnotes of Table 2. Actual travel speeds are estimated at 46 mph for No Pricing and 55 mph for FEE highways.

Highway lane capacity is estimated as follows:

- For existing 12-foot-wide general purpose lanes, capacity is estimated at 2,350 vph, based on Exhibit 13-6 in the HCM.
- Base lane capacity for 11-foot-wide general purpose lanes is estimated as follows, based on Exhibit 23-3 of the HCM:
    - Capacity of 11-foot-wide lane = 1,700 + (10 x Free flow speed) = 1,700 + (10 x 60.7) = 2,307 vph.

Table 2: Traffic Impacts on a 10-mile long 6-Lane Freeway Reconfigured into Four FEE Lanes per Direction

	Existing Configuration	Four FEE Lanes
Traffic volume (demand)	6,930	7,069
Number of lanes per direction	3	4
Traffic volume per lane	2,310	1,767
Capacity per lane	2,350	2,307
Ratio of service flow volume at LOS C* to capacity	0.65	0.65
Service flow volume per lane at LOS C	1,528	1,500
Ratio of traffic volume to service flow volume at LOS C	1.51	1.18
BPR alpha (α) coefficient**	0.02	0.02
BPR beta (β) coefficient**	10	10
Free flow speed (mph)	65.00	60.70
Travel time per mile with free flow speed (min.)	0.92	0.99
Actual travel time per mile (min.)	2.08	1.09
Time saved per mi. relative to Existing Configuration (minutes)		0.99
Speed (mph)	28.87	55.01

*LOS C = Level of Service C. Engineers grade quality of service on roadway facilities at six levels, A through F, with LOS A representing the best operating conditions and LOS F the worst. LOS C is considered an operating service level "acceptable" to users. LOS F represents conditions where traffic flow has broken down on a freeway.

** The Bureau of Public Roads (BPR) formula estimates travel time on a facility based on free-flow speed and the ratio of traffic volume to service flow volume at LOS C. Based on the BPR formula, actual travel time = free-flow travel time x (1+ α (ratio of traffic volume to service flow volume at LOS C)$^{\beta}$).

Other Impacts

Tables 4 and 5 present other impacts of the FEE highway and the No Pricing alternatives, respectively. Safety impacts have not been estimated. Since this is a prototypical analysis, user benefits (other than safety) are estimated using national average coefficients. For any specific urban area, coefficients and results would be different. The analytic procedures were as follows:

- **Time savings**. Time saved by induced and diverted travelers on free lanes for the No Pricing alternative is estimated at half the time saved by prior travelers on the free lanes, consistent with consumer surplus theory. Value of time saved is calculated based on the nationwide average value of time per vehicle hour from the *Urban Mobility Report* (Texas Transportation Institute 2007).
- **Fuel cost savings**. Fuel consumption savings due to reduced delays is estimated at 0.69 gallons per hour of delay saved based on the relationship between congestion delay and fuel consumption in Exhibit 1 of the 2007 *Urban Mobility Report* (Texas Transportation Institute 2007). Fuel cost is estimated at $2.50 per gallon.
- **Annualization of benefits**. Congestion levels are assumed to prevail only during four hours per direction on each weekday, on 250 working weekdays per year. For example, the *inbound* direction of an urban radial freeway may be congested for three hours in the morning, but for only one hour in the afternoon, for a total of four hours each day.
- **Toll Rates**. Recall that FEE highways seek to ensure that the benefits (primarily travel time savings) to tolled vehicles are at least equal to the value of the toll paid, to ensure that

Table 3: Traffic Impacts on a 10-mile 6-Lane Freeway Reconfigured to Four Free Lanes per Direction with No Pricing

		Four Free Lanes with No Pricing		
	Existing Configuration	**Free Lanes (Initial)**	**Free Lanes incl. Induced traffic**	**Total**
Traffic Impacts				
Traffic volume (demand)	6,930	6,930	7,914	7,914
Number of lanes per direction	3	4	4	4
Traffic volume per lane	2,310	1,733	1,978	
Capacity per lane	2,350	2,307	2,307	
Ratio of service flow volume at LOS C to capacity	0.65	0.65	0.65	
Service flow volume per lane at LOS C	1,528	1,500	1,500	
Ratio of traffic volume to service flow volume at LOS C	1.51	1.16	1.32	
BPR alpha coefficient	0.02	0.02	0.02	
BPR beta coefficient	10	10	10	
Free flow speed (mph)	65.00	60.70	60.70	
Travel time per mile with free flow speed (minutes)	0.92	0.99	0.99	
Actual travel time per mile (minutes)	2.08	1.07	1.30	
Time saved per mi. relative to Existing Configuration (minutes)		1.01	0.77	
Speed (mph)	28.87	55.96	46.00	
Induced Travel Estimation				
Elasticity of demand relative to travel time		-0.28		
Initial VHT* at before expansion speeds		240.02		
Initial VHT after facility expansion		123.84		
Initial VHT savings		116.17		
Delay VHT per additional VMT**		0.05		
Induced travel		984		

*VHT = Vehicle hours of travel, i.e., travel time x traffic volume
** VMT = Vehicle miles of travel, i.e., traffic volume x distance

they are not made worse off with pricing. This means that the toll rate will be governed by the value of time of the tolled vehicle with the lowest value of time. This is used to estimate toll rates. Toll rates are estimated based on travel time saved by tolled vehicles (relative to the base case six-lane highway configuration), average value of time from the *Urban Mobility Report* (Texas Transportation Institute 2007), the distribution of the value of time of all freeway motorists under No Pricing (HDR Inc. 2009), and the percentage of vehicles that would be tolled off with pricing based on the difference between the total traffic on the highway with No Pricing from Table 3 (i.e., 7,914 vph) and the number of vehicles that can be accommodated on the FEE highway while meeting the 55 mph speed target (i.e., 7,069 vph).

- **Annual revenue**. For annual revenue estimation, it is assumed that toll-exempt vehicles average 200 passenger car equivalents per hour per lane. Congestion levels (and the estimated fees) are assumed to apply only during four hours per direction on each weekday, on 250 working weekdays per year. However, revenue estimates based on these four hours are increased by 25% to account for revenues during shoulder hours. Tolling will be needed during shoulder hours to manage traffic that shifts from the peak hours to avoid the higher tolls during those hours.
- **Greenhouse gas emissions**. To estimate greenhouse gas emissions, it is assumed that 10% of total fuel consumed during peak periods is diesel fuel, and 90% is gasoline. The following relationships (HDR Inc. 2009) are used:
    - o Fuel efficiency (miles per gallon) is estimated based on speed, using the relationship: 8.8 + 0.25* speed.
    - o CO2 emissions per gallon are estimated based on 8,788 grams per gallon of gasoline and 10,084 grams per gallon of diesel.

The results of the analyses for the two alternatives, for the first year after implementation, are presented in Tables 4 and 5 and summarized in Table 6. The results suggest that in the first year the FEE highway alternative will provide much higher user benefits, i.e., almost $38 million annually, vs. $31.25 million for No Pricing. It will also result in more greenhouse gas reductions, i.e., a 28% reduction in peak period emissions, vs. a 10% reduction for No Pricing. Long-term analysis results, also presented in Table 6, show more striking differences. They show that as traffic demand increases and throws the facility into breakdown flow conditions, user benefits from FEE highways would be about twice the user benefits from No Pricing, i.e., $66 million vs. $33 million annually. In the long term, the No Pricing alternative will *increase* peak period greenhouse gas emissions by 7%, vs. a 39% reduction for FEE highways.

Table 4: Other Impacts on a 10-mile 6-Lane Freeway Reconfigured to Four FEE Lanes per Direction

	Existing Configuration	Four FEE Lanes
User Benefits on 10-mile Facility		
Facility length (mile)	10.00	10.00
Total travel time on facility per mile (min)	14,401	7,711
Total delay reduced per mile (minutes)		6,980
Average value of time ($/hour)		$14.60
Value of delay savings/mile/peak hour/direction		$1,698.47
Number of directions	2	2
Peak hours per weekday for both directions	4	4
Number of weekdays per year	250	250
Value of delay savings per mile annually		$3,396,934
Value of annual delay savings for full facility		**$33,969,345**
Annual delay savings for 10-mile facility (hours)		2,326,667
Annual delay savings per 10-mile trip on the facility (hours)		41
Fuel cost per gallon		$2.50
Gallons of fuel saved annually		1,605,401
Value of fuel saved		**$4,013,501**
Total delay and fuel cost saved annually		**$37,982,846**
Toll Revenue on 10-mile Facility		
Travel time saved in priced lanes (min/mi)		0.99
Percent of traffic volume reduction relative to No Pricing		-12%
Percentile in WTP distribution		12%
Average toll rate per mile		$0.10
Average toll for 10-mile trip		$0.97
No. of exempt vehicles/hour (passenger car equivalents)		200
Factor to account for revenue during shoulder hours		1.25
Annual toll revenue		**$16,593,749**
Greenhouse gas emissions on 10-mile Facility		
Fuel consumption rate (mpg)	16	23
Annual fuel consumption (gallons)	8,652,619	6,269,372
CO2 emissions per gallon	8,918	8,918
CO2 emissions annually (metric tons)	77,161	55,908
Percent change with 4 FEE lanes		**-28%**

Table 5: Other Impacts on a 10-mile 6-Lane Freeway Reconfigured into Four Free Lanes with No Pricing

	Existing Configuration	Four Free Lanes with No Pricing
User Benefits on 10-mile Facility		
Facility length (mile)	10.00	10.00
Total travel time on facility per mile (min)	14,401	10,323
Total delay reduced per mile (minutes)		5,742
Average value of time ($/hour)		$14.60
Value of delay savings/mile/peak hour/direction		$1,397.18
Number of directions	2	2
Peak hours per weekday for both directions	4	4
Number of weekdays per year	250	250
Value of delay savings per mile annually		$2,794,363
Value of annual delay savings for full facility		**$27,943,633**
Annual delay savings for 10-mile facility (hours)		1,913,947
Annual delay savings per 10-mile trip on the facility hours)		30
Fuel cost per gallon		$2.50
Gallons of fuel saved annually		1,320,624
Value of fuel saved		**$3,301,559**
Total delay and fuel cost saved annually		**$31,245,192**
Greenhouse gas emissions on 10-mile Facility		
Fuel consumption rate (mpg)	16	20
Annual fuel consumption (gallons)	8,652,619	7,797,092
CO2 emissions per gallon	8,918	8,918
CO2 emissions annually (metric tons)	77,161	69,531
Percent change with four Free Lanes		**-10%**

Table 6: Summary of Short-Term and Long-Term Impacts of Alternative Reconfigurations of a 10-Mile Six-Lane Freeway

Short-Term	Four FEE Lanes	No Pricing
User Benefits		
Annual delay reduced per 10-mile trip (hrs.)	41	30
Annual value of user benefits ($millions)	$37.98	$31.25
Revenues		
Charge for 10-mile trip ($)	$0.97	$0.00
Annual revenues ($millions)	$16.59	$0.00
Sustainability		
Change in peak period CO2 emissions (%)	-28%	-10%
Long-Term		
User Benefits		
Annual delay reduced per 10-mile trip (hrs.)	71	29
Annual value of user benefits ($millions)	$65.93	$33.65
Revenues		
Charge for 10-mile trip ($)	$1.97	$0.00
Annual revenues ($millions)	$33.78	$0.00
Sustainability		
Change in peak period CO2 emissions (%)	-39%	7%

REGIONWIDE APPLICATION OF THE FEE HIGHWAY CONCEPT

A typical large metropolitan area such as Seattle, WA, or Washington, DC, has a freeway network of about 300 miles or 1,800 lane miles. Table 7 shows results from an extrapolation of the impacts estimated for a 10-mile highway to a 300-mile highway network. Preliminary cost estimates for conversion of standard highways to FEE highways, including tolling and active traffic management, have been prepared for a study in progress by Booz Allen Hamilton for the Federal Highway Administration. Based on this study, annualized capital and operating costs (in real dollars) are estimated at about $500 million for the No Pricing alternative and $610 million for the FEE highway alternative.

Estimated user benefits, shown in Table 7, for both alternatives exceed these costs by 87% in the first year (short term), but benefits increase by significant amounts in the long term with growth in demand. User benefits in the long term are almost $2 billion annually for FEE highways, resulting in a benefit-cost ratio of 3.24. For No Pricing, the comparable user benefits are over $1 billion, resulting in a benefit-cost ratio exceeding 2.02.

Table 7 suggests that, for a 300-mile congested network, the FEE highway concept could generate almost $0.5 billion in toll revenue in the first year after implementation. In the longer term, annual revenues could be as much as $1 billion annually. All revenue estimates are in real dollars, and the long-term revenue estimates are conservative because the value of time parameter was not adjusted to account for the increases in real wage rates that can be expected over time. The No Pricing alternative generates no toll revenue. Revenues for the 300-mile FEE highway network concept fall short of costs in the first year by about $110 million (i.e., $610 million in costs vs. about $0.5 billion in toll revenue), but estimated long-term revenues are more than adequate to cover the annualized costs, generating over $400 million in toll revenue surpluses annually (i.e., $610 million in costs vs. about $1.01 billion in revenue). The stream of surpluses could be leveraged to fund multimodal transportation investments and other strategies to benefit the tolled off and the untolled.

Table 7: Summary of Short-Term and Long-Term Impacts of Alternative Reconfigurations of a Regional Network of Freeways

Short-Term	**Four FEE Lanes**	**No Pricing**
Number of miles of freeway (avg. 6 lanes)	300	300
Annual value of user benefits ($millions)	$1,139	$937
Annual revenues ($millions)	$498	$0
Annual costs ($millions)	$610	$500
Benefit/cost ratio	1.87	1.87
Annual surplus (+) or deficit (-)	($112.19)	($500.00)
Long-Term		
Number of miles of freeway (avg. 6 lanes)	300	300
Annual value of user benefits ($millions)	$1,978	$1,010
Annual revenues ($millions)	$1,013	$0
Annual costs ($millions)	$610	$500
Benefit/cost ratio	3.24	2.02
Annual surplus (+) or deficit (-)	$403.39	($500.00)

ADDRESSING THE ISSUES

Safety

A summary of research on the safety impacts of narrower lanes is provided by Ng and Small (2009). The authors find that both theoretical and empirical evidence linking narrow lane design to safety are ambiguous. According to the authors, it is an open question whether narrow lanes would in fact reduce safety. They suggest that it would depend on factors that vary from case to case, especially the speeds chosen by drivers. They cite several innovations in Europe that offer hope that roads designed for lower speeds could be accompanied by measures to ensure that lower speeds in fact prevail. For example, they point out that variable speed limits have been used for many years in Germany and the Netherlands, and recently in Copenhagen, primarily to smooth traffic during the onset of congestion—but also with a strong reduction in injury accidents in one German implementation.

A study by the Midwest Research Institute (Bauer et al. 2004) suggests that the observed increases in accident frequency when freeways in the United States have been re-striped to narrower lanes cannot necessarily be attributed to the use of narrower lanes or the conversion of a shoulder to a travel lane. The authors suggest that the use of the added narrow lanes as HOV lanes introduces a difference in speed between adjacent lanes, which may be another explanation for the increase in accidents. The FEE highway concept eliminates these speed differentials, since vehicles in all lanes travel at essentially the same free-flow speed when all lanes are priced.

Nonetheless, the use of shoulders as travel lanes does present safety challenges, which need to be addressed. For example, protocols will need to be developed for emergency response when the shoulder is being used as a travel lane. The first project in the United States to use a shoulder as travel lane "dynamically" during peak periods was implemented on I-35W in Minneapolis in October 2009, and the United States will learn from that experiment.

Impacts on Alternative Free Routes

The FEE highway's capacity will be higher than the existing highway's capacity, due to the extra FEE lane. If the FEE highway were available for use *free of charge* during peak periods, it could be expected that some drivers who currently drive on parallel facilities to avoid congestion on the highway would shift to the FEE highway, reducing traffic on the parallel facilities. Peak period fees charged on the FEE highway will need to be high enough to keep excessive traffic away and maintain free flow of traffic on the expanded capacity. However, fees must also be low enough to ensure that the FEE highway is fully utilized. This will ensure that net diversion to alternative free routes does not occur.

Some drivers who do not value their time very highly (and have low willingness to pay) are likely to shift to free parallel surface facilities. However, this shift in traffic to free facilities is likely to be counterbalanced by shifts in traffic from free facilities to the FEE highway by those who value the time savings more highly than the fees charged on the FEE highway.

A redistribution of traffic will occur based on driver values of time, but not necessarily a significant increase in traffic either on the FEE highway or on parallel free facilities. Only drivers who have very low values of time will be tolled off. This does pose an equity issue with regard to low-income drivers, and is discussed later in this paper (see Public Acceptance Issues).

Overall Congestion Reduction Impacts

In a typical large city, VMT on limited access highways tends to exceed VMT on arterials. For example, Seattle's 2007 daily VMT on limited access highways was 30.6 million, vs. only 27.1 million for arterials (Texas Transportation Institute 2009). However, a major portion of the delay faced by travelers tends to be on limited access highways, so FEE highways could alleviate a majority of the recurring congestion delay in metropolitan areas.

The advantage of limiting pricing to limited access highways is that it can be implemented in the short term using existing proven transponder-based technologies which are already widely deployed in the United States. Recent research under the National Cooperative Highway Research Program (Sorenson et al. 2009) explored other technology that could potentially be used to collect ubiquitous user charges. When they are further developed and deployed in the future, they could potentially be used to complement the FEE highway concept by deploying pricing on roads off the FEE highway system.

Public Concerns About Fairness

Imposing new charges on existing toll-free roads raises fairness concerns. Many people believe they have already paid for existing roads with their taxes, and new charges would amount to double taxation. However, experience shows that if the public understands the high costs for reconstruction, rehabilitation, and expansion of highway facilities, they may accept congestion pricing as one way to help pay for these costs while receiving improved mobility from congestion relief. The SR 520 floating bridge in the Seattle metropolitan area will become the first existing toll-free facility in the United States to charge new tolls that will vary by time of day to achieve performance targets. The public in Seattle understands the high costs to replace the SR 520 bridge, and that tax dollars will be insufficient to pay for these costs.

Public opinion surveys suggest that the public is more accepting of new charges if they are convinced that revenues from new charges will go directly towards providing improved transportation service on the facilities on which the charges are imposed. The public accepts toll roads because they can see that the revenues go to support the specific facility on which they pay tolls (Zmud and Arce 2008).

The public will need to understand that tolls would only be charged during peak hours with FEE highways – unlike regular toll roads. If road users understand that FEE highways are not regular toll roads, they could be convinced that the new charges will be a just and fair "fee for service" that goes to provide improved peak period service made possible by the extra physical capacity being provided. In South Africa's Johannesburg/Pretoria metropolitan area, 115 route miles of existing freeway are being widened and an additional 236 route miles are being upgraded, all of which will be funded through new electronically charged tolls on the existing free system (Poole 2009). This demonstrates that concerns about "double taxation" can be alleviated when the new charges are directly related to the extra benefits being provided.

Concerns About Equity for Low-Income Drivers

Due to the increase in capacity with FEE highways, few motorists (if any) would need to be "tolled off' – so tolls would be relatively low and quite affordable. Like vehicle property taxes in some states, tolling could be made more progressive by charging toll rates that vary based on value of the vehicle being tolled. Yet valid concerns may be expressed that the new fees may make it too expensive for low-skilled workers to get to their jobs, because low-skilled jobs are often not well served by public transit. Also, low-income workers tend to have jobs with fixed schedules, and pricing may be particularly unfair to those with less flexible work schedules since they are unable to shift their time of travel.

A well-designed FEE highway pricing strategy can be less burdensome to low-income citizens with appropriate use of toll revenue. With a portion of the toll revenue dedicated to transit, low-income transit riders can benefit significantly from toll-financed transit improvements. Policymakers can also include protections for low-income individuals, such as "life-line" credits or toll discounts, or reimbursements for tolls paid. For example, a proposal for congestion pricing in New York City by its mayor in 2008 would have provided reimbursements to low-income individuals qualifying for the federal earned income credit. They would be reimbursed for any charges in excess of the fare for a transit trip.

Concerns About Administrative Costs

Concerns are often expressed about the high administrative costs for implementing and operating a congestion pricing scheme relative to other ways of generating revenue from highway users. Operating costs amount to about 20% of revenue for a normal toll facility, based on a study in progress for the National Cooperative Highway Research Program (NCHRP project 19-08). By comparison, collection of fuel taxes costs only about 1% of revenue.

Yet congestion pricing may be justified despite the high costs, because costs for implementation and operation are significantly exceeded by the benefits. The purpose of congestion pricing is not simply to collect revenue, but also to achieve other economic, social, and environmental goals. If the same amount of revenue as generated from congestion pricing were instead to be raised by fuel taxes, costs for revenue collection would indeed be lower, but the loss of benefits would be far greater than the cost savings. For decision making, the costs for implementation and operation of FEE highways must be compared with benefits, not with revenues collected. As Table 7 shows, the value of incremental user benefits alone (i.e., time and fuel cost savings) would far exceed incremental costs for implementation and operation of FEE highways relative to the No Pricing alternative. Long-term incremental user benefits for a 300-mile FEE highway network relative to a No Pricing network with the same configuration are almost $1 billion annually (i.e., $1.9 billion minus $1.0 billion), vs. incremental costs of only $110 million annually (i.e., $610 million minus $500 million).

Synergistic combinations of FEE highways and supporting strategies should be subjected to a comparative benefit-cost or cost-effectiveness evaluation and compared with other traditionally funded alternatives. The long-range transportation planning process undertaken by metropolitan planning organizations (MPOs) could be used to inform the public about costs and benefits of alternative approaches. This would begin a discussion about the tradeoffs between conventional transportation investment approaches and approaches involving congestion pricing. The MPO in Seattle is showing the way in the United States. Five synergistic pricing alternatives were recently analyzed and presented for consideration by the public for the Year 2040 Regional Transportation Plan for the Seattle area (Puget Sound Regional Council 2009). This led to adoption by policymakers of a fully-priced highway network as part of the Year 2040 Regional Transportation Plan in May 2010.

Concerns About Privacy

Proposals to price the entire network of highways and surface streets using global positioning systems (GPS) have raised significant privacy concerns. Even if the detailed information on travel never leaves the vehicle (as proposed with some concepts), there is a concern that information stored electronically on in-vehicle units is insecure and may become available to unauthorized persons. However, privacy does not appear to be a major concern with pricing of *only* the limited access highway network using transponder-based toll collection technology as envisioned for FEE highways. In a focus group study on such a strategy (Petrella, Biernbaum, and Lappin 2008), privacy was not an issue that resonated strongly or generated much discussion with most participants, perhaps because vehicle identification technology would be restricted to limited access highways only, and information on trip origins and destinations off the tolled system would not be collected.

All operating HOT lane projects in the United States and more than 250 other toll facilities across the country use electronic toll collection (ETC). Yet, some people continue to be concerned about the privacy impacts of ETC. In response, tolling agencies have devised a method to protect the public's privacy by linking the transponder and the driver's personal information with a generic, internal account number that does not reveal the driver's identity and that is not disclosed to other organizations. In many cases, a motorist can open an anonymous pre-paid account if he or she so chooses, although there have been very few who have signed up. Singapore has alleviated privacy concerns by collecting tolls using smart cards with stored value that are inserted into an in-vehicle unit. These "electronic purses" are replenishable at Automated Teller Machines (ATMs), and may also be used for other purchases unrelated to tolling. However, the in-vehicle units needed for such systems are far more expensive than the typical transponders used in the United States.

SUMMARY AND CONCLUSIONS

A sketch-level analysis of the FEE highway concept presented in this paper suggests that it could prove to be a financially feasible way to create a high-performance metropolitan transportation system. However, more detailed assessments of the opportunities and options in specific metropolitan areas will be needed, and safety challenges will need to be addressed. The intent of this paper is to stimulate further exploration and discussion of the concept and to generate other ideas.

References

Bauer, Karin M., Douglas W. Harwood, Warren E. Hughes, and Karen R. Richard. "Safety Effects of Narrow Lanes and Shoulder-Use Lanes to Increase Capacity of Urban Freeways." *Transportation Research Record 1897.* Transportation Research Board, National Research Council, (2004): 71-80.

DeCorla-Souza, Patrick and Harry Cohen. "Accounting for Induced Travel in Evaluation of Urban Highway Expansion." *Transportation* 24(2), (1997): 107-123. Http://www.fhwa.dot.gov/steam/doc.htm.

DeCorla-Souza, Patrick. "Implementing Congestion Pricing on Metropolitan Highway Networks with Self-Financing Public-Private Partnerships." *Journal of the Transportation Research Forum* 45(1), (2006): 5-22.

DeCorla-Souza, Patrick. "A Freeway Congestion Pricing Paradox." Presented at the Transportation Research Board Annual Meeting, Washington, D.C., 2010.

Dupuit, J. "On the Measurement of the Utility of Public Works." *Annales des Ponts et Chaussses*, 2nd series, 8, 1844, reprinted in *Transport*, edited by D. Munby, (1968): 19-57.

Federal Highway Administration. *2006 Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance. Appendix A: Highway Investment Analysis Methodology.* HERS Highway Improvement Cost Inputs, 2007. Http://www.fhwa.dot.gov /policy/2006cpr/appa.htm.

Federal Highway Administration. *Current Toll Road Activity in the U.S.: A Survey and Analysis*, 2009. Http://www.fhwa.dot.gov/ipd/pdfs/2008_toll_activity_white_ paper.pdf.

Goodwin, P. "The Rule of Three: A Possible Solution to the Political Problem of Competing Objectives for Road Pricing." *Transport Engineering and Control* 30, (1989): 495-497.

Hau, Timothy. "Economic Fundamentals of Road Pricing: A Diagrammatic Analysis, Part 1 – Fundamentals." *Transportmetrica* 1(2), (2005): 81-117.

HDR, Inc. *TRUCE-ST User's Guide*. Prepared for the Federal Highway Administration, 2009. Http://ops.fhwa.dot.gov/tolling_pricing/value_pricing/tools/index.htm.

Knight, F. H. "Some Fallacies in the Interpretation of Social Cost." *Quarterly Journal of Economics* 38, (1924): 582-606.

Lee, Douglass B., Jr. "Toward the Evaluation of Value Pricing." *Transportation Research Record No. 2079,* Transportation Research Board, (2008): 71-88.

Merugu, Deepak, Balaji S. Prabhakar, and N.S. Rama. "An Incentive Mechanism for Decongesting the Roads: A Pilot Program in Bangalore." Http://netecon.seas.harvard.edu/NetEcon09/Papers/Merugu_09.pdf.

Minett, Paul. "Enabling a Flexible Ridesharing Response to Congestion Pricing." Presentation at the Innovations in Pricing of Transportation Systems Conference & Workshop, Orlando, Florida, May 2010a. Http://www.flexiblecarpooling.org /fcdecongestionpricing.pdf.

Minett, Paul. "Not Only Seeing, But Also Knowing Who is in the HOV." Presentation at the Innovations in Pricing of Transportation Systems Conference & Workshop, Orlando, Florida, May 2010b. Http://www.flexiblecarpooling.org /fcdecongestionpricingtechnology.pdf.

Ministry of Transport. *Road Pricing: the Economic and Technical Possibilities (the Smeed Report)*, 1964.

Nakamura, K. and K. Kockelman. "Congestion Pricing and Roadspace Rationing: An Application to the San Francisco Bay Bridge Corridor." *Transportation Research* 36A (5), (2002): 403-417.

Ng, Chen Feng and Kenneth Small. "Tradeoffs Among Free-flow Speed, Capacity, Cost, and Environmental Footprint in Highway Design, 2009." Http://www.economics.uci.edu/docs/2008-09/small-04.pdf.

Petrella, Margaret, Lee Biernbaum, and Jane Lappin. "Exploring a New Congestion Pricing Concept: Focus Group Findings from Northern Virginia and Philadelphia." Volpe National Transportation Systems Center, Cambridge, Massachusetts, 2008. Http://ops.fhwa.dot.gov/tolling_pricing/resources/report/cpcfocusgrp/congestion_focus_grp.pdf.

Pigou, A. *The Economics of Welfare*. Macmillan, London, 1920.

Poole, Robert. "Surface Transportation Innovations." *Reason Foundation*. Issue No. 74, Los Angeles, California, 2009.

Puget Sound Regional Council. "Transportation 2040: Draft Environmental Impact Statement, 2009." Http://psrc.org/projects/trans2040/deis/index.htm.

Small, K. "Using the Revenues from Congestion Pricing." *Transportation* 19, (1992): 359-381.

Sorensen, Paul, Liisa Ecola, Martin Wachs, Max Donath, Lee Munnich, and Betty Serian. National Cooperative Highway Research Program (NCHRP) Web-Only Document 143: "Implementable Strategies for Shifting to Direct Usage-Based Charges for Transportation Funding." Transportation Research Board, Washington, D.C., 2009.

Texas Transportation Institute. *2007 Urban Mobility Report: Appendix A*. Texas A & M University, College Station, Texas, 2007.

Texas Transportation Institute. *2009 Urban Mobility Report.* Texas A & M University, College Station, Texas, 2009. Http://mobility.tamu.edu/ums/congestion_data/tables /seattle.pdf.

Transportation Research Board. *Highway Capacity Manual*. National Research Council, Washington, D.C., 2000.

Vickrey, W. S. "Congestion Charges and Welfare: Some Answers to Sharp's Doubts." *Journal of Transport Economics and Policy* 95, (1968):107-118.

Walters, A. A. "The Theory and Measurement of Private and Social Cost of Highway Congestion." *Econometrica* 29, (1961): 676-699.

Zmud, Johanna and Carlos Arce. "Compilation of Public Opinion Data on Tolls and Road Pricing." NCHRP Synthesis 377. Transportation Research Board, 2008. Http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_377.pdf.

***Patrick DeCorla-Souza** is the Tolling and Pricing Program manager at the U.S. Department of Transportation's Federal Highway Administration. In this role, Patrick coordinates and oversees road pricing research and initiatives across the United States, and works with public and private sector partners nationwide to implement innovative road pricing strategies. Patrick has master's degrees in transportation planning and civil engineering, and co-chairs the Transportation Research Board's Congestion Pricing Committee.*

Productivity and Deregulation in the Canadian For-Hire Trucking Industry, 1985-1996

by James Nolan, Pamela Ritchie, and John Rowcraft

Many policy makers today take for granted that economic deregulation yields net benefits to industry, particularly with respect to helping improve productivity and efficiency. Certain structural issues about industrial deregulation are not well understood, including the impact of technical change or innovation. In this paper, we measure the extent to which innovation affected productivity within the Canadian for-hire trucking industry during deregulation. Using a unique data set spanning the trucking deregulation era in Canada, non-parametric Malmquist productivity indices are computed and de-composed. Results of the analysis indicate that the industry was affected by deregulation in a manner consistent with public policy goals.

INTRODUCTION

Policy makers today take for granted the idea that economic deregulation yields net benefits to industry, particularly with respect to helping improve firm level productivity and efficiency. Productivity improvements in deregulated industries can often be attributed to the fact that both incumbents and entrants must seek and implement technological innovations in order to survive and prosper. This paper will illustrate the degree to which innovation plays a part in productivity change in a large, deregulated transportation industry. To achieve this goal, a unique dataset from the Canadian for-hire trucking industry is used to compute a non-parametric industry level productivity index through the time of deregulation. This index is subsequently decomposed to identify the sources of productivity change in the industry.

Transportation is a vital component of the Canadian economy, and by value, trucking continues to be the most important commercial mode of transportation. As an industry, trucking can be separated into two primary sectors. These are i) the private sector, where trucking happens to be an integrated part of firm activity, and ii) the for-hire sector, comprised exclusively of carriers who only provide transportation services. These sectors are not mutually exclusive; large shippers often use both private and for-hire carriers to provide transportation services (Wood and Johnson 1996). Whereas private trucking accounts for slightly more total transportation activity than for-hire trucking in Canada, the for-hire sector is still critical to the economy. While total trucking movement in Canada roughly tripled from 1987 to 1997 (Transport Canada 1999), for-hire trucking grew at a remarkable rate after deregulation. In the decade following deregulation, total for-hire trucking revenues doubled, while approximately 14% of for-hire trucking firms had revenues in excess of $1 million per year (Nix 1996).

The following section contains a brief institutional and historical review of the Canadian trucking industry. Then the relationship between regulation and productivity is described. A brief descriptive theory section highlights why productivity changes are crucial to the study of deregulation. Next, a detailed account of the methodology used in the study is presented. After a discussion of the empirical results, a brief set of conclusions ends the paper.

REGULATION AND THE TRUCKING INDUSTRY IN CANADA

The development and evolution of the Canadian for-hire trucking industry relates closely to its counterpart in the United States. The industry in each country conducts much of its operations as part of the complete North American market. From a regulatory perspective, they have been treated historically in a similar manner by their respective governments.

The policy of economic regulation in the U.S. trucking industry, which took the form of entry restrictions and rate setting, began as far back as the 1930s. Political pressure from rail interests forced increased regulation of the trucking industry because the rail freight industry was itself heavily regulated at that time. Thus, some authors have concluded that the U.S. trucking industry was regulated solely in the interest of creating a "level playing field" for all firms and modes operating in the freight transportation market (Viscusi et al. 1996).

Trucking regulation in Canada was also implemented as a response to pressure from the regulated Canadian rail companies, but similar regulatory policies also permitted easier access to the important U.S. freight market (Woudsma and Kanaroglou 1994). In both countries, the primary impetus for regulating the for-hire trucking industry was not due to typical market problems like excessive concentration or conditions of natural monopoly. In spite of this unusual situation, regulatory policy towards the trucking industry in Canada and the United States remained unchanged for almost 50 years.

Economic regulation in the Canadian trucking industry most often took the form of restrictions on entry by new firms (Bonsor 1995). Certain special freight rates were also pre-determined by regulatory fiat. This prolonged policy of regulatory protection of almost every trucking market against entry certainly weakened incentives for firms to innovate in the provision of trucking services.[1] But by 1980, once the U.S. rail industry had been partially deregulated, the U.S. trucking industry was fully deregulated as well. Subsequently, the trucking industry in Canada followed suit through the early 1980s.[2] By 1988, domestic trucking markets and routes in Canada were no longer protected by law against entry (Dionne et al. 1998). Most analysts agree that by 1993, for-hire trucking deregulation in Canada was complete.

In the years since deregulation, the Canadian trucking industry has continued to undergo substantial change. By way of example, 38 of the top 100 trucking companies (measured by total revenues) operating in 1990 were no longer operating by 1995, most often transformed or subsumed through buy-outs or mergers. [3,4] But evidence from the latter part of this sample shows that the importance of larger carriers in the industry declined, and the for-hire industry itself became less concentrated after deregulation (Transport Canada 1996). However, some industry observers believe that for-hire trucking deregulation in Canada has been characterized by extreme instability in prices and output (Nix 1996). In summary, it can be very difficult to judge the merits of deregulation by simple indicators of market performance. In fact, the very nature of the delivered good in trucking has changed due to several major technological changes that occurred during deregulation.

The most obvious example of technological change in the trucking industry is the development and use of satellite/Global Positioning System (GPS) tracking systems, begun in the late 1980s. Among the largest carriers in particular, GPS has replaced short-range dispatch systems to ensure that delivery occurs when and where desired. The spread of this particular innovation has been driven by increased use of just-in-time logistics with precise delivery windows. Without a doubt, innovations of this sort have substantially altered the freight transportation industry, both rail and trucking, throughout North America.

If we define productivity as the ability of a firm to produce outputs from a given set of inputs over time, then policy shocks such as deregulation affect productivity at the industry level by changing behavior at the firm level. In this research, the following questions are addressed: i) Are industry level productivity changes through a transition to deregulation different from those that can be expected to occur in a regulated industry, and ii) in what manner has this particular industry reacted to the process of deregulation? In the next section, the first question is addressed along with a description of how productivity changes within an industry constitute a litmus test as to the "appropriateness" of deregulation for the industry.

DEREGULATION AND PRODUCTIVITY

Trucking companies in Canada did not need to be particularly innovative or even economically efficient in order to survive in the regulatory era. Heavy-handed and long-term regulation in Canada meant that trucking firms possessed few incentives to innovate in order to gain competitive advantage (Viscusi et al. 1996). As might be expected of firms under such regulation, there is strong evidence in Canada that considerable resources were devoted to entry seeking and forestalling (Bonsor 1995).

The fact that this industry survived under such conditions does not imply that firms were productively efficient in an absolute sense. But without loss of generality, it can be assumed that as the regulatory era drew to a close, the largest Canadian for-hire trucking firms operated with similar levels of productive efficiency. This was simply due to the fact that throughout the era of regulation, trucking companies had plenty of time to discover those procedures and processes required to efficiently produce output within the constraints (for example, inflexible route structures and rates) inherent in the regulatory regime.

Under these conditions, it is likely that a relative lack of innovation in production technology existed in the regulated Canadian trucking industry compared with what would have been the case under a competitive environment characterized by frequent entry and exit of firms. Without foreseeable changes in market risk or profit levels, trucking firms in Canada maintained the status quo with respect to industry structure. Such a stylized description of events also implies that the "best practice" technology in regulated for-hire trucking was well established and that all firms producing the same services would likely employ similar technologies. Thus, by the time of deregulation, for-hire trucking was a mature industry in Canada characterized by stable prices and output levels with few differences among the operating procedures of firms.

Now consider the nature of productivity changes in this regulated industry. In the prolonged period prior to regulatory change, it is very likely that the dominant component of temporal productivity change for any individual trucking firm consisted of *emulating* those technological processes that were already well established. In the productivity measurement literature, this type of productivity change is referred to as a change in *technical efficiency* (Lovell 1993). More formally, the technical efficiency effect is defined simply as a change in efficiency relative to a static industry production possibility frontier.

Economists assert that industrial deregulation generates numerous benefits. But aside from demand side considerations, many authors argue it is extremely important that open entry also stimulate technical innovation in industry (Scherer and Ross 1990; Gallamore 1999). Therefore in contrast to the situation that is argued to exist under prolonged regulation, it can be argued that changes in productivity over time through the period of deregulation are dominated (at least) initially by shifts in technology. This is known as technical change in the productivity literature. As opposed to the technical efficiency effect, technical change is defined as a shift in the level of best practice technology represented by the industry production possibility frontier. Among incumbent firms in the industry, those that survive the shock of deregulation will do so because they innovate and adapt to the new market situation.[5]

Using these basic arguments as a foundation for evaluating the merits of de-regulation, sources of year-to-year productivity change among the largest firms in the Canadian trucking industry during the window of deregulation are identified by computing and decomposing a non-parametric Malmquist productivity index for these firms. To the best of our knowledge the Canadian data employed here are novel, but the methodology mirrors the work of McMullen and Okuyama (2000) measuring and interpreting productivity changes in the U.S. trucking sector.

THE MEASUREMENT OF TECHNICAL EFFICIENCY AND PRODUCTIVITY

To evaluate non-parametric productivity changes in the Canadian for-hire trucking industry, static technical efficiency is measured using data envelopment analysis (DEA). DEA is a non-parametric efficiency estimation methodology consisting of a series of specially configured linear programs. DEA is a useful analytic tool to measure efficiency because it does not require any prior assumptions about the structure of production technology in the industry. This flexibility is important if the researcher is somewhat ignorant about the structural nature of the production technology in question (Seiford and Thrall 1990). It would be imprudent to make any a priori assumptions about the structure of technology in this industry, especially considering the significant technological changes that most certainly occurred during and after deregulation.

In this paper, output oriented DEA scores are computed. Output-oriented DEA generates a hypothetical measure of how much a firm's outputs could have been proportionately increased, with inputs held fixed, such that it transformed inputs into outputs as efficiently as the best firms in the sample. Thus, output oriented DEA efficiency scores fall in a range between 1 and infinity, where an efficiency score of 1 means that the firm is technically efficient (or equivalently, the datum falls on the production possibility frontier). An efficiency score greater than 1 indicates that a firm could theoretically increase outputs by a factor of *{efficiency score - 1}* percent to produce as efficiently as the best practice firms in the sample. The linear program used to compute the efficiency scores (the basic output oriented, constant returns to scale version of DEA) is the following:

$$
(1)\quad \begin{aligned}
\text{Min}_{u,v}\quad & \nu X_1 \\
\text{s.t.}\quad & \mu Y_1 = 1 \\
& -\mu Y + \nu X \geq 0 \\
& \mu \geq 0,\ \nu \geq 0
\end{aligned}
$$

Here, Y and X in the second constraint in (1) represent vectors of outputs and inputs respectively across the entire sample, whereas X_1 and Y_1 are those inputs and outputs unique to the particular firm under examination. Also, μ and ν are vectors of so-called "virtual" multipliers, and are coefficients necessary to solve the linear program.[6] The linear program is solved once for each observation in the sample (with the non-negativity and second constraint in (1) being common to each LP) and yields the technical efficiency measure for each firm.

By definition, the reference data set of firms used to calculate the relative efficiency scores (represented by the second constraint in (1)) consists of all other firms operating in the same time period. Therefore, the level of industry technology in a given year is defined exclusively by firms in the dataset at that time, and not on a theoretical level dependent upon a chosen functional form of production.

Table 4 shows the average DEA scores for all trucking firms in the sample spanning 1985-1997. The average scores fluctuate somewhat over time, as might be expected given the policy changes in the industry. Note that the highest (least average efficiency) score occurs in 1994 and the lowest (highest average efficiency) occurs in 1997. These efficiency scores are subsequently used as inputs into the computation of the productivity indices.

THE MALMQUIST INDEX

The productivity index used in this paper is the Malmquist index. The index facilitates identification of temporal sources (or components) of productivity change. These components are: i) technical efficiency, or the convergence of technology among the firms in the sample over time towards the best practice frontier, and ii) technical change, or shifts in the production frontier due to changes in the overall level of technology used by firms in the industry.

In terms of the technical efficiency scores computed with DEA, the Malmquist index can be written in the following manner (Lovell 1993):

$$(2)\quad M_0^{t+1}(x^{t+1},y^{t+1},x^t,y^t) = \left[\frac{D_0^t(x^{t+1},y^{t+1})}{D_0^t(x^t,y^t)}\ \frac{D_0^{t+1}(x^{t+1},y^{t+1})}{D_0^{t+1}(x^t,y^t)}\right]^{1/2}$$

And this equation can be decomposed into the following components:

$$(3)\quad M_0^{t+1}\ x^{t+1},y^{t+1},x^t,y^t) = \frac{D_0^{t+1}(x^{t+1},y^{t+1})}{D_0^t(x^t,y^t)}\left[\frac{D_0^t(x^{t+1},y^{t+1})}{D_0^{t+1}(x^{t+1},y^{t+1})}\ \frac{D_0^t(x^t,y^t)}{D_0^{t+1}(x^t,y^t)}\right]^{1/2}$$

In this notation, D_o is a standard output distance function (Shephard 1953), which is simply the inverse of the output oriented DEA technical efficiency scores. This allows for a straightforward computation of the Malmquist index under the chosen DEA specification. The superscripts on D_o indicate the reference time period (or technology) with which the scores are calculated. The superscripts on **x** (inputs) and **y** (outputs) indicate the time period of the data used in the index calculation. For example, if the superscripts for each distance function are the same, the efficiency score is calculated in the standard way using the data of other firms operating in the same time period. However, if the superscripts in the distance functions are not the same (i.e., t vs. t+1), the DEA efficiency score for an observation is computed instead by using, as the reference set, industry data from the adjacent time period.

The first term on the right-hand side in equation (3) measures the change in technical efficiency between time periods, or the change in the DEA score from period *t* to period *t+1*. The second term measures pure technical change over time by computing modified DEA scores for a given firm in a particular year relative to the technology sets (the data on inputs and outputs) of both the previous ($D_o^t(x^{t+1}, y^{t+1})$) and succeeding (($D_o^{t+1}(x^t, y^t)$) years.

Intuitively, the Malmquist index compares the level of industry technology in the current year with those from previous and future years. The changes in the technical efficiency scores induced by changing the reference technology set are used to examine whether or not individual firms contribute either to the shifting of the production possibility frontier, or if firms are simply "catching up" with the prevailing level of technology.

As in McMullen and Okuyama (2000), the decomposition of the industry Malmquist index will be interpreted in the following manner. The technical efficiency term is called the "catching up" term (see also Fare et al. 1994), and measures the effect on productivity of firms catching up to the performance of the most efficient firms in the sample.

This term:

- equals 1 when no change in technical efficiency has occurred;
- is less than 1 when technical efficiency has declined over time;
- is greater than 1 when technical efficiency has increased over time.

The second term in the Malmquist decomposition is often referred to as the technical change or "innovation" term, measuring how much overall productivity change is due to the efficiency frontier shifting from one time period to the next. If this term dominates the index, the level of best practice technology in the industry has changed and the best performing firms in the industry became more efficient. This term:

- equals 1 when no technical change has occurred;
- is less than 1 when there has been regressive technical change;
- is greater than 1 when there has been progressive technical change.

The Malmquist index itself is interpreted in the same manner as the individual terms. Thus the index:

- equals 1 when no change in productivity has occurred;
- is less than 1 when productivity has declined over time;
- is greater than 1 when productivity has increased over time.

By using this decomposition, the structure of productivity changes among this set of firms during deregulation in the Canadian for-hire trucking industry can be identified. And since average indices of productivity and decompositions are computed, the results can be interpreted with respect to the entire sample on aggregate.

DATA

The sample comprises a set of the largest for-hire trucking firms operating across Canada.[7] The data were compiled from several sources, including the Motor Carriers of Freight (MCF) survey conducted by Statistics Canada (1993).[8] Even though there are more Class I firms in Canada (about 350 in 1993) than in the sample, only about one-third of these firms were actually surveyed in any given year (Statistics Canada 1993). The sub-sample consists of those firms for which reliable data were available from the MCF survey through the entire period of trucking deregulation.[9]

The empirical analysis is composed of two parts. First, DEA scores and Malmquist indices are computed using a balanced panel of 14 trucking firms over 12 years, from 1985 to 1996. Balanced panels are best employed in the computation of Malmquist indices so that reference technology sets do not change; this ensures consistency of the results between specific firms (Lovell 1993). But since only overall productivity results are of interest here and not results at the firm level, such consistency is not as critical. Subsequently, the second portion of this study computes Malmquist indices for a slightly larger unbalanced panel of 21 firms over 13 years, from 1985-1997. See Table 1 for a comparison of sample sizes for the balanced and unbalanced panels.

Table 1: Sample Sizes Used for the Balanced and Unbalanced Panels

Year	**85**	**86**	**87**	**88**	**89**	**90**	**91**	**92**	**93**	**94**	**95**	**96**	**97**
Balanced	12	12	12	12	12	12	12	12	12	12	12	12	-
Unbalanced	19	19	19	20	20	19	20	20	20	17	18	18	15

Table 2 provides summary statistics for the input and output variables for the full (unbalanced) sample, whereas Table 3 provides the summary statistics for the balanced panel estimates. With respect to variable selection, non-parametric studies of productivity and efficiency of transportation modes have used similar variables to those chosen here (Bookbinder and Qu 1993; McMullen and Okuyama 2000). This set of input and output variables constitutes a parsimonious but reasonable representation of the structure of production in the trucking industry.

Table 2: Full (Unbalanced) Sample Descriptive Statistics, 1985-1997

Inputs	**Number of employees**	**Total vehicles**	**Fuel Consumption (liters)**
Mean	668.87	1,010.84	9,883,616.59
S.D.	53.54	71.03	607,683.97
Max	5,906	7,666	53,388,743
Min	49	25	111,810
Outputs	**Total number of shipments**	**Tonne-Kilometers**	
Mean	457,771.78	508,692,461.9	
S.D.	34,143.24	28,701,722.58	
Max	2,607,683	2,469,461,919	
Min	4,246	1,177,340	

Table 3: Balanced Sample Descriptive Statistics, 1985-1996

Inputs	Number of employees	Total vehicles	Fuel Consumption (liters)
Mean	727.56	1,072.84	10,690,223.04
S.D.	84.16	104.54	852,775.41
Max	5,906	7,666	53,388,743
Min	112	155	1,015,000
Outputs	**Total number of shipments**	**Tonne-Kilometers**	
Mean	430,460.62	544,778,399.4	
S.D.	41,599.41	38,492,130.47	
Max	2,412,023	2,469,461,919	
Min	5,001	4,796,814	

Some size discrepancies among the firms in the sample open the possibility that not all of these firms are operating at constant returns to scale, and the assumption of CRS facilitates the computation of Malmquist indices (Lovell 1993). However, parametric studies of the U.S. trucking industry offer a strong consensus that the for-hire industry operates at constant returns to scale (Braeutigam 1999). Given the large number of firms operating in the Canadian for-hire trucking industry (over 8,000 operators of all sizes as of 1993; Nix 1996), it is likely that post-deregulation trucking technology operates effectively at constant returns to scale. Furthermore, it is very unlikely that this sample of large companies includes any firms operating at increasing returns to scale. Therefore, the constant returns to scale formulation used to compute the DEA efficiency scores and the Malmquist indices seem to be a reasonable assumption for this analysis.

Table 4: Average DEA Scores, All Firms, Full Sample (1985-1997)

Year	CRS DEA Score	Year	CRS DEA Score
1985	1.22	1992	1.09
1986	1.11	1993	1.10
1987	1.16	1994	1.37
1988	1.21	1995	1.08
1989	1.15	1996	1.28
1990	1.22	1997	1.02
1991	1.09		

As a final point about the data, there is no way to directly control for provincial regulatory differences because in order to obtain a reasonable sample size, the data necessarily include firms operating all over Canada.[10] After the process of trucking deregulation began, provinces continued to exert varying levels of control over certain aspects of the industry, particularly safety (Nix 1996). But entry and rate deregulation of trucking was mandated from the federal level and therefore, was relatively homogenous across the country. Therefore, it is safe to assume that any regulatory variability found from province to province will not affect the basic results obtained here.

RESULTS AND INTERPRETATION

Table 5 shows the average of the Malmquist index for each year computed with both the balanced and unbalanced panels of data. Table 6 and Figure 1 show the average of the two productivity components and the Malmquist indices for all firms computed with the balanced panel sample. Since the Malmquist productivity indices are calculated year to year, the productivity gains observed in a given year are relative to productivity levels in the previous year only.

As a test of data sensitivity, the full (unbalanced) sample was used to calculate another series of averaged Malmquist indices and components. These are shown in Figure 2 and Tables 5 and 6. Overall, industry productivity increases steadily over time in the sample (Table 5). Both sets of indices (balanced and unbalanced) show increases in overall productivity from 1986-1988. The effects of the 1990-91 recession can be seen as well in Figures 1 and 2, with productivity slipping in both samples. In fact, this recession hit the trucking industry particularly hard (Nix 1996). Finally, another noticeable increase in productivity appears starting 1994-1995, diminishing slightly towards the end of this sample.

Table 5: Malmquist Average Scores for All Firms, Both Samples (1985-1997)

Year	Index, balanced sample	Index, unbalanced sample	Year	Index, balanced sample	Index, unbalanced sample
1986	1.14	1.09	1992	1.08	1.23
1987	1.65	1.36	1993	1.12	1.17
1988	2.80	1.49	1994	0.89	0.72
1989	1.06	0.98	1995	1.28	1.58
1990	1.01	0.83	1996	1.63	1.02
1991	1.03	0.98	1997	—	1.10

Table 6: Average Malmquist De-composition Components, Both Samples

Balanced Sample			Unbalanced Sample		
Year	Technical Efficiency	Technical Change	Year	Technical Efficiency	Technical Change
1986	1.14	1.00	1986	1.09	1.01
1987	1.27	1.28	1987	0.97	1.42
1988	0.64	4.83	1988	0.97	1.60
1989	4.17	0.30	1989	1.07	0.94
1990	0.98	1.01	1990	0.98	0.86
1991	1.06	0.95	1991	1.12	0.90
1992	0.97	1.12	1992	1.00	1.23
1993	1.05	1.07	1993	0.99	1.18
1994	1.00	0.89	1994	0.91	0.83
1995	0.90	1.41	1995	1.35	1.22
1996	1.32	1.18	1996	0.93	1.21
1997	—	—	1997	1.24	0.95

Figure 1: Balanced Sample Malmquist Productivity De-composition, 1985-1996 (vertical axis is the index value)

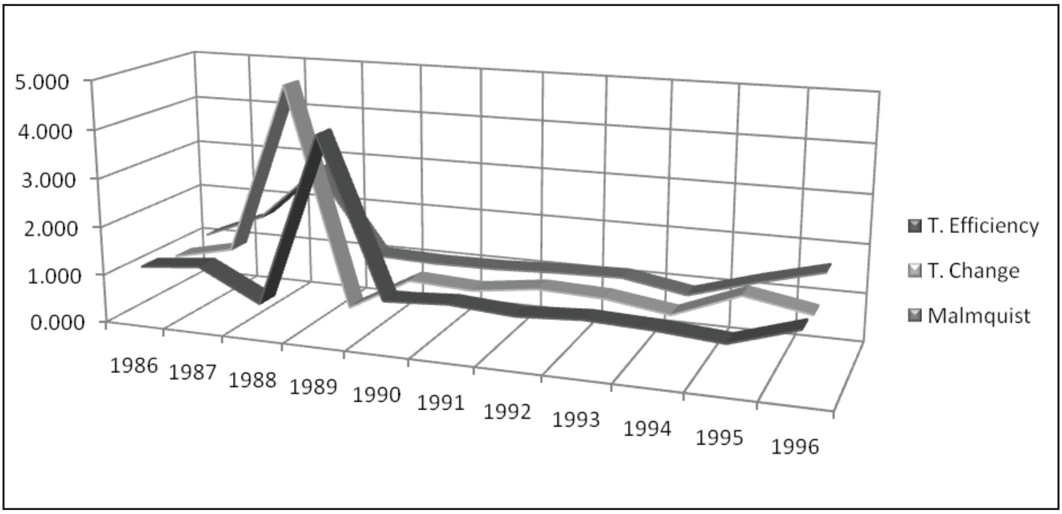
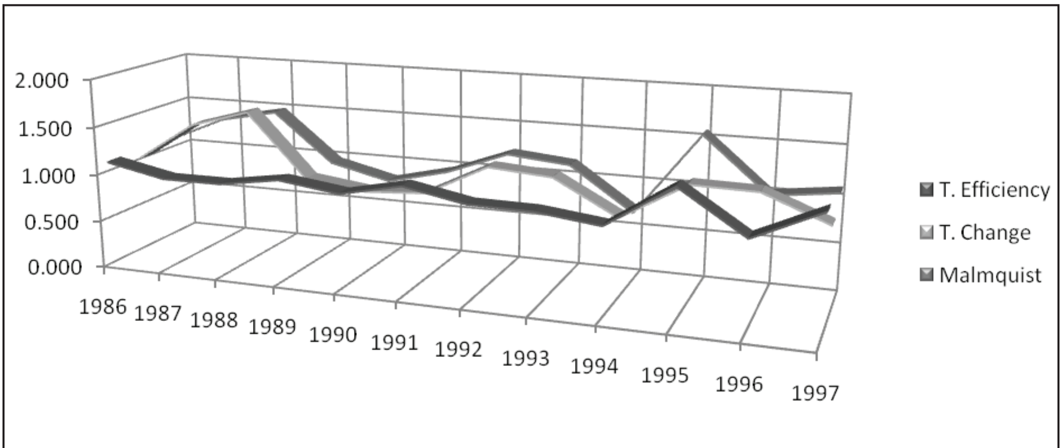


Figure 2: Full (Unbalanced) Sample Malmquist Productivity De-composition, 1985-1997 (vertical axis is the index value)



The index de-compositions indicate that the technical change or innovation term dominates the Malmquist index over most of the sample, and especially in the initial period following deregulation. To help identify relationships between productivity and innovation in the data, sample correlations and a Spearman non-parametric rank test (for independence) between the Malmquist index and the de-composed terms were performed. One reason for computing the Spearman rank statistic is that as a non-parametric test, it allows a more general assessment of the relationship between the non-parametric Malmquist index and its components. The Spearman statistic indicates which of the decomposition factors is more critical to the calculation of the index. Both the sample correlations and the Spearman rank tests are shown in Table 7.

Table 7: Malmquist Component Correlations with the Malmquist Index

Balanced Panel 1985-1996	Entire Sample	1988-1993	1994-1996	Spearman test of Independence (entire sample)*
Technical Change	0.923	0.983	0.586	ρ = 0.8273 (0.0017)
Technical Efficiency	-0.242	-0.311	0.698	ρ = -0.0364 (0.915)
Unbalanced Panel 1985-1997	**Entire Sample**	**1988-1993**	**1994-1997**	**Spearman test of Independence (entire sample)**
Technical Change	0.830	0.984	0.735	ρ = 0.9021 (0.0001)
Technical Efficiency	0.406	-0.468	0.868	ρ= 0.2657 (0.4038)

* The corresponding P-value is listed under each test statistic in brackets.

For both Malmquist indices (balanced and unbalanced), the full sample correlation between the technical change term and the index was extremely high (over 0.8 in both cases), exceeding the measured correlation associated with the technical efficiency term. In addition, in both samples the technical change term is significantly correlated with the Malmquist index (i.e., the null hypothesis of independence between the data series is strongly rejected).

Conversely, the technical efficiency term is found to be statistically independent of the Malmquist index in both samples (i.e., we cannot reject the null hypothesis of independence), although weakly so in the unbalanced panel (P-value = 0.40). But overall, the statistical evidence presented here shows that the technical change or innovation term tended to dominate productivity change in both samples. However, variability in the values of each de-composition component across the sample (see Table 6) indicated that specific periods of interest needed closer examination.

The sample was split into two crucial periods in these initial stages of Canadian trucking deregulation: i) the immediate post deregulation period from 1988-1993; and ii) the latter stages of the deregulation process from 1994-1996. Malmquist correlations in Table 7 show that productivity change at the beginning of the deregulation era was dominated by technical change or innovation. For instance, the technical change term drives the productivity index up to 1993 (with correlations of 0.983 [balanced sample] and 0.984 [unbalanced sample]). This finding is consistent with consensus about the time it took for the industry to adjust to the deregulation transition (Dionne et al. 1998). But the correlations in Table 7 also show that by 1994, the process of innovation driving productivity improvements was slowing. As time progressed, it is clear that the "catching up" or technical efficiency term eventually became the driver (i.e., there is a higher correlation between the technical efficiency term and the productivity indices from 1994-1996 than from 1988-1993 within both samples) behind productivity changes in the Canadian for-hire trucking industry. Thus, the Malmquist index decomposition confirms that innovation among the largest for-hire trucking firms in Canada slowed over time after initial de-regulation, with subsequent productivity changes realized instead through the emulation of extant best practices or technology.

It must be noted that in contrast to these Canadian findings, McMullen and Okuyama (2000) performed a similar analysis on trucking and found that technical change in the early U.S. post-deregulation period declined. The industries in each country are quite similar and there was no evidence of any fundamental differences between the two industries at the time of de-regulation (Bonsor 1995). In this light, one explanation for the observed difference in technical change that is consistent with the interpretation above is that most of the large Canadian for-hire carriers in fact served markets throughout all of North America. The deregulation of U.S. trucking markets predated Canadian deregulation. Adjustments to U.S. policy changes may have induced certain operational changes among the Canadian carriers in this sample, including the need to improve their technical efficiency in order to better compete in the liberalized U.S. market. And while of potential interest, identifying the exact source of the changes in technical efficiency in Canada is beyond the scope of this study.

CONCLUSION

A unique Canadian dataset is used to examine the relationship between productivity and the process of deregulation in the for-hire trucking industry. This research adds to the growing body of empirical work concerning the nature of firm behavior by decomposing a Malmquist productivity index computed through the time of economic deregulation in this industry.

One desirable consequence of economic deregulation is that after implementation, firms will act upon a set of new incentives and ultimately compete by innovating. We identified such an effect within the early stages of trucking deregulation in Canada. It was also found that innovation in the industry ultimately declined over time, likely as a result of firms adjusting to the new regulatory environment by emulating extant technology rather than continuing to pursue innovation. Analyzing whether or not productivity improvements in Canadian for-hire trucking attributable to innovation simply fade after the act of deregulation or are cyclical in nature awaits further data. An extension in this direction would clearly contribute to the extensive and growing literature on innovation and industry life cycles (see for example, Cefis and Marsili 2005).

Events of the past two decades have shown that technology improvements often occur in ways that were unforeseen before changes in economic regulation (Winston 1993). Sometimes an unforeseen technical change is obvious to the casual observer, such as the ubiquitous advent of satellite tracking systems, but in other cases technical improvements and their consequences are more subtle. It is in the latter cases in which the methodology and interpretation implemented here would help measure and identify changes in firm behavior brought about by deregulation. Ideally, it would be instructive to compare these findings with those from industries that have undergone a similar regulatory transition and examine whether differences exist across productivity de-compositions.

Endnotes

1. See Viscusi et al. (1996) for a detailed overview of regulation and its effects on the U.S. trucking industry. The history and evolution of trucking regulation in Canada is similar.

2. The Canadian provinces enforced various forms of their own economic regulation, but federal level deregulation was seen as a crucial step to improving competition in this industry.

3. Since the focus in this research is on the behavior of the largest (Class 1) carriers, the total sample size is limited. This increases the risk of “survivor bias” over the sample (especially in the balanced sample) since these firms remained intact throughout the duration of the regulatory transition. But such bias is not likely to be an issue since the point of the research is to examine those firms who survived the shock of deregulation.

4. Through the sample period, alliances were becoming increasingly important to Canadian carriers looking to expand service into the United States (Transport Canada 1996).

5. On a theoretical note, this discussion implicitly assumes that potential property rights enforcement issues in innovation did not hinder deregulation. Those innovations that became popular in this segment of trucking (e.g., hub and spoke networking, GPS based logistics)

were commonly known in the industry and clearly part of the public domain around the time of trucking deregulation in Canada. This is not to say that property rights issues in innovation will never become important in trucking, only that these issues do not affect the results from the period under analysis.

6. Microsoft Excel Solver was used to compute both the efficiency scores and the Malmquist indices.

7. Class 1 trucking firms were defined in Canada as those firms earning more than C$5 million in revenue per year. This classification was changed in 1996, but the old classification is relevant to the data used here.

8. Some of the data used here are also described in Nolan (1998). The authors would like to also thank R. Laroque at Statistics Canada for compiling much of this dataset.

9. We have made every effort to ensure that the firms in the sample do not rely upon "owner-operators," the use of which could affect the reported inputs for the firm, depending on how the firm chooses to report operating data to Statistics Canada.

10. The data used are confidential, so identifying individual firms is not possible. However, both samples were examined to ensure that they did not include merged firms.

References

Bonsor, N. "Competition, Regulation and Efficiency in the Canadian Railway and Highway Industries." Filip Palda, ed. *Essays in Canadian Surface Transportation.* Vancouver, B.C.: The Fraser Institute (1995): Ch.2.

Bookbinder, J.H. and W.W. Qu. "Comparing the Performance of Major American Railroads." *Journal of the Transportation Research Forum* 34, (1993): 70-85.

Braeutigam, R. "Learning About Transport Costs." J. Gomez-Ibanez, W.B. Tye and C. Winston, eds. *Essays in Transportation Economics and Policy.* Washington, D.C.: Brookings Institution Press (1999): Ch. 3.

Cefis, E. and O. Marsili. "A Matter of Life and Death: Innovation and Firm Survival." *Industrial and Corporate Change* 14, (2005): 1167-1192.

Dionne, G., R. Gagne, and C. Vanasse."Measuring Technical Change and Productivity Growth with Varying Output Qualities and Incomplete Panel Data." *Journal of Econometrics* 87, (1998): 303-327.

Fare, R., S. Grosskopf, M. Norris, and Z. Zhang. "Productivity Growth, Technical Progress and Efficiency Change in Industrialized Countries." *American Economic Review* 84, (1994): 66-83.

Gallamore, R. "Regulation and Innovation: Lessons from the American Railroad Industry." J. Gomez-Ibanez, W.B. Tye, and C. Winston, eds. *Essays in Transportation Economics and Policy.* Washington, D.C.: Brookings Institution Press (1999): Ch. 15.

Lovell, C.A.K. *Production Frontiers and Productive Efficiency*. H.O. Fried, C.A. Lovell, and S.S. Schmidt, eds. The Measurement of Productive Efficiency: Techniques and Applications. New York, N.Y.: Oxford University Press (1993): Ch. 1.

McMullen, B.S. and K. Okuyama. "Productivity Changes in the U.S. Motor Carrier Industry Following Deregulation: A Malmquist Index Approach." *International Journal of Transport Economics* 27 (2000): 335-354.

Nix, F. "Trucking in Canada: A Profile." The Canadian Trucking Research Institute, Ottawa, Canada, 1996.

Nolan, J.F. "Productivity and Deregulation in the Canadian For-Hire Trucking Industry." 77th Annual Transportation Research Board (TRB) Conference, Washington, D.C., pre-print CD-ROM (1998).

Scherer, F.M. and D. Ross. *Industrial Market Structure and Economic Performance*, 3rd Ed. Houghton Mifflin Company, Boston, 1990.

Seiford, L. M. and R.M. Thrall. "Recent Developments in DEA: The Mathematical Programming Approach to Frontier Analysis." *Journal of Econometrics* 46, (1990): 7-38.

Shephard, R.W. *Cost and Production Functions*. Princeton University Press, Princeton, NJ, 1953.

Statistics Canada. *Trucking in Canada*. Transportation Division, Report #53-222, Ottawa, 1993.

Transport Canada. *Transportation in Canada: Annual Report*. Ministry of Public Goods and Services, Ottawa, 1996.

Transport Canada. *Transportation in Canada: Annual Report*. Ministry of Public Goods and Services, Ottawa, 1999.

Viscusi, W.K., J. M. Vernon, and J. Harrington Jr. *Economics of Regulation and Anti-Trust*, 2nd Ed. MIT University Press, Cambridge, Mass., 1996.

Winston, C. "Economic Deregulation: Days of Reckoning for Microeconomists." *Journal of Economic Literature* 31, (1993): 1263-1289.

Wood, D. and J. Johnson. *Contemporary Transportation*, 5th Ed. Prentice Hall, Upper Saddle River, NJ, 1996.

Woudsma, C. G. and P.S. Kanaroglou."Deregulation of the Motor Carrier Industry: a Canadian Example." *Environment and Planning* A 26, (1994): 343-360.

James Nolan *is an associate professor in the Department of Bioresource, Policy, Business and Economics (formerly Agricultural Economics) at the University of Saskatchewan. James is currently developing a set of spatial and computational economic models for transportation and agricultural applications. In addition, James has been a co-editor for the Canadian Journal of Agricultural Economics since 2008.*

Pamela Ritchie *is professor and dean in the Faculty of Business and IT at the University of Ontario Institute of Technology in Oshawa. She received her Ph.D. in accounting and finance from Lancaster University. Her current research interests include transportation and logistics, financial and economic performance, and efficiency studies.*

John Rowcroft *was, until recently, professor of economics at the University of New Brunswick. He is currently an adjunct professor at the University of Ontario Institute of Technology in Oshawa. He received his Ph.D. in economics from Simon Fraser University. His current research interests include transportation and logistics, data envelopment analysis and the mathematical modelling of economic behaviour.*

Empirical Examination of Passing Lane Operational Benefits on Rural Two-Lane Highways

by Ahmed Al-Kaisy and Zachary Freedman

This paper presents an empirical investigation into the operational benefits of passing lanes on rural two-lane highways. Two study sites in the state of Montana were used in this investigation. Performance was examined at a single location upstream and multiple locations downstream of the passing lane at each study site. Using percent followers as a performance measure, operational benefits right after the passing lane ranged between 33% and 42% at one study site and 12% to 19% at the other study site under prevalent traffic levels. Study results also suggest that operational benefits persist for a remarkable distance beyond the end of the passing lane.

INTRODUCTION

Two-lane highways constitute the vast majority of the highway system in the United States. Those highways are known for higher level of interaction between vehicles traveling in the same as well as in opposing directions. Specifically, passing maneuvers are typically performed using the opposing lane when sight distance and gaps in the opposing traffic stream permit. This has serious implications for traffic operations and safety. From a traffic operations perspective, limited passing opportunities would result in higher impact of slow-moving vehicles on mobility and performance.

As traffic level increases and passing opportunities become more restricted, passing lanes are used in practice as one of the most effective means in providing passing opportunities, thus breaking up platoons and improving operations. In practice, the use of passing lanes is often considered more appropriate than the costly alternative of upgrading to continuous four-lane highway sections. Also, it is a practice to provide passing lanes periodically at regular intervals on long stretches of rural two-lane highways.

For the proper use of passing lanes on two-lane highways, their effect on performance needs to be accurately assessed both at the location of the passing lane and for some distance downstream, i.e., beyond the end of the passing lane. The term "effective length" of passing lane is used by the Highway Capacity Manual (HCM) to refer to the total length of the passing lane and the distance downstream, where the effect of the passing lane persists (TRB 2000). Knowing the operational benefits and the effective length of a passing lane would allow the selection of the most appropriate length and frequency of passing lanes on longer stretches of two-lane highways. In this time of increasing fiscal constraints, this information is critical in achieving optimum use of the limited highway improvement funds.

BACKGROUND

The HCM 2000 includes a procedure for the analysis of passing lanes on two-lane highway segments (TRB 2000). For two-lane highways, the HCM uses two service measures, percent time-spent-following (PTSF) and average travel speed (ATS). PTSF is defined as "the average percentage of travel time that vehicles must travel in platoons behind slower vehicles because of an inability to pass" (TRB 2000). The procedures assume that the PTSF within the passing lane is 58% to 62% of its upstream value (value immediately before the passing lane) while ATS within the passing lane is 8 to 11% higher than its upstream value. Also, the procedure provides values for the effective length of passing lanes, which varies between 3.6 and 13 miles based on traffic volume for PTSF, versus a constant 1.7 miles for ATS.

While there are quite a few studies on passing lanes in the literature, only a small number of studies attempted to assess the operational benefits and effective length of passing lanes. Most of those studies were largely dependent on traffic simulation, and some used limited field data in the calibration of simulation models. The most relevant of those studies are discussed in this section.

In a relatively recent study (Woolridge et al. 2002), researchers investigated the optimal passing lane length and spacing on two-lane highways. Field data were collected from study sites in Kansas and Minnesota and used later in calibrating the simulation model TWOPAS. Simulation results showed that the percent time-delay (measured as the percentage of headways less than five seconds) decreased as the passing lane length increased. Based on traffic volumes and type of terrain, the optimal length of passing lane varied from 0.8 to 2.0 miles and the optimal spacing between passing lanes varied from 3.5 to 11 miles.

May (1991) used a TRARR traffic simulation model to investigate the operational benefits and the effective length of passing lanes. The results suggested that for passing lanes between 0.5-2.0 miles in length, benefits tended to diminish after about 2.0 miles. The study also investigated five passing lane locations with different volumes and lengths in California. Overall, passing lanes reduced the percent time-delay and the percentage of headways less than two seconds with the exception of one site where slight increases were observed.

A study by Harwood and Hoban (1987) investigated the optimal length and spacing of passing lanes using simulation. The study used different traffic volumes and lengths of passing lanes, which varied in the range 0.25 mile to 2.0 miles. Results showed that for a flow rate of 400 vehicles per hour (vph) the benefits persisted for at least seven miles from the beginning of the passing lane, while for a flow rate of 700 vph almost all the benefits diminished after five miles. Another earlier study evaluated the effect of passing lanes on operations using the TWOPAS simulation model that was calibrated using field data (Harwood and St. John 1986). The researchers found that the effective length of passing lanes was in the range of three to eight miles. They also found that higher flow rates led to shorter effective lengths. The effective length refers to the total length of the passing lane and the distance downstream, where the effect of the passing lane persists (TRB 2000). The same researchers investigated 12 passing lane sites in several states in an attempt to determine the effective length and spacing of passing lanes (Harwood and St. John 1985). At each site, automatic traffic recorders were placed at locations upstream, within, and up to one mile downstream of the passing lanes where six hours of field data were collected. Using four-second headway or less in identifying platooned vehicles, the study showed that the number of vehicles in a platoon decreased from 35.4 % directly upstream of the passing lane to 20% within the passing lane. Just downstream of the passing lane the percentage increased to 29%. One mile downstream of the passing lane the number of platooned vehicles was 3.5% lower than the upstream value. The average platoon length decreased from 2.9 vehicles to 2.7 vehicles before and after the passing lane, respectively.

Mutabazi, Russell, and Stokes (1999) conducted a simulation study to determine the optimum passing-lane configuration to reduce delay. Traffic data were gathered from two sites in Kansas, US-54 and US-50, for calibration of TWOPAS simulation software. The simulation was then used to determine the best configuration of passing lane. The results showed that percent time delay was equivalent for the side-by-side configuration and head-to-head configuration; however, percent delay was higher for all other configurations.

RESEARCH MOTIVATION

Results from the above mentioned studies show that there is little guidance, if any, in answering the following questions: what are the operational benefits of passing lanes and how far downstream do these benefits persist? The answer to these questions is very important in the planning and design of passing lanes on two-lane highways.

Results from previous studies lack consistency as they varied in too wide of a range to be of any meaningful use. For example, while a study by May (1991) claims that for passing lanes between 0.5-2.0 miles in length, benefits tended to diminish after about 2.0 miles. The HCM (TRB 2000) suggests that benefits may exist for 13 miles beyond the passing lane. Other studies suggested different values between the former two values (Harwood and Hoban 1987, Harwood and St. John 1986). Similar inconsistencies exist in the reported benefits of passing lanes that are measured immediately downstream from the passing lane taper. Those inconsistencies may largely be attributed to the following two reasons:

1. The measures of performance used in estimating the benefits of passing lanes in previous studies are not consistent. Specifically, different values for headways were used to identify platooned vehicles by previous research efforts. For example, some studies used three-second headway (Al-Kaisy and Karjala 2008, Gattis et al. 2006), four-second headway (Harwood and St. John 1985), and five-second headway (Woolridge et al. 2002, May 1991; Taylor and Jane 1991, Harwood et al. 1985) in identifying platooned vehicles.
2. The majority of the already limited published research on this topic relied largely on using traffic simulation models. Many of those studies are somewhat outdated and used early versions of simulation models that may not provide realistic results. Further, the difficulty in obtaining appropriate field data may have compromised the necessary validation of these models using field observations.

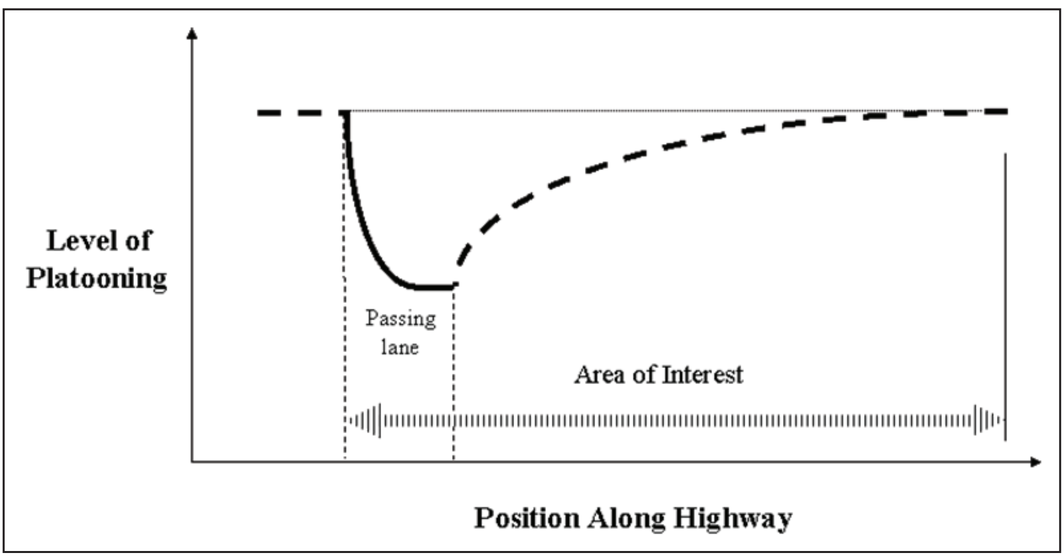
In light of the above, there is a need to conduct empirical research that would provide a more reliable assessment of passing lane benefits using the performance measures currently used in practice.

PASSING LANE EFFECT ON PLATOONING: THEORETICAL PERSPECTIVE

This study aims to investigate the effect of passing lanes on two-lane highway operations and the distance downstream of the passing lane where this effect persists, commonly referred to as the passing lane effective length.

Figure 1 shows the hypothesized performance on two-lane highways in the presence of a passing lane. It is somewhat simplistic and very consistent with a similar figure provided in Exhibit 20-22 in the HCM (2000) where the percent time-spent-following is used as a platooning measure. The figure shows the theorized change in platooning level versus position prior, within, and past the passing lane. Upon arriving at the passing lane, traffic on a two-lane highway with restricted

Figure 1: Theoretical Graph Showing Platooning Level Within and Downstream of a Typical Passing Lane on a Two-Lane Highway



passing opportunities is expected to have a relatively high platooning level. Upon entering the passing lane section, traffic is distributed on the two available lanes, with slower vehicles typically using the passing (right) lane and faster vehicles remaining on the normal (left) lane. This speed-based segregation usually results in passing maneuvers, improved speeds, and decline in platooning level. At the end of the passing lane section, vehicles in the passing lane have to move back into the normal lane, merging with the rest of traffic using a taper.

At this point, all vehicles share again the original travel lane; however, slower vehicles are not necessarily leading platoons at this point. This denotes less impedance to traffic, higher average speeds, and improvement in quality of service. As traffic moves downstream farther from the passing lane, faster vehicles catch up with slower vehicles and platooning level is expected to steadily increase until it eventually reaches its original level (pre-passing-lane value). The stable original level of platooning is based on the assumption that there are no changes in geometric features or passing restrictions throughout the site except for the presence of the passing lane.

STUDY APPROACH

This study used empirical observations to investigate the effect of passing lanes on platooning and performance on rural two-lane highways. Only highways that are located in rural settings, typically characterized by relatively low volumes and high speeds, are considered within the scope of this study. Further, climbing lanes (passing lanes on extended upgrades) are considered beyond the scope of the study.

The study approach required the selection of passing lane sites where the passing lane and a significant distance downstream are free from major access points and are generally located in level terrain. Several data collection stations were installed upstream and at various distances downstream of the passing lane using automatic traffic recorders. Three performance measures were estimated to discern the effect of the passing lane and the change in that effect as traffic traverses and moves away from the passing lane. Those performance measures are:

1. **Percent Followers:** Percent followers represent the percentage of vehicles with short headways in the traffic stream. Specifically, a follower is a vehicle following another vehicle in the traffic stream with a short headway in between. This performance indicator can easily be measured in the field and as such is used by the HCM as a surrogate measure for PTSF field estimation. The headway cut-off value used by the current HCM is three seconds, which is the same value used in this study.
2. **Follower Density:** Follower density is the number of followers in a directional traffic stream over a unit length such as a mile or a kilometer. It is estimated as the product of density and percent followers, as in the following equations:

   (1) Follower Density = Density * Percent Followers
   (2) Follower Density = (Hourly Flow Rate / Average Speed) * (% Headways < three seconds)

   The argument behind using this performance indicator is that a road with low average daily traffic (ADT) and high PTSF should have a lower level of service (LOS) than the same road with a higher ADT and PTSF equal to the low ADT road (Van As 2003). This is particularly true in the context of highway improvement decision making for which the LOS and other operational analyses are important inputs.
3. **Average Travel Speed as a Percentage of Free-Flow Speed (ATS/FFS):** This is an indicator of the amount of speed reduction due to platooning on two-lane highways. In the context of this research, if the percentage is high, then the interaction among successive vehicles in the traffic stream is small, and a low platooning level is expected. By the same token, a lower percentage indicates a higher level of platooning.

For details on the above measures, please refer to a recent study on two-lane highway performance measures (Al-Kaisy and Karjala 2008).

DATA COLLECTION AND REDUCTION

To investigate the effect of passing lanes on a platooning level, it was necessary to consider the following criteria in the selection of study sites:

1. No major driveways or intersections exist within the passing lane and for some distance upstream and downstream.
2. As practically as possible, the site should be free of tangible grades (grades greater than 2%) or other features that are believed to affect platooning on two-lane highways.
3. The site should possess geometric features that are considered typical for other passing lanes on rural two-lane highways.

The above criteria followed in the selection of study sites made it difficult to identify appropriate sites for this study. Upon examining seven candidate sites, only two sites largely met the criteria above and were therefore selected for the study. As-built roadway drawings, video logs, and Google earth images were all used in screening candidate study sites.

Study Site I. This site is located on Highway US 287 between the towns of Three Forks and Townsend in the state of Montana. Townsend is approximately 5.5 miles upstream of this site and therefore the impact of any traffic interruption on platooning is considered minimal, if any. The total length of the site is about 8.25 miles including the length of the passing lane, which is approximately 1.5 miles (including tapers). The major merit of this site is the relatively long distance downstream that met the selection criteria. Traffic level at this site is relatively low.

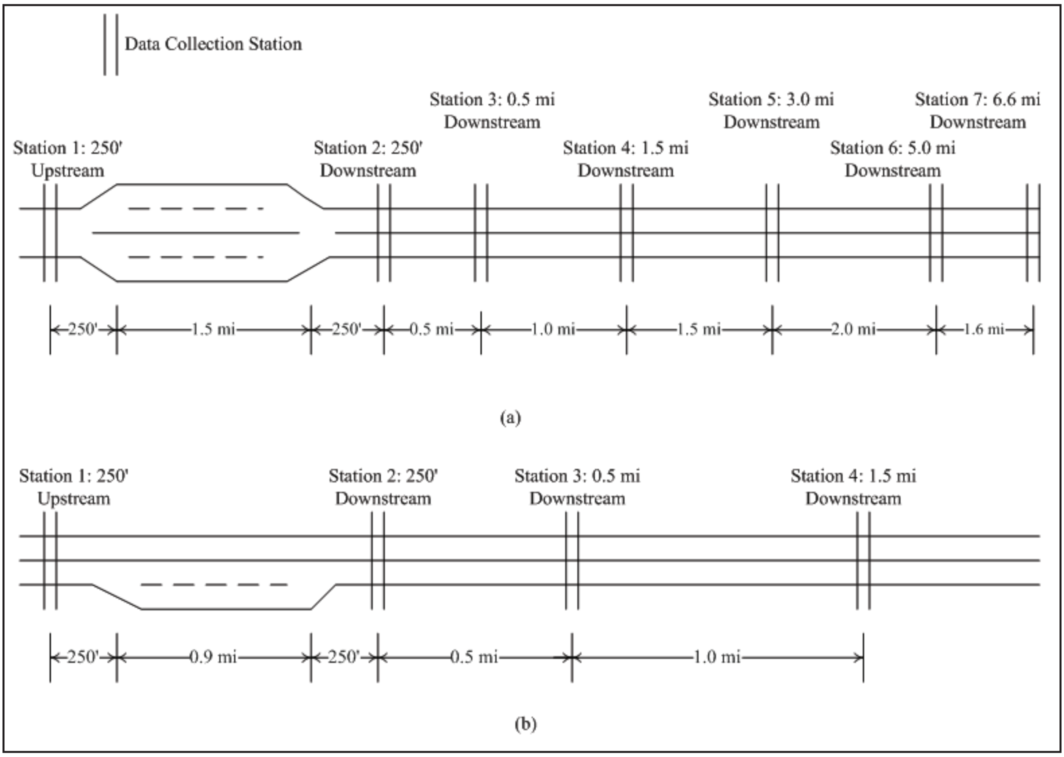
Study Site II. This site is located on Highway US 191 just north of the Gallatin canyon in southwest Montana. The upstream end of this site is about 30 miles from the nearest intersection, and as such no effect on platooning is expected. The total length of this site is approximately 2.5 miles including the length of the passing lane, which is about 0.9 miles (including tapers). The distance downstream is limited to 1.5 miles, which is relatively short. However, traffic level at this site is notably higher than that at the US-287 site, which is important to this investigation.

Data collection setups at the two study sites are shown in Figure 2. Seven data collection stations were installed at the US-287 site versus four stations at the US-191 site. The first station at both sites is located 250 feet upstream of the passing lane. As applicable, the other stations are located at roughly 250 feet, 0.5 miles, 1.5 miles, 3 miles, 5 miles, and 6.6 miles downstream of the passing lane, respectively. The distances mentioned were measured from the end of the passing lane tapers. Automatic traffic recorders were used to measure traffic counts, individual vehicle speeds, vehicle classification, and headways.

Data were collected over a one-week period in May 2008. Due to a malfunction in the traffic recorder at the upstream station at the US-287 site, additional data from the upstream and downstream stations were collected over a one-week period in June 2008. A total of 1,583 hours (about 66 days) of traffic records with more than 222,000 observations were collected at the two study sites.

Traffic records were thoroughly screened to make sure that any defective data were eliminated from the data sets. Also, to avoid any possible effect of light condition and rain on performance measures, only daytime periods with dry weather were used in this study.

Figure 2: Data Collection Setup (a) US 287 Study Site (b) US 191 Study Site



STUDY RESULTS

The hypothesized platooning level graph shown in Figure 1 was tested using analyses that aim at exploring the association between performance, the presence of passing lane, and the distance past the end of the passing lane. The terms "platooning" and "performance" will be used interchangeably in the following sections. While two sites were used in this study, only the US-287 site was used throughout the analysis. The other site, the US-191, was used for the preliminary examination and analysis of headway distribution only. The main reason for not including the US-191 site in the statistical analyses is that this site did not allow the investigation of adequate distance downstream of the passing lane, which was necessary for examining the effect of the passing lane on platooning.

Preliminary Examination

The first step in testing the research hypothesis was to examine performance measures at data collection stations to discern trends in platoon dispersion and formation as related to the presence of passing lanes.

To assess the immediate effect of the passing lane on performance, a comparison of performance measures between the upstream and downstream stations is provided in Table 1. Traffic levels at the two sites were established based on the observed traffic volumes at the two study sites. Upon carefully examining this table, the following observations can be made.

Table 1: Changes in Performance Measures over the Passing Lanes at Study Sites

Performance Measure	Upstream	Downstream	Percent Change
US 287 Site			
Average Hourly Volume = 130 vph			
Percent Followers	36.3	21.4	-41.05
Follower Density	0.697	0.400	-42.67
ATS/FFS	0.98	0.99	0.81
Average Hourly Volume = 155 vph			
Percent Followers	40.8	27.1	-33.58
Follower Density	0.922	0.564	-38.83
ATS/FFS	0.98	0.96	-1.43
US 191 Site			
Performance Measure	**Upstream**	**Downstream**	**Percent Change**
Average Hourly Volume = 110 vph			
Percent Followers	25.32	21.03	-16.94
Follower Density	0.37	0.33	-11.71
ATS/FFS	0.984	0.969	-1.52
Average Hourly Volume = 200 vph			
Percent Followers	38.79	31.49	-18.82
Follower Density	1.04	0.888	-14.62
ATS/FFS	0.987	0.948	-3.95
Average Hourly Volume = 330 vph			
Percent Followers	53.31	46.61	-12.57
Follower Density	2.45	2.28	-6.94
ATS/FFS	0.984	0.954	-3.05

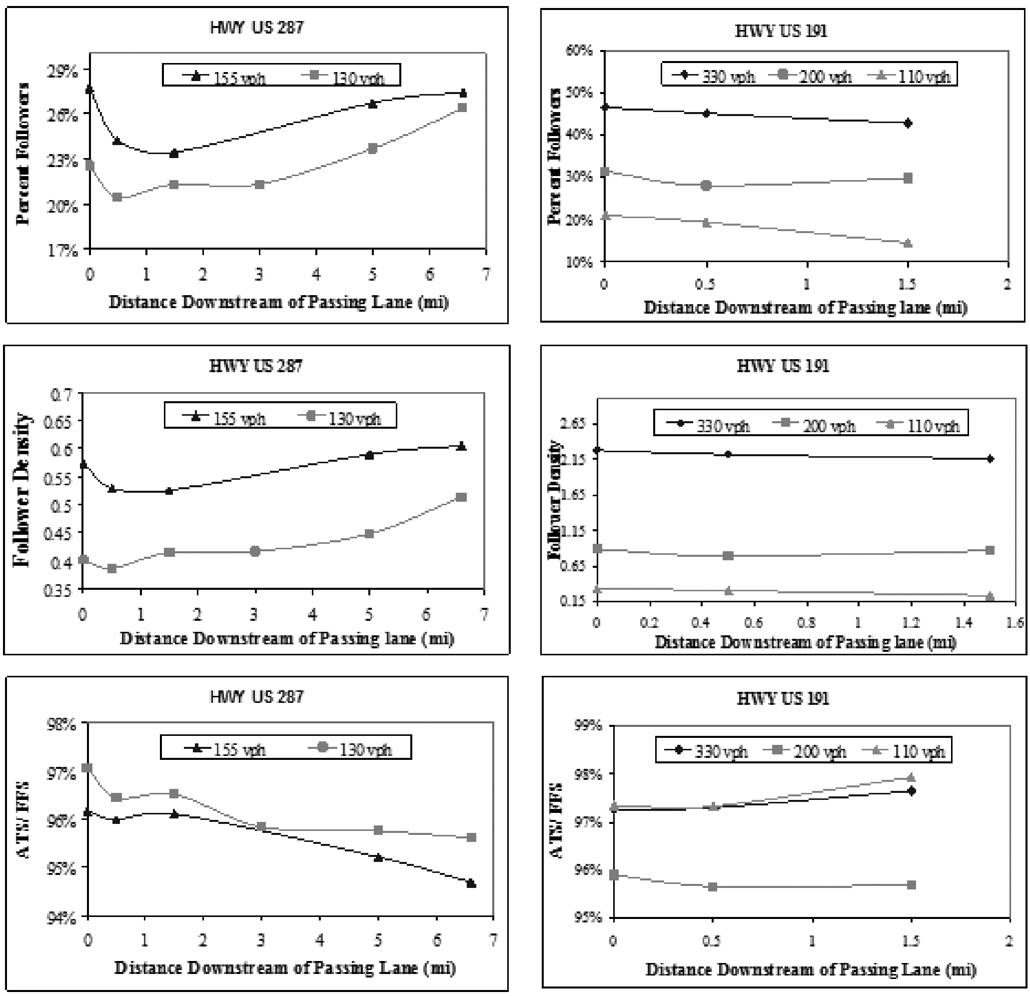
1. The percent changes in performance measures are greater at the US-287 site relative to the US-191 site. This is consistent with the fact that the passing lane at the US-287 site is considerably longer than that at the US-191 site.
2. The headway-related measures, i.e., percent followers and followers' density, have shown remarkable improvements at the two study sites under various traffic conditions. These improvements are more notable at lower traffic levels.
3. Changes in the third performance measure (ATS/FFS) are relatively small particularly at lower traffic levels.

To examine the part of the platooning graph beyond the passing lane, the three performance measures were plotted at various stations downstream of the passing lane as shown in Figure 3. The following important trends are clearly exhibited in this figure:

1. Percent followers and follower density exhibit improvement in performance for a relatively short distance downstream of the passing lane. While this trend applies to the two study sites, it is more evident at the US-287 site, where a considerable downstream distance was investigated. Also, the improvement is somewhat more notable at higher traffic levels.

2. Beyond the 1.5 mile station (US-287 site), platooning level increases steadily as traffic moves away from the passing lane. This trend is common to all performance measures used in this study.
3. The increase in platooning level in these graphs does not flatten out or reach a plateau as suggested by the hypothesized platooning curve. This may be an indication that the operational benefits of the passing lane continue well beyond 6.6 miles, the farthest distance investigated in this research.

Figure 3: Change in Performance Measures Downstream of Passing Lanes at Study Sites



While the first observation may initially sound unexpected, it can easily be understood if one thinks about the dynamics of breaking up platoons by passing lanes more at the microscopic level. Initially, vehicles arrive at the passing lane location with relatively high level of platooning. As vehicles enter into the passing lane, a portion of vehicles, including the slower moving vehicles, usually take the passing (right) lane. This maneuver normally takes place at the taper location, but some vehicles may change lanes later into the passing lane section. This will allow faster vehicles on the normal (left) lane to travel at higher speeds, not impeded by the once platoon leaders, the slower moving vehicles. Eventually, many of the faster vehicles will pass the slower vehicles within the passing lane. As the passing lane comes to an end, vehicles in the passing lane have to merge back into the normal lane. This merge activity creates short headways, as many vehicles are forced to accept shorter headways moving back into the normal lane. Those vehicles may need to adjust their speed, which is normal in a similar merge situation. As a result, platooning level (denoted by the presence of short headways) becomes "relatively" high right after the passing lane. However, platoons at the downstream end of the passing lane are different in attributes from those at the upstream end, as the former are largely associated with the aforementioned merge activity, while the latter are mostly associated with slower vehicles impeding the movement of faster vehicles. As vehicles move farther from the passing lane, some vehicles adjust the short headways to their comfort level, thus reducing the percentage of short headways and hence platooning level. This activity takes place over a short distance beyond the passing lane, which could be a function of traffic level. On the other hand, the impedance of slower vehicles to faster vehicles, which was largely diffused by the passing lane, starts to build up again as vehicles move farther from the passing lane and faster vehicles eventually catch up with the slower vehicles. This explains the increase in platooning level at the stations beyond the 1.5-mile station at the US-287 site.

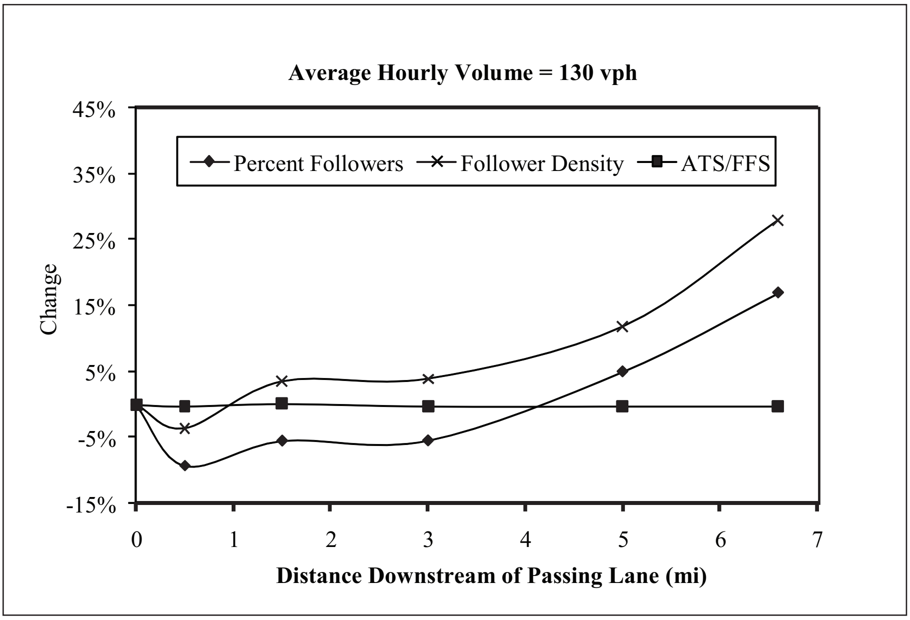
To better understand the magnitude of change of performance measures as traffic moves farther from the passing lane, the percent of change for each performance measure is plotted downstream of the passing lane at the US-287 site as shown in Figure 4. While the change in the ratio of average speed to free-flow speed is minimal, changes in headway-related measures are quite notable. This indicates that the headway-related measures investigated are more sensitive to platooning on two-lane highways than the speed ratio.

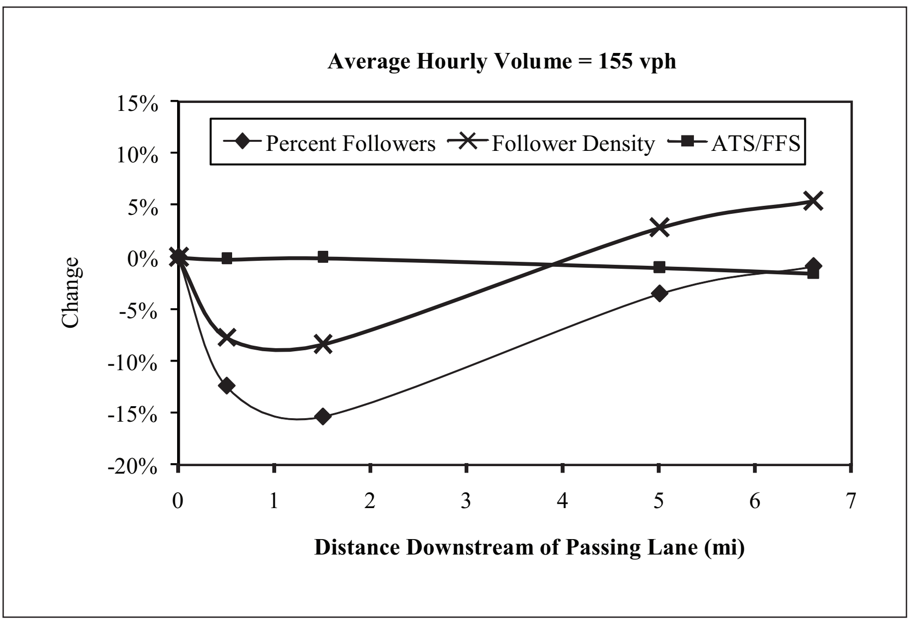
Table 2 shows the difference in performance between the upstream station and the farthest downstream station at the US-287 site using headway-related performance measures. The figures presented in this table show that the residual operational benefits at 6.6 miles downstream of the passing lane are fairly high. This suggests that the effective length of the passing lane could extend a considerable distance beyond the location of this station.

Table 2: Residual Operational Benefits at 6.6-Miles Station (US-287 Study Site)

Average Hourly Volume = 130 vph			
Performance Measure	**Upstream**	**6.6 miles Downstream**	**Residual Benefits**
Percent Followers	36.3	25.8	28.93%
Follower Density	0.697	0.510	26.85%
Average Hourly Volume = 155 vph			
Performance Measure	**Upstream**	**6.6 miles Downstream**	**Residual Benefits**
Percent Followers	40.8	27.7	32.11%
Follower Density	0.922	0.617	33.08%

Figure 4: Percent Change in Performance Measures Downstream of the Passing Lane at US-287 Study Site



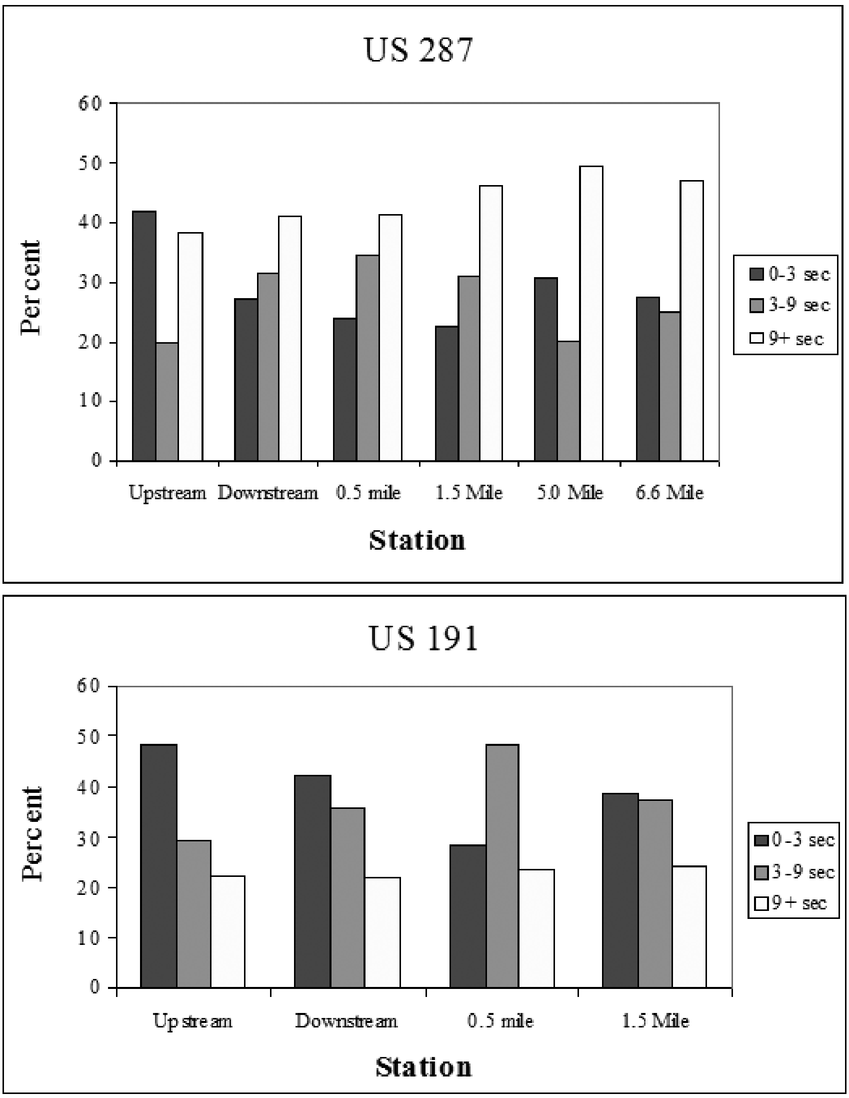


Examination of Headway Distribution

As platooning is directly related to headways, headway distributions were plotted at the two study sites using a three-interval scheme as shown in Figure 5. Upon examining this figure, it is evident that short headways (less than three seconds) have dropped remarkably at the downstream station in comparison with that just before the passing lane at the two study sites.

However, there is sufficient evidence that the actual improvement in performance could be even greater if one considers the merge effect and its associated short headways as discussed earlier. This explains the further decrease in short headways over the following station(s): 0.5-mile station at the two study sites and 1.5-mile station at the US-287 site. Beyond this point, the percentage of short headways generally increases, indicating an increase in the platooning level.

Figure 5: Headway Distribution at Data Collection Stations at Study Sites



At the US-287 site, this percentage remains lower than the original percentage at the upstream station. This may well indicate that, for the traffic level investigated, operational benefits of the passing lane persist for a distance longer than the distance investigated in this research (6.6 miles). On the other hand, the percentage of long headways (nine seconds or greater) generally increases, and the medium size headways (three to nine seconds) decreases beyond the downstream station. This may indicate an increased level of vehicle segregation into platooned and non-platooned vehicles and hence into longer and shorter headways. At the US-191 site, trends are consistent with their counterparts at the US-287 site for the limited distance investigated downstream of the passing lane.

Statistical Analysis

To examine the effect of distance downstream of the passing lane on platooning and performance, data from the US-287 site were further analyzed to get a better understanding of their association. Data from the US-191 site were not used in the statistical analysis due to the short downstream distance investigated by this study. Table 3 provides the coefficients of correlation among the three performance measures and all "platooning" variables, including the distance downstream of the passing lane, referred to as "distance" in the table. The percentage trucks refers to the percentage of heavy vehicles in the traffic stream.The merge effect refers to the effect of merge activity on headways and platooning that only exists at the data collection station immediately downstream of the passing lane. It is included in this analysis as an indicator (dummy) variable. An indicator variable is one that takes the values 0 or 1 to indicate the absence or presence of some effect, in this case the merge activity, which may be expected to shift the outcome. While the study site is generally located in level terrain, the exact percentage grade from as-built drawings at various data collection stations was used in the analysis. Percent no passing in this study refers to the percentage of length of highway where passing is prohibited within one mile upstream of the data collection station.

Table 3: Correlation Coefficients Between Performance Measures and Major Platooning Variables

US 287						
Performance Indicator	**Volume**	**% No Passing**	**Distance**	**Merge Effect**	**Grade**	**% Trucks**
Percent Followers	0.800	-----	0.169	-----	-----	0.109
Follower Density	0.832	-----	-----	0.104	-----	-----
ATS/FFS	-0.744	-----	-0.248	0.145	-0.194	-----

Note: Cells marked with "-----" refer to coefficient of correlation less than 0.1.

The results show that traffic volume has the highest correlation with performance measures (all coefficients are above 0.74). The results also suggest a weak correlation among the three performance measures and other platooning variables, including the distance downstream of the passing lane. Overall, the percent-no-passing showed the lowest correlations with the three performance measures. All correlation coefficients exhibited logical relationships, i.e., platooning increases with the increase in traffic volume, grade, distance downstream of passing lane, percentage trucks, and the presence of merge effect.

To get a closer look at the association between performance measures and distance downstream of the passing lane, it was deemed appropriate to control the effect of volume in the analysis. As such, the analysis of correlations was repeated using data that roughly reflects the same traffic level. The new data set only involves hourly volumes that are in the range 125 vph to 135 vph, which constitute the majority of the observations. At this traffic level, the correlation coefficients were calculated between performance measures and all other platooning variables as shown in Table 4. This time, the distance downstream of the passing lane showed relatively high correlations with performance measures as shown in this table (all coefficients are above 0.73). All distance coefficients exhibited logical relationships with performance measures. Specifically, as the distance downstream of the passing lane increases, percent followers and follower density increase as well. On the other hand, as this distance increases, the ratio of average travel speed to the free-flow speed decreases. Only one illogical (and low) correlation coefficient was found between follower density and percent no-passing. What this tells us is that follower density was not reasonably sensitive to percent no-passing at this study site. In case studies where a specific site is investigated, it is likely to have an illogical weak correlation given the amount of field data used and the many variables that could affect the measure of interest.

Table 4: Correlation Coefficients Between Performance Measures and Platooning Variables at Same Traffic Level

US 287					
Performance Indicator	**% No Passing**	**Distance**	**Merge Effect**	**Grade**	**% Trucks**
Percent Followers	-----	0.732	-----	0.338	0.189
Follower Density	(-0.233)	0.735	0.139	0.21	0.197
ATS/FFS	-0.399	-0.94	0.711	-0.68	0.579

Note: Cells marked with "-----" refer to coefficient of correlation less than 0.1
Values in brackets are those that exhibited illogical relationships

Besides correlation coefficients, multivariate linear regression, another tool used to explore linear relationships, was also performed to examine the functional relationships between performance indicators and platooning variables with special emphasis on the distance downstream of the passing lane.

Table 5 summarizes the results of the regression analysis, which show the relationship between the response variable (performance indicators) and predictor (platooning) variables. Upon examining this table, the following observations can be made:

1. All three models are significant at the 95% confidence level as affirmed by the F-test results.
2. The coefficients of determination (R-square) are relatively high, which shows that much of the variation in the performance measures is explained by the respective models. Traffic volume and distance downstream of the passing lane were found significant in all regression models.
3. This indicates that distance beyond the end of the passing lane has significant effect on all performance measures investigated in this study. Merge effect and percent no-passing were found only significant in the model for estimating percent followers.
4. Grade was found insignificant in all regression models, which shows that it had no tangible effect on any of the performance measures. This is mainly due because the study site is located in generally level terrain and the amount of changes in grade among data collection stations was minimal.
5. All predictor variables that were found significant using the t-test exhibited logical relationships with the dependent variable.

Table 5: Results from Multivariate Linear Regression Analysis at Study Site US 287

US 287									
Performance Indicator	**Regression Model**			**Coefficient and P-value from t-test[2, 3]**					
	P value F-test[1]	**R-squared**	**SE**	**Volume**	**% No Passing**	**% Trucks**	**Dist**	**Merge Effect**	**Grade**
Percent Follower	*<0.001*	0.74	2.66	**0.0741 <0.001**	**0.0323 0.01**	**0.055 (0.09)**	**0.789 <0.001**	**3.42 <0.001**	0.626 0.43
Follower Density	*<0.001*	0.72	0.09	**0.00023 <0.001**	0.00002 0.76	-0.0154 0.92	**0.0122 0.02**	**0.0877 (0.08)**	-0.0081 0.76
ATS/FFS	*<0.001*	0.67	0.77	**0.00008 <0.001**	-0.0012 0.78	0.201 0.59	**-0.174 <0.001**	-0.048 0.20	-0.194 0.11

[1] Values underlined in italic refer to models that were found significant using the F-test
[2] Values in bold are for coefficients that were found significant using the t-test
[3] Values in brackets are those that passed significance testing at the 90% confidence level only

To further examine the effect of distance on performance measures, another regression analysis was conducted where all variables that were found less important were dropped from further analysis. This has narrowed the list of variables to traffic volume and distance downstream of the passing lane. To control for the effect of volume, only field data that roughly represents the same traffic level (125-135 vph) were used in the regression. This allowed dropping volume from the regression analysis leaving distance as the only predictor variable.

The objective is to answer the question: for a given traffic volume, what is the effect of distance on traffic performance and platooning? The results are presented in Table 6. Upon examining the F-test and t-test results, it is evident that the three models are significant at the 95% confidence level. Also, the only predictor variable, distance downstream of the passing lane, was found significant in all three models. The R-square values suggest that distance explained approximately 54% of the variation in percent followers and follower density and approximately 89% of the variation in the speed ratio. These results are generally consistent with those from the correlation analysis.

Table 6: Results from Multivariate Linear Regression Analysis with Distance as the Only Predictor Variable

Performance Indicator	**Regression Model**			**P-value from t-test**
	F-test	**R-squared**	**SE**	**Distance**
Percent Follower	<0.001	0.536	1.598	<0.001
Follower Density	<0.001	0.540	0.023	<0.001
ATS/FFS (%)	<0.001	0.887	0.175	<0.001

CONCLUDING REMARKS

Most studies in the literature either relied heavily on simulation or investigated a short distance downstream of the passing lane. Therefore, there were conflicting findings in the literature in regards to the distance where the effect of the passing lane persists, also called the effective length of the passing lane. The current research utilized empirical data to investigate the effect of passing lanes on rural two-lane highways. Analyses were conducted to assess the operational benefits of passing lanes and how those benefits diminish as traffic moves away from the passing lane. The major findings of this study are:

1. Using headway-related performance measures, platooning level decreases beyond the downstream end of the passing lane for a short distance before it starts to increase again as traffic moves away from the passing lane. For the study sites investigated, results suggest that this distance could be as long as 1.5 miles beyond the end of passing lane taper. This phenomenon is related to the merge activity at the end of the passing lane, which may force many drivers to accept headways shorter than what they would comfortably select. This shows a limitation in using the headway-related performance measures as indicators of platooning and impedance on two-lane highways.
2. Despite the effect of the merge activity, the change in headway-related performance measures before and after the passing lane is notable in most instances and is a function of passing lane length and traffic level. Changes in headway-related performance measures ranged between 33% and 43% at the US-287 site versus 7% and 19% at the US-191 site. The true improvement in performance caused by the passing lane is expected to exceed these values should the merge effect discussed earlier be considered.
3. Research results showed considerable residual operational benefits (around 30% reduction in platooning level) at 6.6 miles downstream of the passing lane at the US-287 study site. This suggests that the passing lane operational benefits at this site may well persist for a few more miles beyond the distance investigated, and consequently the effective length is likely to be in the order of 10 miles (or even longer). This finding may contradict other studies in the literature, but it is somewhat consistent with the value of the passing lane effective length provided in the current HCM. Specifically, the HCM suggests that the passing lane benefits in breaking up platoons could persist for up to 13 miles depending on traffic volumes. While this study could not examine platooning for such a long distance, the residual benefits at 6.6 miles suggest that such benefits may persist well beyond the distance investigated in this research.

Given the many variables that could affect the passing lane effective length, the authors strongly recommend further empirical studies on passing lanes which consider other ranges of traffic volumes and passing lane lengths. This will help to develop a better understanding of the operational benefits of passing lanes, which is critical in the proper planning and design of passing lanes on rural two-lane highways.

Acknowledgements

The authors would like to acknowledge the financial support to this research by the Western Transportation Institute of Montana State University.

References

Al-Kaisy, A. and S. Karjala. "Indicators of Performance on Two-Lane Rural Highways: An Empirical Investigation." *Transportation Research Record 2071*, (2008): 87-97.

Gattis, J., R. Bhave, and L. Duncan. "Alternating Passing Lane Length." *Transportation Research Record 1961,* (2006): 16-23.

Harwood, D. and C. Hoban. "Low Cost Methods of Operational Improvements on Two-Lane Highways." Midwest Research Institute. Report No. FHWA/IP-87/2, 1987.

Harwood, D. and A. St. John. "Operational Effectiveness of Passing Lanes on Two Lane Highways." Federal Highway Administration, Report No. FHWA/RD-86/195, 1986.

Harwood, D. and A. St. John. “Passing Lanes and Other Operational Improvements on Two-Lane Highways.” Federal Highway Administration, Report No. FHWA-RD-85-028, 1985.

May, A. “Traffic Performance and Design of Passing Lanes.” *Transportation Research Record* 1303, (1991): 63-73.

Mutabazi, M., E. Russell, and R. Stokes. “Review of the Effectiveness, Location, Design, and Safety of Passing Lanes in Kansas.” *Transportation Research Record* 1658, (1999): 25-33.

Taylor, W. and M. Jain. “Warrants for Passing Lanes.” *Transportation Research Record* 1303, (1991): 83-91.

Transportation Research Board. *Highway Capacity Manual*. Fourth Edition. TRB, National Research Council, Washington, D.C., 2000.

Van As, C. “The Development of an Analysis Method for the Determination of Level of Service on Two-Lane Undivided Highways in South Africa.” Project Summary. South African National Roads Agency, 2003.

Woolridge, M., C. Messer, B. Heard, S. Rughupathy, A. Parham, M. Brewer, and S. Lee. “Design Criteria for Improved Two-Lane Section (Super 2).” Texas Transportation Institute, Project Report 0-4064, 2002.

Dr. Ahmed Al-Kaisy *is an associate professor and the principal transportation faculty member in the Department of Civil Engineering at Montana State University. He is also program manager for the Safety and Operations Focus Area at the Western Transportation Institute and a registered professional engineer in the state of Montana. Dr. Al-Kaisy received a Ph.D. degree in transportation engineering from Queen’s University in Ontario, Canada.*

Zachary Freedman *received an M.Sc. degree in transportation engineering from Montana State University in 2009, and a B.Sc. in civil engineering from the University of Utah in 2007. His graduate research was in the area of traffic performance on two-lane highways. He currently works as a transportation engineer with the South Dakota Department of Transportation.*

Studying the Tailgating Issues and Exploring Potential Treatment

by Miao Song and Jyh-Hone Wang

A human factors study consisting of a vehicle headway analysis and a questionnaire survey was conducted in Rhode Island (RI) to investigate tailgating issues and possible means for tailgating treatment. Vehicle headways were collected from highway surveillance videos and serious tailgating issues were identified on RI highways. The results of the questionnaire survey further confirmed the observations made in the vehicle headway analysis that most RI drivers maintained insufficient vehicle headways on highways. Among a few tailgating treatments presented in the survey, most subjects preferred a system consisting of equally spaced, white horizontal bars marked on pavement and overhead graphic-aided dynamic message signs.

INTRODUCTION

This paper presents a human factors study that investigated the tailgating issues and possible means for tailgating treatment. "Tailgating" is defined as following a vehicle with insufficient vehicle headway where vehicle headway is the time interval that the two consecutive vehicles passed the same reference point. Following with a vehicle headway less than two seconds is considered insufficient and unsafe (Hutchinson 2008). Tailgating is one of the most dangerous and aggressive driving behaviors, and is a major cause of rear-end crashes which resulted in 1.8 million incidents in the United States in 2006 (National Highway Traffic Safety Administration 2008). To help reduce crashes caused by tailgating, effective tailgating treatments need to be found to advise drivers about maintaining proper vehicle headways.

In this study, a vehicle headway analysis was first conducted to assess the tailgating situation on Rhode Island highways. It next surveyed drivers' opinions regarding the causes and effects of tailgating, their experiences and perceptions on tailgating behavior, and their preferences for possible tailgating treatments consisting of advisory signs and road markings. Findings from the vehicle headway analysis and the survey lead into the next phase of the project where a driving simulation experiment is developed to allow a real-time test on drivers' responses to different tailgating treatments. It is expected, upon completion of the project, that an effective tailgating treatment system would be recommended to Rhode Island and other states. The findings of this project might interest practitioners in this field and possibly contribute to the development of a standard tailgating treatment systems to be included in the Manual on Uniform Traffic Control Devices (MUTCD). Ultimately, it could help ease the tailgating issues and facilitate safer driving on U.S. highways.

BACKGROUND

Reviews of literature and past studies involving tailgating issues and tailgating treatments are provided below.

Tailgating Issues

Tailgating is generally considered a form of aggressive driving (Teigen 2007). The National Highway Traffic Safety Administration (NHTSA 2008) defines "aggressive driving" as "an individual committing a combination of moving traffic offenses so as to endanger other persons or property."

While many driving patterns are considered aggressive, tailgating is among the most dangerous ones and is a major cause of rear-end crashes.

Out of about 5.9 million police-reported automobile accidents in the United States during the year 2006, rear-end collisions ranked the highest, with more than 1.8 million cases (30.4%), and resulted in more than 2,100 fatalities and approximately 500,000 injuries (National Highway Traffic Safety Administration 2008). Data from athe Federal Highway Administration (FHWA) indicate that each year, approximately 2.2% of total licensed drivers in the United States are involved in rear-end crashes (Singh 2003). Two factors are primarily responsible for rear-end crashes: inattention and tailgating (Dingus et al. 1997), while the latter is the major contributing cause with a deadly consequence (Carter et al. 1995).

Some past researchers showed that a wide range of factors such as drivers' behavior, traffic condition, road condition, roadway design, state law and regulation, and even personality had effects on vehicle headway (Ayres et al. 2001; Aycin and Benekohal 1998; Brackstone and McDonald 1999; Hogema 1999; Brackstone 2003; Rajalin, Hassel, and Summala 1997). Based on these factors, various car following models were developed to describe the interaction between individual vehicles or the whole traffic dynamic. However, none of them compared these factors and identified factors that have major effects on vehicle headway.

While driving on highways, a driver's reaction time varies from 0.5 second for simple situations to 4 seconds for complex situations and the reaction time in braking is about 2.5 seconds (American Association of State Highway Officials [AASHO] 1973). Green (2000) and Summala (2000) reported that simple reaction time was often less than one second while decision reaction time could take much longer. According to this, quantified safe following distance has been written into rules of the road. It varies from state to state, but is mostly in the form of a "two-second rule." Drivers are advised to keep a vehicle headway of at least two seconds from the vehicle ahead driving in the same direction. Rear-end crash risk increases as vehicle headway decreases. When vehicle headway reduces to zero, a rear-end crash occurs.

Tailgating Treatments

Hutchinson (2008) conducted an in-depth investigation in Australia of rear-end crashes, tailgating, and the correlation between them. Calculations about how tailgating could lead to rear-end crashes were presented in his study and the results proved that a safe vehicle headway should be at least two seconds. Although he mentioned that inattention in various forms is a more frequent cause than tailgating as shown from some rear-end crash investigations, inattention could naturally lead to tailgating. Measures to counter tailgating such as advisory signs, pavement markings, and enforcement by the police could help reduce rear-end crashes and were recommended in his study.

Rama and Kulmala (2000) conducted a field study in Finland and investigated the effects of two variable message signs (VMS) on drivers' car-following behavior. The signs warned of slippery road conditions and to keep a minimum following distance. It was performed as a before-and-after experiment at three test sites. Results showed that the slippery road conditions sign reduced the mean speed by one-two km/hour in addition to the decrease caused by the adverse road conditions. The minimum following distance sign reduced the proportion of cars with following distance of less than 1.5 seconds, in addition to a speed reduction of one km/hour.

To help drivers gauge their following distances, research was conducted to assess the effects of regularly-spaced markings on highway pavement. Lertworawanich (2006, 2009) conducted a study to estimate safe car-following distance according to speed limit, and developed the "dot" treatment pavement markings. Headways in term of distance were examined before and after the implementation of "dot" markings. He found that headways were increased after the marking implementation at a given flow rate and the likelihood of rear-end collisions was reduced at the study site. Arrows spaced 131 feet apart, implying a gap of about three seconds at 60 miles per hour, were painted on a U.K. motorway in a study by Helliar-Symons, Webster, and Skinner (1995). They found that, because of the markings, crashes were reduced by 56% at the study site.

A few tailgating treatment programs were pilot-tested in several states such as Pennsylvania, Minnesota, and Maryland. PENNDOT's Tailgating Treatment Program (Safety Improvements) was considered the most successful and was honored in 2001 with the National Highway Safety Award. On a portion of US Route 11 that previously experienced high rates of tailgating, aggressive driving and tailgating has dropped a significant 60% after equipping with reflective dots on the roadway, pavement markings, and signs that help motorists gauge their distance behind moving vehicles (Roadway Safety Foundation 2001). Before the implementation, there were 135 crashes a year costing approximately $1.9 million. After the implementation, yearly crashes decreased to 60 at a reduced cost of $1.3 million. The cost of implementation in the first year is estimated at just over $11,000, including enforcement. After eight to nine months, statistics indicated that crash reductions remained fairly constant, pointing to the success of the program.

Given the successes, relatively low implementation cost, and the measurable benefits of the PENNDOT program, Minnesota DOT and Public Safety piloted a similar project in 2006. The project was viewed as a tool to educate motorists on how to identify and maintain a minimum safe following distance, and to ultimately reduce rear-end crashes. Minnesota used similar engineering elements from the Pennsylvania program: elliptical pavement dots, informational signs, and a strong public information campaign. A section of State Highway 55 in Wright County was painted with 94 elliptical dots, spaced 223 feet apart, along a two-mile single-lane segment of the rural roadway with a 55 mph speed limit. Vehicle headway data collected prior to and after the treatments showed that the average headway increased from 2.36 to 2.62 seconds, or 22.89 feet at the mid-point of the test site (Minnesota Department of Public Safety 2006).

Michael, Leeming, and Dwyer (2000) implemented a method to collect tailgating data in an urban setting and assessed the effectiveness of two hand-held roadside signs admonishing drivers not to tailgate. One was "Please Don't Tailgate" sign and the other "Help Prevent Crashes Please Don't Tailgate." Data collected from over 25,000 drivers were studied. They found that the latter had a significantly positive impact on drivers' tailgating behavior compared with the former sign, increasing the average headway by 0.18 seconds.

Advisory signs could be part of tailgating treatments to help mitigate tailgating behavior. However, improper use of advisory message signs could distract drivers and cause inattention leading to rear-end crashes. To reduce the risk of distracting drivers, the use of graphical images to convey the meaning on roadway signs has been employed in many European countries. It is found in many studies that graphically presented information allowed faster responses than information presented by words (Staplin, Lococo, and Sim 1990; Hanowski and Kantowitz 1997; Bruce, Boehm-Davis, and Mahach 2000). Wang et al. (2007) conducted a pioneer study on the use of graphics on the dynamic message sign (DMS), a 1024x270 full matrix tri-color LED overhead message sign that is capable of displaying pre-programmed text and graphic messages. They found that most drivers preferred graphics over text and responded faster to graphic-aided messages than text-only messages. Due to these findings, proper graphics need to be designed and integrated into advisory signs in the development of feasible tailgating treatment systems.

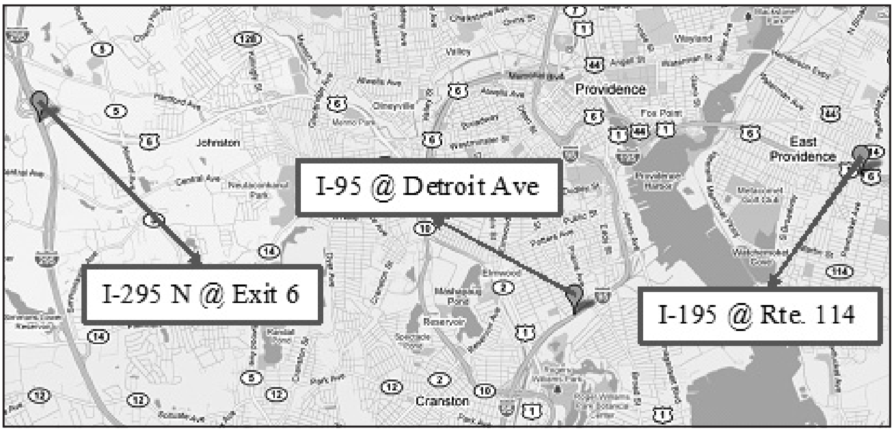
DESCRIPTION OF THE STUDY

Two approaches were employed in this study: a vehicle headway analysis to assess the tailgating situation on highways, and a questionnaire survey to find out the public's perceptions on tailgating and their preferences on possible treatments.

Vehicle Headway Analysis

In this approach, traffic at three test sites on I-95, I-195, and I-295 in Rhode Island were analyzed. The study examined highway traffic surveillance videos that captured eight-lane traffic taken by three highway surveillance cameras on I-95 at Detroit Ave, I-195 at Rte. 114, and I-295 North at Exit 6 (see Figure 1).

Figure 1: Locations of Highway Traffic Surveillance Cameras



Videos taken during both rush hour and non-rush hour on weekdays during a two-week period in December 2008 were provided by Rhode Island Department of Transportation (RIDOT). Videos taken on weekdays between 7:30 and 8:00 AM during rush hour and between 10:00 and 10:30 AM during non-rush hour were analyzed. Three 5-minute segments were randomly selected from each of the 30-minute video clips in the analysis. To determine the vehicle headway for a vehicle, a reference line was drawn in the recorded scene, and based on the time stamp (in 1/100 second) embedded in the video, the time when the front bumper of a vehicle reaching the reference line was recorded. By calculating the time difference that two consecutive vehicles crossed the reference line, vehicle headway in hundredths of a second was determined for the following vehicle. Traffic volumes were also collected by counting all the vehicles that appeared in the surveillance videos during the randomly chosen 5-minute time intervals. The traffic volumes confirmed the selection of the time periods for rush hour and non-rush hour traffic.

All vehicles analyzed were classified by their vehicle headways in increments of seconds. Percentages of vehicles that broke the two-second rule, which means their headways were between zero and two seconds, were calculated. These vehicles were noted as tailgators. To find out if time of the day was a significant factor affecting tailgating, a hypotheses test using a paired t-test was employed to compare the tailgating situations (percentages of tailgators) between rush hour period and non-rush hour period on three test sites. The hypotheses are:

$$H_0 : \mu_d = 0$$

$$H_1 : \mu_d > 0$$

where $\mu_d = \mu_{\text{rush hour}} - \mu_{\text{non-rush hour.}}$

These analyses were further stratified by lane and bound (all lanes in the same direction) since all three highway segments were eight-lane highways with traffic in opposite directions (I-295 South bound was not well captured by the highway surveillance camera, thus its data were missing in this study).

In addition, a correlation analysis was conducted to investigate the functional relationship between tailgating and traffic volume. The percentages of tailgators at the randomly chosen time intervals during both rush hour and non-rush hour were regressed against respective traffic volumes in these time intervals. Analyses were made on different lanes as well as on different test sites.

Questionnaire Survey

Following the vehicle headway analysis, a questionnaire survey was designed and deployed in two phases to help find the causes of tailgating and to identify drivers' preferences on a few proposed tailgating treatments. The first phase was designed to identify the causal factors of tailgating and to gain insights about drivers' experiences and perceptions regarding tailgating on Rhode Island highways. Various components of tailgating treatments were presented to drivers on the second phase to obtain drivers' preferences on potential tailgating treatment systems. Some components presented in the survey were adopted from existing tailgating treatments in the United States as well as those currently used in some European countries.

The questionnaire survey was designed using Microsoft PowerPoint® and Visual Basic macros to present the questions and to collect subjects' answers on laptop computers. Questions in the survey might require either a single or multiple answers. The first phase of the survey contained 19 questions designed to collect drivers' opinions on tailgating and to find out drivers' perceptions of tailgating and its causes and effects. It also surveyed drivers' behavior when they were following other vehicles or being followed. The second phase of the survey was designed to find drivers' preferences on components of a tailgating treatment system. These components include: reference marking (pavement marking, roadside marking) and advisory message sign (dynamic message sign, variable message sign, or fixed road sign). This phase was presented with both auditory and written instructions starting with information about the proper headway to maintain while driving on highways. Six questions were presented in the second phase. The first question surveyed a subject's preference on reference markings. Four simulated driving videos with different reference markings built in (painted dots, painted arrows, neon/hot-pink roadside panels, and painted bars, as shown in Figure 2) were presented in a random sequence. According to the subject's answer, the preferred marking would then be incorporated in subsequent questions. The second question showed several text messages on fixed signs that could be used in conjunction with the markings (Figure 3). The third question inquired a driver's preference regarding graphic-aided messages. The subject was asked to choose between the advisory sign with only the preferred text message and one integrated with graphics (Figure 4). Questions four and five were essentially the same as the second and the third but the road-side fixed signs were replaced with overhead dynamic message signs. The last question was designed to capture a subject's preference on the type of advisory sign to be used in conjunction with the selected reference marking. By completing the second phase of the survey, a preferred tailgating treatment system would be identified.

The two phases of the survey were conducted at multiple locations in Rhode Island in order to obtain a representative sample of the Rhode Island driving population. The University of Rhode Island and the Warwick Mall were among several sites where the survey took place. Subjects with a valid driver's license were randomly recruited at the survey sites with voluntary participation. Prior to beginning the survey, each participant read and gave their consent on an electronic consent form, approved by the university's Institutional Review Board. The subject would then start taking the survey presented as PowerPoint slides on a laptop computer. Survey questions were presented one at a time with no time limit.[1] Answers could be made via a computer mouse and keyboard or via verbal communication with the survey assistant.

Figure 2: Screenshots of Driving Simulation Videos in the Questionnaire

A total of 210 subjects participated in the first phase. Among them, 91 (43.3%) were between 18 and 40 years old, 72 (34.3%) were between 41 and 60, and 47 (22.4%) were older than 60; there were 107 females (51.0%) and 103 males (49.0%). A total of 142 subjects participated in the second phase. Among them, 76 were females (53.5%) and 66 were males (46.5%); 63 (44.4%) were between 18 and 40 years old, 45 (31.7%) between 41 and 61, and 34 (23.9%) were older than 61. Age and gender percentages of the survey resembled the Rhode Island population.

Figure 3: A Sample Survey Question

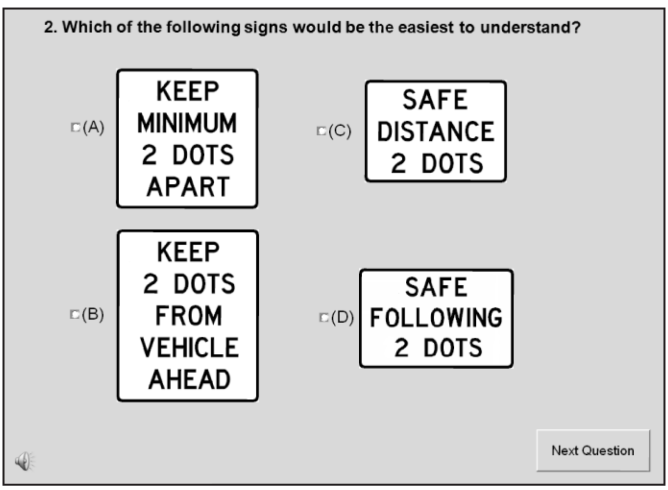
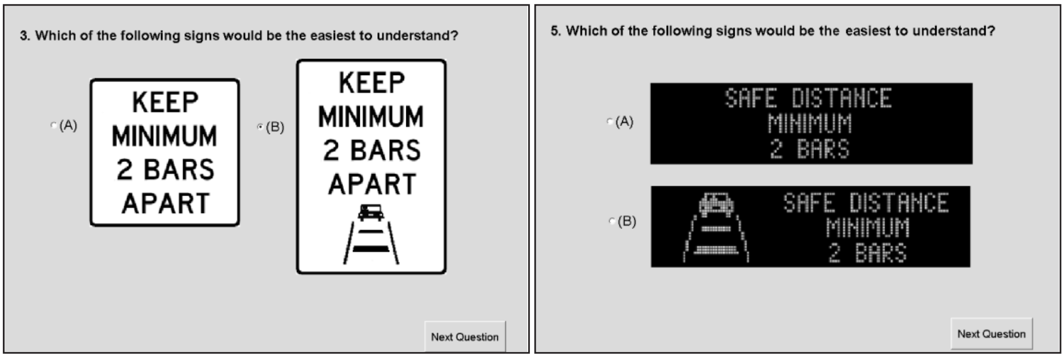


Figure 4: Text Signs and Graphic-aided Message Signs



RESULTS AND DISCUSSION

Vehicle Headway Analysis

The proportions of vehicles following with less than two seconds of vehicle headway, i.e., tailgators, on the three test sites on Rhode Island interstate highways were tabulated in Table 1. The statistics were shown by test sites, by day of the week, and by time of the day. From the analysis, 61.23% of vehicles were tailgating during rush hour and 39.21% during non-rush hour. Tailgating situations during rush hour and non-rush hour were compared through paired t-tests and were found to be significantly different (p values = 0) at all three locations. The rush-hour and non-rush hour tailgating percentages on different days of the week were not significantly different. Small variations in total tailgating percentage during each of these two periods were observed among different days of the week. This also held true within each site. The rush-hour tailgating percentage ranged from 62.24% to 70.96% on I-95, from 56.81% to 62.59% on I-195, and from 46.28% to 56.50% on I-295. The non-rush-hour percentage ranged from 40.62% to 45.24% on I-95, from 33.04% to 38.91% on I-195, and from 19.76% to 27.99% on I-295. The percentages on I-295 were lower than the other two sites since its traffic volume is much lower.

The distributions of vehicle headways for both rush hour and non-rush hour at the three test sites are shown in Figure 5. Vehicle headways collected ranged from less than one second to more than 30 seconds. It should be noted that large vehicle headways were not generally considered "following" and thus the distributions displayed here included only up to 10 seconds of vehicle headways. It is noticed that the majority drove with vehicle headways of less than two seconds during rush hour where more than half of that group were following with headways of less than one second. During non-rush hour, less tailgating behaviors were observed and most occurred in the one-second to two-second interval.

To further assess the tailgating situations, the percentages of the tailgators were classified by lane and bound and were tabulated in Table 2. It showed that vehicles in the high speed (innermost) lane exhibited the worst tailgating situation (highest tailgating percentage during rush hour) while the outermost lane had the lowest tailgating percentage (except the test site on I-295). This could be due to the fact that tailgating is correlated with speed, and vehicles travelling in high-speed lanes tend to follow the leading vehicles. Compared with their opposite bounds, higher percentages of tailgators were mostly observed on I-95 north bound and I-195 west bound, especially during rush hour. This might be due to the large vehicle volumes entering the Providence metropolitan area during rush hour as observed in surveillance videos (see traffic volume data in Table 2).

Table 1: Percentages of Tailgators

	Highway sections		I-95 @ Detroit Ave	I-195 @ Rte. 114	I-295 N @ Ex. 6	Total
Day of the week	Monday	RH*	70.96%	62.59%	54.58%	64.24%
		NRH*	45.24%	37.35%	25.50%	40.61%
	Tuesday	RH	66.54%	57.39%	56.50%	61.52%
		NRH	44.62%	34.46%	27.17%	39.79%
	Wednesday	RH	64.22%	60.92%	46.28%	58.85%
		NRH	41.58%	38.91%	27.99%	38.65%
	Thursday	RH	67.14%	56.81%	47.87%	63.84%
		NRH	43.71%	35.64%	19.76%	39.59%
	Friday	RH	62.24%	59.52%	53.09%	59.17%
		NRH	40.62%	33.04%	26.16%	37.02%
P value (α= 0.05)			0.000	0.000	0.000	
Total		RH	66.93%	60.50%	51.81%	61.23%
		NRH	45.52%	39.00%	25.31%	39.21%

*RH: Rush Hour, NRH: Non-rush Hour

It was noted that the tailgating situation was more serious during rush hour than non-rush hour due to the significant differences in their traffic volumes. A correlation between the percentage of tailgators and the traffic volume could exist. Therefore, linear regression analyses were next conducted to investigate the functional relationship between tailgating and traffic volume at all test sites. Figure 6 showed that strong correlations existed at all three test sites since their R^2 values were all greater than 0.85. It could be concluded that tailgating percentage increases as traffic volume increases among the three test sites. I-295 exhibited the most serious tailgating situation as it had the largest slope.

To assess the correlations by different lanes, separate regression analyses were conducted and the results are shown in Figure 7.

As seen from Figure 7, strong correlations existed between tailgating and traffic volume in all lanes (R^2 ranging from 73.7% to 84.3%). Although the innermost lane had the lowest slope among all four lanes, its large intercept value (about twice the value of any other lane) made it the lane with the most serious tailgating problem.

The vehicle headway analysis provided strong evidence that serious tailgating occurred on major highways in Rhode Island and posed severe traffic safety concerns for highway driving. This finding strongly supported the need of a tailgating treatment system on Rhode Island highways. Since the tailgating situation was worse in the innermost lane, especially during rush hour, it might be more effective to place tailgating treatments, including roadside markings and advisory signs, on the left side of the innermost lane to mitigate tailgating behavior.

Figure 5: Distributions of Vehicle Headways

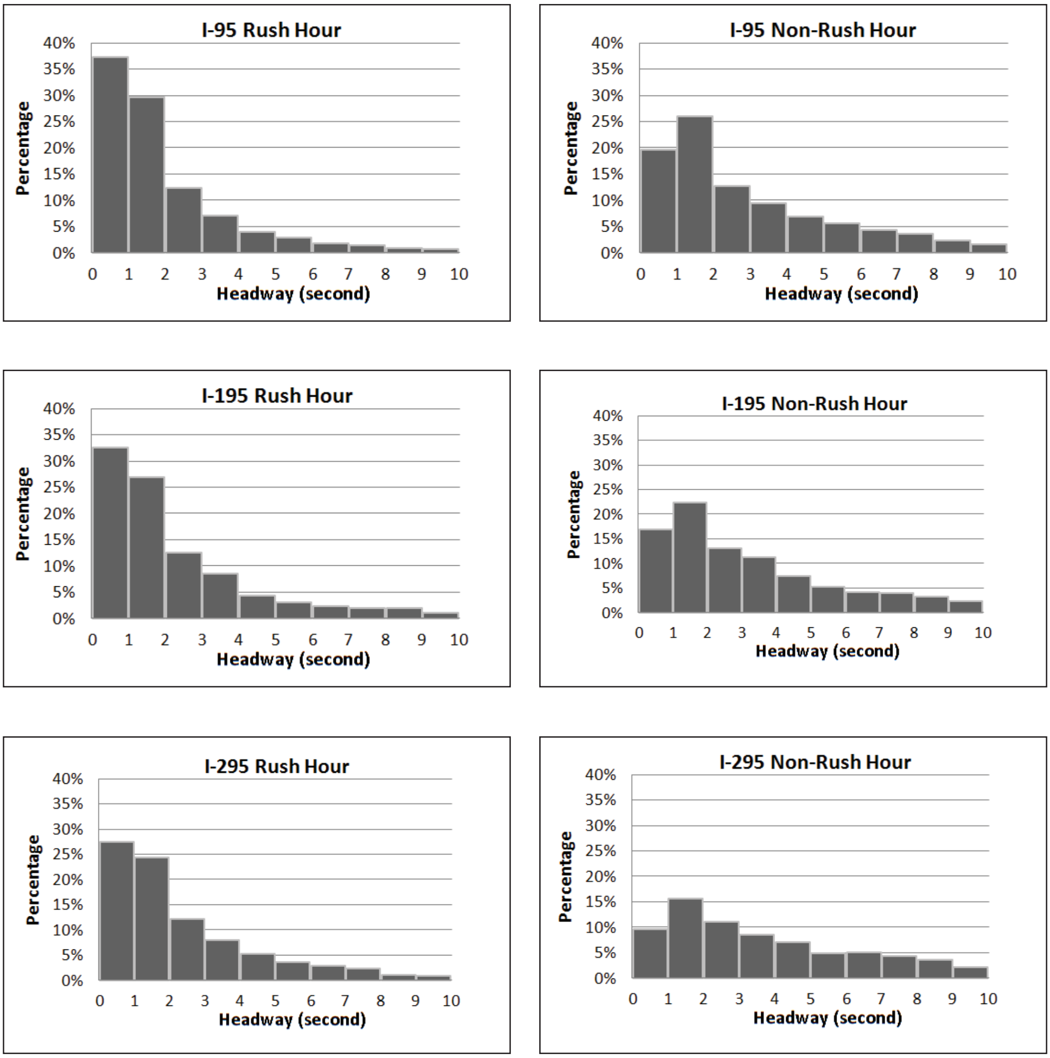


Table 2: Percentages of Tailgators and Traffic Volumes by Lane and Bound

Lane	Time of the day	Highway Test Sites									
		I-95 @ Detroit Ave				I-195 @ Rte. 114				I-295 N @ Ex. 6	
		North		South		West		East		North	
		Percentage	Volume*	Percentage	Volume	Percentage	Volume	Percentage	Volume	Percentage	Volume
Innermost lane	RH	79.62%	40.13	70.27%	28.40	72.38%	30.17	67.76%	18.57	70.27%	33.28
	NRH	47.46%	10.60	46.09%	13.26	43.14%	11.90	36.47%	8.50	31.21%	8.62
2nd Lane	RH	72.96%	42.40	63.76%	29.07	65.84%	32.10	59.49%	22.30	28.08%	12.79
	NRH	46.67%	19.47	47.06%	18.33	43.35%	16.53	36.97%	14.06	17.68%	8.73
3rd Lane	RH	68.28%	38.47	53.93%	24.60	55.36%	28.00	55.81%	20.37	16.12%	5.21
	NRH	54.43%	21.13	45.31%	16.67	33.59%	12.80	36.62%	14.20	21.67%	5.48
Outermost lane	RH	62.58%	33.13	41.35%	13.87	52.78%	22.17	36.58%	9.70	44.62%	17.35
	NRH	41.78%	23.00	33.93%	15.33	26.55%	9.67	17.37%	5.57	29.97%	7.64
Total	RH	71.29%	154.13	59.93%	95.94	62.41%	112.44	57.47%	70.94	51.81%	68.63
	NRH	47.48%	74.20	43.23%	63.59	37.65%	50.90	34.17%	42.33	25.31%	30.47

*Traffic volume was measured in vehicles per minute (vpm).

Figure 6: Correlation Plots Between Percentage of Tailgators (P) and Traffic Volume (V) at Three Test Sites

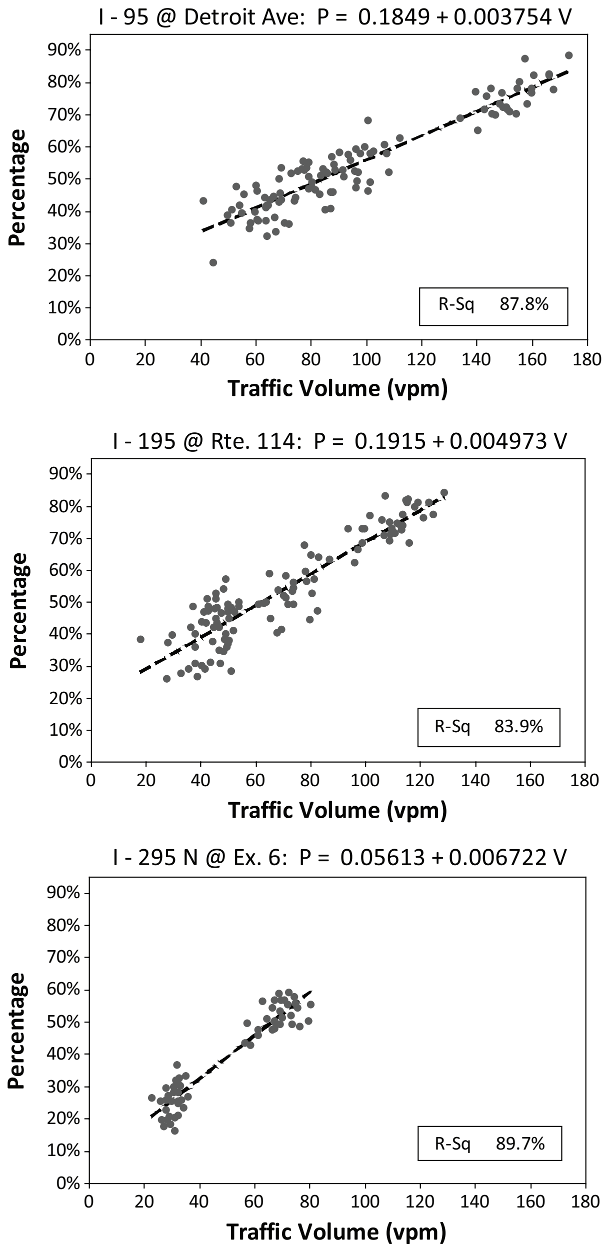
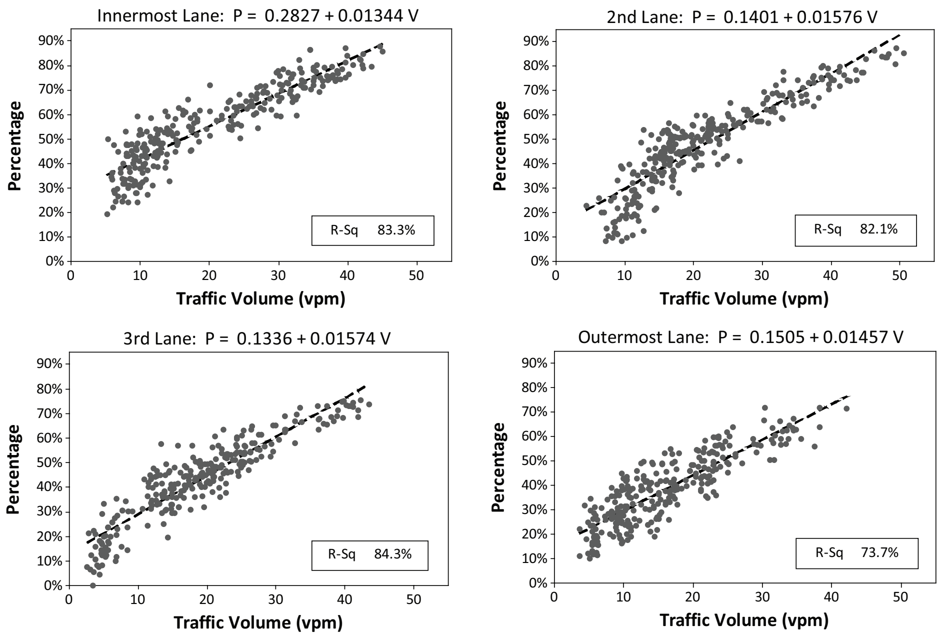


Figure 7: Correlation Plots Between Percentage of Tailgators (P) and Traffic Volume (V) in Four Different Lanes



Questionnaire Survey

With the tailgating problem confirmed on Rhode Island highways, drivers' understandings and perceptions of the tailgating issues and their preferences for components of a potential tailgating treatment system were next collected in the questionnaire. The participants were first asked in the survey to select and rank the top three causes of crashes among 13. According to the weighted scores (3 points for the first ranked, 2 points for second and so forth), the top three leading causes of crashes were distraction, speeding, and tailgating (Figure 8).

When subjects were asked about the possible reasons for following other vehicles, the majority (70.0%) indicated that they usually did not follow others; when asked if they intentionally followed other vehicles while driving on highways, 67.6% of subjects indicated that they never did; when asked about the reasons for not following other vehicles, 49.5% of subjects indicated that it's safer not following others (Table 3).[2]

When asked about the definition which best described tailgating, 76.2% of the subjects chose "following too close to the vehicle ahead" and only 11.4% chose "insufficient following distance." This indicated that most drivers had only a qualitative idea of what tailgating means instead of a quantitative one. "Heavy traffic," "slow car ahead of my vehicle," and "I am in a hurry" were among the top three choices of tailgating causes (Table 4).

Figure 8: Pareto Chart of Major Causes of Crashes

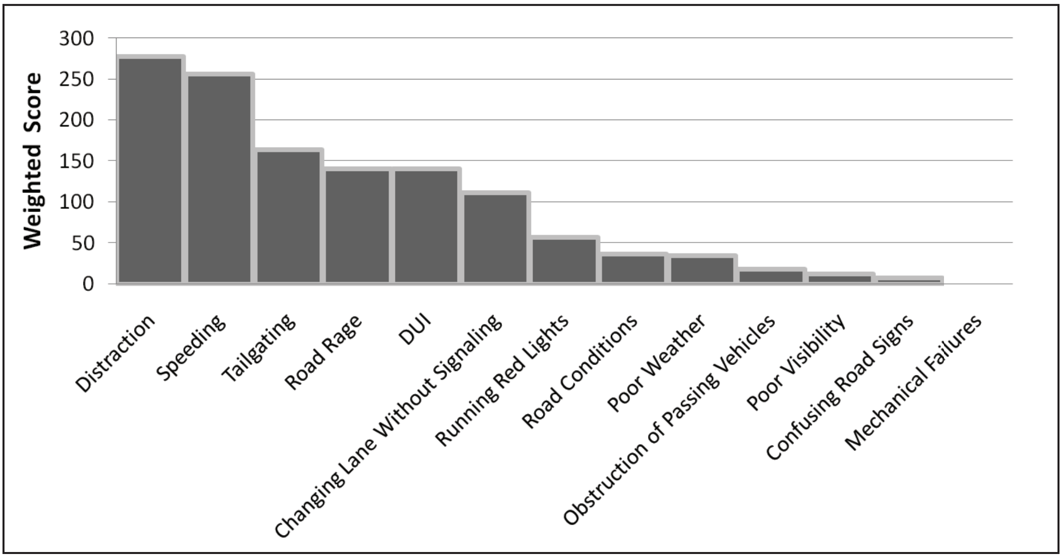


Table 3: Survey Questions Regarding Following Behavior on Highways

Why would you follow other vehicles while driving on highways?	It is easier to maintain speed.	13.8%
	It is safer.	1.4%
	It is less likely to get a speeding ticket.	14.8%
	I don't do that.	70.0%
Do you intentionally follow other vehicles while driving on highways?	Always.	0.0%
	Sometimes.	32.4%
	Never.	67.6%
Why would you <u>not</u> follow other vehicles while driving on highways?	It obstructs my view.	17.1%
	Others drive too slow.	3.4%
	It is safer.	49.5%
	It is against the law.	30.0%

Table 4: Applicable Causes of Tailgating (Multiple-choice)

Which of the following could cause you to tailgate while driving on highways? Mark all that apply.

Heavy traffic	Slow car ahead of my vehicle	In a hurry	Poor visibility	Distraction	Weather conditions	Hypermiling*
36.1%	20.7%	19.5%	5.8%	9.8%	7.1%	1.0%

* The act of driving using techniques that maximize fuel economy such as following a truck closely to reduce wind resistance.

When asked about their reactions when being tailgated, most indicated they were affected by tailgators, and most drivers affected by tailgators reacted passively (Table 5). The top choice was "change lanes to let the tailgator pass" (34.1%).

Table 5: Reactions When Being Tailgated (Multiple-choice)

How would you react if you were followed too closely? Mark all that apply.						
Change lanes to let the tailgater pass	Slow down to force the tailgater to get away	Speed up	Make a gesture at the tailgator	Ignore the tailgator	Tap the brake	Call the police
34.1%	17.2%	12.1%	4.9%	15.7%	12.5%	3.6%

When asked about safe vehicle headways, the majority of subjects (77.1%) indicated they knew what the proper vehicle headway was, 73.8% indicated that keeping a safe vehicle headway was very important, and 90.5% believed they always kept a safe vehicle headway or at least most of the time (Table 6). Along with the results shown in Table 3, it did not appear that serious tailgating problem existed on Rhode Island highways.

Table 6: Survey Questions Regarding Safe Vehicle Headway

Do you know the proper vehicle headway to maintain while driving on highways?			
Yes	77.1%	No	22.9%

How important is it for you to keep a safe vehicle headway while driving on highways?				
Very important	Important	Average	Slightly important	Not important
73.8%	17.6%	2.4%	6.2%	0.0%

Do you maintain a safe vehicle headway while driving?				
Always	Most of the time	Sometimes	Rarely	Never
35.7%	54.8%	8.6%	1.0%	0.0%

When questioned "how much distance do you maintain when driving at 60 mph on highways," 95% of subjects indicated they maintained a vehicle headway less than 11 car lengths, and almost half maintained less than four car lengths (Figure 9). These answers exposed a serious tailgating issue by showing that the majority of drivers who took the survey did not know what the proper vehicle headway was, and drove with insufficient following distance. When driving at 60 mph, a two-second vehicle headway requires a following distance of 11 car lengths (assuming a car length of 15 feet). Although 75.2% of subjects indicated in another question that they maintained a vehicle headway equal to or greater than three seconds, it is not likely they kept three seconds of headway, which is about 16 car lengths. Subjects' opinions on vehicle headway expressed in car lengths could be more reliable since 78.6% of them preferred using car lengths to measure vehicle headway. This finding from the survey, in fact, did not contradict but confirmed the serious tailgating situation identified in the vehicle headway analysis.

When asked about components in an effective tailgating treatment system, 63.8% of subjects preferred a combination of both advisory message signs and reference markings. A total of 37.3% of subjects selected the painted horizontal bars as pavement markings, followed by painted arrows and neon panels (Table 7). Some thought the roadside neon panels were distracting while others felt that the arrows might suggest drivers to speed up. Compared with other age groups, drivers over 60 preferred the painted arrows (32.3%) over painted bars (29.4%) although the difference was not significant. Female and male drivers did not show any difference in their answers.

Figure 9: Vehicle Headways Maintained by Drivers When Driving at 60 mph

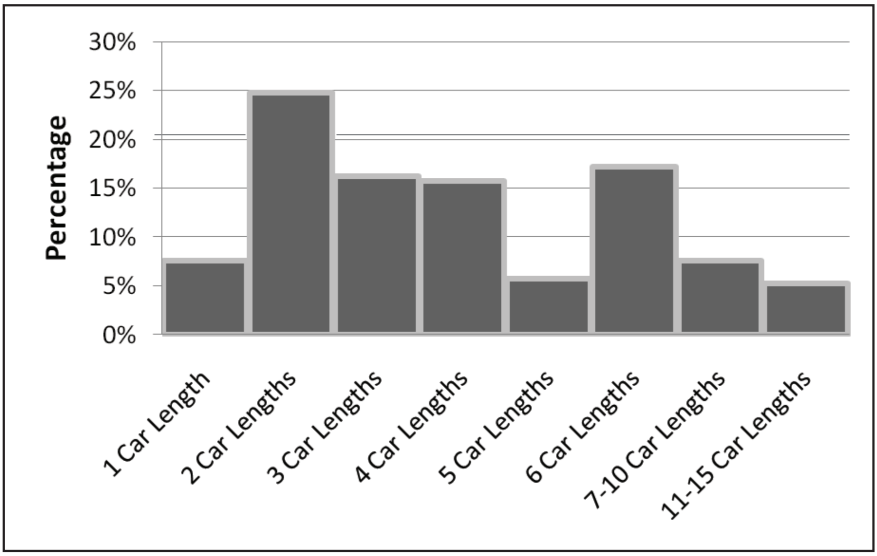


Table 7: Subjects' Preference on Reference Markings

Which of the following would be most effective in helping you maintain a proper following distance?	Painted dots	9.4%
	Painted arrows	33.2%
	Neon panels	20.1%
	Painted bars	37.3%

When asked which sign message would be the easiest to understand, most subjects (40.2%) chose the lengthiest-worded one (for example, "Keep 2 bars from vehicle ahead") over others. "Safe distance 2 bars" ranked second (Table 8). This might be due to the fact that subjects were in a static environment where they could take as much time as they want to read the messages, and in that case, more information in the message helped their understanding. Since lengthier messages could demand more drivers' attention in real driving, graphics could be used to help drivers better understand the message while demanding less attention. When graphics were added to text sign message, graphic-aided message signs were mostly preferred (86.6%) over text-only ones.

Table 8: Subjects' Preference on Messages (If Painted Bars Were Preferred)

Which of the following signs would be easiest to understand?	Keep Minimum 2 Bars Apart	21.8%
	Keep 2 Bars from Vehicle Ahead	40.2%
	Safe Distance 2 Bars	33.1%
	Safe Following 2 Bars	4.9%

When similar questions were asked with regard to overhead dynamic message signs, similar responses were obtained. Among responders, 45.1% preferred the sign with the lengthiest wording and 81.7% preferred it with graphics. When asked about which traffic sign drivers most likely would pay attention to while driving, subjects preferred the overhead dynamic message signs (46.7%) over the fixed road signs (29.6%) and the roadside variable message signs (23.9%).

The findings of the first phase survey showed that the majority of drivers considered tailgating a serious offense and one of the top three major causes of highway crashes. Most of them, however, had only a qualitative sense of what tailgating was about and did not know what the proper vehicle headway was while following other vehicles on highways. As most indicated in the survey, they did not maintain sufficient vehicle headways. This finding confirmed the observations made in the vehicle headway analysis. From the second phase of the survey, it was found that the majority preferred regularly-spaced horizontal bars painted on pavement as reference marking to help drivers gauge safe following distance. Coupled with the pavement marking, most of them preferred employing overhead graphic-aided dynamic message signs to communicate to drivers about safe following distance.

CONCLUSION AND FUTURE WORK

This study examined the tailgating issues in Rhode Island, identified causes and effects of tailgating, and surveyed drivers' preferences on possible tailgating treatments. Through the vehicle headway analysis, a serious tailgating situation (61.23% vehicles were tailgating) was identified during rush hour. Less tailgating was observed during non-rush hour but there were still 39.21% of vehicles following with insufficient headways. Tailgating percentages classified by lane and bound showed that tailgating was worse in the innermost lanes and on bounds with high traffic volume. Tailgating percentages were correlated with the traffic volumes in the correlation analysis, and the serious tailgating situation in innermost lanes was confirmed. The results of the vehicle headway analysis suggested a need for an effective tailgating treatment system in Rhode Island.

The findings of the survey indicated that the majority did not know what the proper vehicle headway was and maintained insufficient vehicle headways while following other vehicles on Rhode Island highways. This confirmed the observations made in the vehicle headway analysis. Among the proposed tailgating treatments, the majority preferred horizontal bars painted on pavement as a means to help maintain safe following distance, and overhead graphic-aided dynamic message signs as a way to advise drivers about safe following distance. Based on this study, it is concluded that a proper tailgating treatment system should be developed and implemented on Rhode Island highways. Survey results suggested horizontal bar pavement markings and overhead dynamic message signs displaying graphic-aided advisory messages.

The components of a tailgating treatment system surveyed will be further examined in a follow-up study where a driving simulation experiment will be designed and conducted. Results obtained from the driving simulation, coupled with the findings from the survey, will lead to the recommendation of an effective tailgating treatment system. Field studies will be conducted after the implementation of the recommended system. Before and after statistics would be collected and analyzed to assess the effectiveness of the system in mitigating tailgating behavior and reducing rear-end crashes. The findings could contribute to the development of a standard tailgating treatment system to be included in MUTCD and help facilitate more efficient and safer driving on U.S. highways.

Acknowledgements

The authors wish to thank the Rhode Island Department of Transportation (RIDOT) and the University of Rhode Island Transportation Center (URITC) for their support and guidance on this study.

Endnotes

1. No time constraint was given since the questionnaire was designed to capture as much of the subjects' responses as possible at this stage.

2. The word "following" was used in the first few questions in the first phase of the survey to investigate drivers' general following behavior, including both tailgating and non-tailgating.

References

AASHO. “A Policy on Design of Urban Highways and Arterial Streets.” American Association of State Highway Officials, Washington, D.C., 1973.

Aycin, Murat F. and Rahim F. Benekohal. “Linear Acceleration Car-following Model Development and Validation.” *Transportation Research Record* 1644, (1998): 10-19.

Ayres, T.J., L. Li, D. Schleuning, and D. Young. “Preferred Time-Headway of Highway Drivers.” IEEE Intelligent Transportation Systems Proceedings, Oakland, CA, (2001): 826-829.

Brackstone, Mark and Mike McDonald. “Car Following: A Historical Review.” *Transportation Research Part F: Traffic Psychology and Behaviour 1* (2), (1999): 181-196.

Brackstone, Mark. “Driver Psychological Types and Car Following: Is There a Correlation? Results of a Pilot Study.” Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, (2003): 245-250.

Bruce, Deborah, Deborah A. Boehm-Davis, and Karen Mahach. “In-Vehicle Auditory Display of Symbolic Information.” Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society: Ergonomics for the New Millennium, San Diego, CA, (2000): 230-233.

Carter, Chris, Andrew J. May, Fred J. Smith, and Stephen H. Fairclough. “An Evaluation of an In-Vehicle Headway Feedback System.” *Contemporary Ergonomics*, (1995): 287-292.

Dingus, Thomas A., et al. “Human Factors Field Evaluation of Automotive Headway Maintenance/Collision Warning Devices.” *Human Factor 39* (2), (1997): 216-229.

Green, Marc. “How Long Does it Take to Stop? Methodological Analysis of Driver Perception-Brake Times.” *Transportation Human Factors 2*, (2000): 195-216.

Hanowski, Richard J. and Barry H. Kantowitz. “Driver Memory Retention of In-Vehicle Information System Messages.” *Transportation Research Record* 1573, (1997): 8-18.

Helliar-Symons, Robin D., P. B. Webster, and A. J. Skinner. “The M1 Chevron Trial.” *Traffic Engineering and Control* 36 (10), (1995): 563-567.

Hogema, Jeroen H. “Modeling Motorway Driving Behavior.” *Transportation Research Record* 1689, (1999): 25-32.

Hutchinson, Paul T. “Tailgating.” Safety Research Report # CASR046, Centre for Automotive Safety Research, University of Adelaide, Australia, 2008.

Lertworawanich, Ponlathep. “Safe-Following Distances Based on the Car-Following Model.” PIARC International Seminar on Intelligent Transport System (ITS) in Road Network Operations, Kuala Lumpur, Malaysia, 2006.

Lertworawanich, Ponlathep. “Evaluation of Dot Following-Distance Pavement Markings in Thailand.” Proceedings of the Transportation Research Board 88th Annual Meeting, 2009.

Michael, Paul G., Frank C. Leeming, and William O. Dwyer. “Headway on Urban Streets: Observational Data and an Intervention to Decrease Tailgating.” *Transportation Research Part F: Traffic Psychology and Behaviour 3* (2), (2000): 55-64.

Minnesota Department of Public Safety. *Minnesota Tailgating Pilot Project, Report and Summary*, 2006.

National Highway Traffic Safety Administration. *Traffic Safety Facts 2006 Data: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System*. U.S. Department of Transportation Report no. DOT HS 810 818, 2008.

Rajalin, Sirpa, Sven-Olof Hassel, and Heikki Summala. "Close-following Drivers on Two-lane Highway." *Accident Analysis and Prevention 29*, (1997): 723-729.

Rama, Pirkko and Risto Kulmala. "Effects of Variable Message Signs for Slippery Road Conditions on Driving Speed and Headways." *Transportation Research Part F: Traffic Psychology and Behaviour 3* (2), (2000): 85-94.

Roadway Safety Foundation. "Saving Lives: RSF Recognizes Nine Outstanding Highway Safety Projects." *Road Safety Reporter*, December, 2001.

Singh, Santokh. "Driver Attributes and Rear-end Crash Involvement Propensity." NHTSA Technical Report no. DOT HS 809 540, March, 2003.

Staplin, Loren, Kathy H. Lococo, and James J. Sim. "Traffic Control Design Elements for Accommodating Drivers with Diminished Capacity." Volume II: Final Technical Report. Publication No. FHWA-RD-90-055, Federal Highway Administration, Washington, D.C, 1990.

Summala, Heikki. "Brake Reaction Times and Driver Behavior Analysis." *Transportation Human Factors 2*, (2000): 218-226.

Teigen, Anne. "Transportation Review: Aggressive Driving." National Conference of State Legislatures, February 2007.

Wang, Jyh-Hone, Siamak G. Hesar, and Charles E. Collyer. "Adding Graphics to Dynamic Message Sign Messages." *Transportation Research Record 2018*, (2007): 63-71.

***Miao Song** is a Ph.D. candidate in Industrial and Systems Engineering at the University of Rhode Island. He holds a B.S. degree in automation from the University of Science and Technology of China and an M.S. degree in mechanical engineering from the University of Rhode Island. His research interests include human factors in transportation and human physiology in injury prevention and driving safety.*

***Jyh-Hone Wang** is a full professor in the Department of Mechanical, Industrial and Systems Engineering at the University of Rhode Island. Prior to joining the University of Rhode Island in 1995, he was a faculty member at California State University, Fresno, and Georgia Southern University. Dr. Wang is generally recognized by his research work in human factors and driving safety in transportation. He has conducted many research projects on the intelligent transportation system and its interaction with human drivers, mostly supported by Rhode Island, New England, and federal transportation authorities. He also performed research works in automobile quality engineering and computer integrated manufacturing. He has authored and co-authored over 60 articles, which have appeared in books, technical journals, and conference proceedings. He received his B.S. degree in industrial engineering from Tunghai University in 1980, his M.S. and Ph.D. degrees in industrial engineering from the University of Iowa in 1986 and 1989.*

Railroad Operational Performance in the United States

by Denver Tolliver, John Bitzan, and Doug Benson

This paper examines the effects of increased car weights, train size, and distance on railroad operational performance, which is measured in revenue ton-miles per train hour. A statistical model is estimated from Class I data that explains most of the variation in this index, while controlling for commodity and fixed network effects. The results suggest that length of haul, net load, and train size contributed significantly to performance gains from 1987 to 2006. However, train miles per track mile increased by 89% during the same period, resulting in line and yard congestion.

Introduction

In 2006, Class I railroads moved 32 million carloads of freight (including 12.3 million trailers and containers) and provided 1.77 trillion ton-miles of service (AAR 2007). Transportation agencies in the United States are hoping to limit the growth of highway congestion in key intercity corridors by keeping freight on rail lines instead of highways (AASHTO 2003). From a multimodal perspective, it is important to understand the factors that have contributed to railroad performance.

Railroad performance could be defined in many ways. In this paper, we have selected a measure that has historically been used by the American Association of Railroads (AAR): net or revenue ton-miles per train-hour. This measure reflects train speed and revenue tonnage and is a composite measure of the capability to efficiently move freight in intercity corridors. Other potential indicators (such as speed, tonnage, or congestion) measure only one aspect of railroad performance.

Many studies have addressed railroad productivity and costs, including Lee and Baumel (1987); Berndt, Friedlaender, Chiang, and Vellturo (1993); Wilson (1997); Martland (1999); Oum, Waters, and Yu (1993); and Bitzan and Keeler (2003). While this paper benefits from this research, it does not attempt to measure railroad multi-factor productivity or cost. Instead, the primary objective is to show the effects of increased car weights, train lengths, and shipment distances on line-haul performance. A better understanding of these factors is important in assessing the performance of surface transportation modes within multimodal corridors.

Performance Index

The Association of American Railroads publishes an index: *net ton-miles per train hour*. This index "reflects both the number of tons hauled and the miles traveled during an average hour of a freight train's operation" (AAR 2007). The numerator of this index (net or revenue ton-miles) is a measure of railroad output. A performance objective is to move a given number of revenue ton-miles with the fewest possible train hours. In doing so, both road and yard hours are important. Most cars require switching and blocking services at origin and destination, as well as at intermediate locations (Figure 1). For this reason, the denominator of our index is *total* road and yard hours (TH), which is different from the AAR's index, which reflects road train hours only.

The trend in revenue ton-miles per train hour (RTM/TH) from 1987 through 2006 is shown in Figure 2. The index increased from 33,108 in 1987 to 40,258 in 1992. The remainder of the period is characterized by ups and downs, with the high value for the period (44,128) being recorded in 2002. The index subsequently declined to 42,618 in 2006. Nevertheless, it increased by 29% from 1987 through 2006.

Figure 1: Abstract Representation of Revenue Ton-Miles/Train Hour

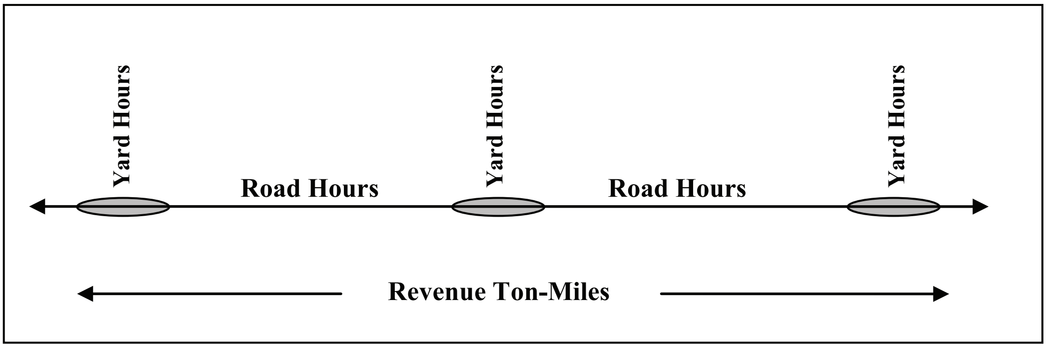
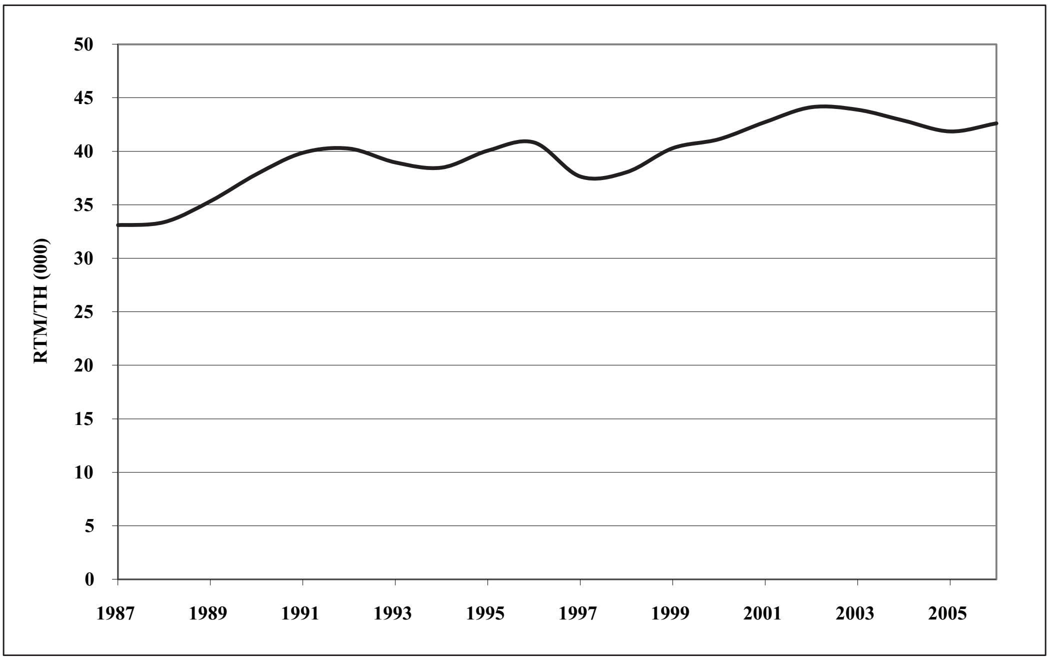


Figure 2: Trend in Operation Performance of Class I Railroads in the United States (Computed from R-1 Reports to the Surface Transportation Board)



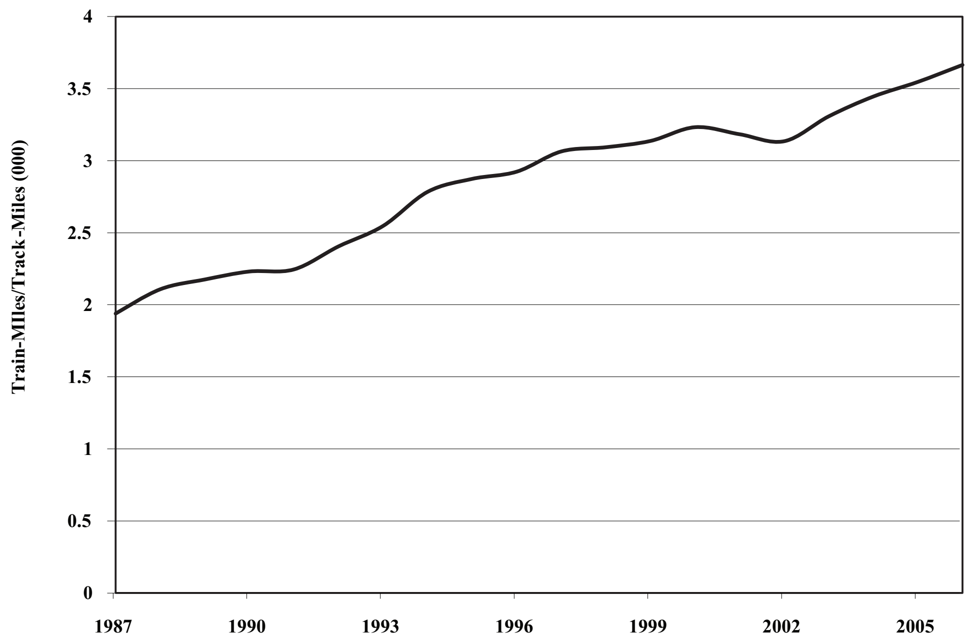
Other performance indicators could have been selected for this study, including train speed, tonnage, or yard dwell time. However, none of these factors are stand-alone or composite indicators. The operational factors (or independent variables) which affect revenue ton-miles per train hour are discussed next.

Density

Density is a measure of utilization of railroad infrastructure. It can be computed by dividing annual train miles by running track miles. Running track miles include first main tracks; second, third, and fourth main tracks; and passing and side tracks. Collectively, these tracks (along with signal aspects[1], block lengths[2], and other traffic control factors) govern the capacities of rail lines.

As shown in Figure 3, train density has risen over time. Train miles per running track mile increased from 1,939 in 1987 to 3,665 in 2006. As traffic increased, railroads added main, passing, and side tracks in high-density corridors. For example, the ratio of second and other main track to first main track increased from 14.7% in 1987 to 19% in 2006 (computed from R-1 reports).[3] However, as the chart shows, train miles have been increasing at a faster pace than track miles.

Figure 3: Trend in Train Density in the United States (Computed from R-1 Reports to the Surface Transportation Board)



While density is beneficial, additional capacity may be needed as traffic grows. A study by Cambridge Systematics, Inc. (2007) indicates that 12% to 13% of U.S. rail lines are at or near capacity. Another 39% have volume/capacity ratios of 0.4 to 0.7. With continued traffic growth, 45% of U.S. rail lines are projected to be near, at, or beyond capacity in 2035 without expansions (Cambridge Systematics, Inc. 2007).

Data Sources

This paper utilizes 20 years of Class I railroad data from R-1 reports, which are submitted annually to the Surface Transportation Board (STB). This database includes 193 observations from 1987 through 2006. Information for the following railroads are used: Atchison, Topeka, & Sante Fe (ATSF), Burlington Northern (BN), BNSF, Chicago & Northwestern (CNW), Conrail (CR), CSX, Grand Trunk Western (GTW), Illinois Central (IC), Kansas City Southern (KCS), Norfolk Southern (NS), Soo Line (SOO), Southern Pacific (SP), and Union Pacific (UP). Some of these railroads were merged during the period and their names are no longer used.

The primary sources of information are Schedules 700 and 755 of the R1 Report and the Quarterly Commodity Statistics (QCS) report. It is a report of the number of carloads and revenue tons handled by each railroad. Operational statistics such as train miles, car miles, and train hours are reported in Schedule 755. Miles of road and running track miles are reported in Schedule 700.[4]

OPERATIONAL PERFORMANCE MODEL

In this section of the paper, a performance model is presented to explain variations in revenue ton-miles per train hour. The model (shown in Equation 1) includes key operational factors such as train size, length of haul, net load, and road and yard congestion.

(1) $RTM/TH = \mathrm{f}\,(ALH, AVCARS, NET, WTR, TD, YD, MR, COMM, RR, T)$

Where:

RTM/TH =	Revenue ton-miles per train hour
ALH =	Average length of haul (miles)
AVCARS =	Average cars per freight train
NET =	Net load per freight car (tons)
WTR =	Average way train trips per carload
TD =	Train density (train miles per running track mile)
YD =	Yard density (locomotive miles per yard track mile)
MR =	Miles of road
COMM =	A vector of commodity variables
RR =	A vector of railroad indicator variables
T =	Time (years since 1987)

The first six variables are of primary importance. The last four are control variables. They are necessary to account for the effects of mergers, rationalization, network effects, and variations in commodity mix (among railroads and across time). The operational effects of the primary variables, and their expected relationships to revenue ton-miles per train hour, are previewed in the following paragraphs.

Average Length of Haul

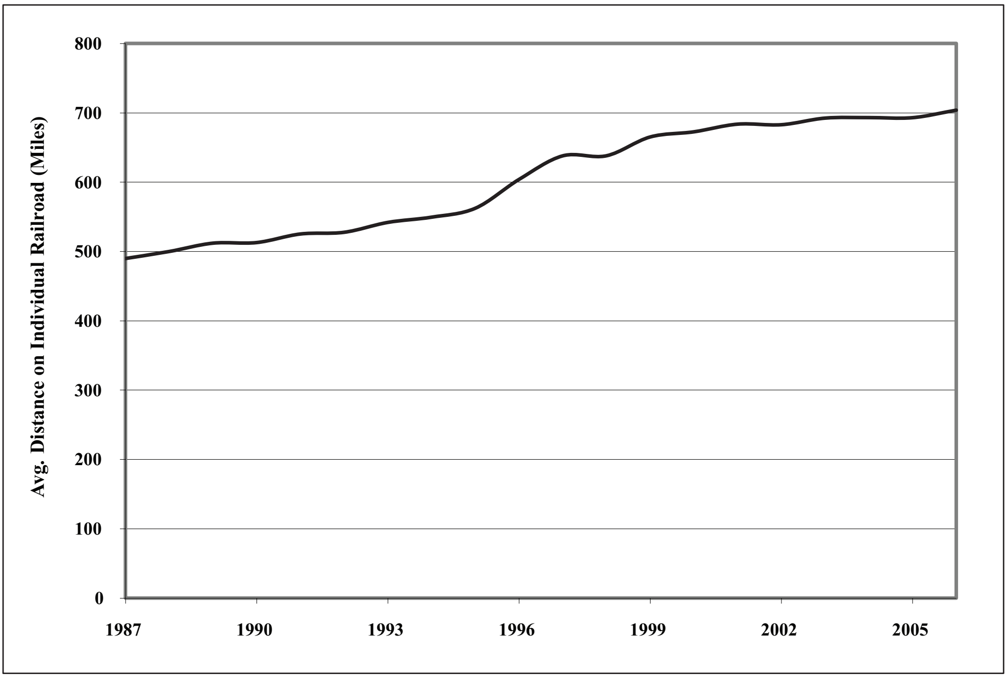
ALH is the average distance that a shipment travels on an individual railroad's network. Because most shipments are interlined, the average distance of a movement on a single railroad is often less than the total shipment distance. As used in this study, ALH is intended to measure the effect of longer hauls on the efficiency of an individual railroad. The sign of ALH is expected to be positive for several reasons: (1) Interchange frequency decreases with mergers that increase ALH. (2) The yard hours required to originate and terminate a shipment are spread over more line-haul miles, resulting in economies of distance. As shown in Figure 4, average length of haul has increased since 1987 as a result of mergers, growth in export/import traffic, and other market factors. The effects of mergers between 1995 and 1999 are apparent in Figure 4.

Average Cars per Train

Maximum train length (i.e., the longest feasible train) is constrained primarily by safety and physical constraints, such as the lengths of side tracks. Maximum train length is the upper bound of average train length. However, average train length (which is measured in cars per train) is influenced by many factors other than maximum length, including network and route densities and commodity mix. Because longer trains utilize time more effectively, AVCARS is expected to have a positive sign.

An analysis of R-1 report data reveals that substantial variations in cars per train were observed during the period—both within and across railroads. Average cars per train ranged from 47.2 to 86.7. The overall mean (for all railroads) was 68.4 cars. The standard deviation was 9.4.

Figure 4: Trend in Average Length of Haul on Individual Railroads in the United States (Source: Computed from R-1 Reports to the Surface Transportation Board)



Net Load

The net load per car (NET) is computed by dividing the revenue tons of freight handled on an individual railroad by the carloads. Logically, this variable is expected to have a positive sign in the model.

An analysis of R-1 report data reveals that substantial variations in tons per car were observed during the 1987-2006 period—both within and across railroads. Average tons ranged from 44.9 to 101.6. The overall weighted mean was 65.1 tons. The standard deviation was 8.0.

Train Service

Generally, railroads operate three types of trains: way, through, and unit. Single cars usually travel in way trains at origin and/or destination. Way trains are small trains that operate primarily between remote stations located on branch lines and railroad yards. They stop frequently at stations to drop off and pick up cars. In contrast, through trains move from yard to yard and perform little or no switching en route. Unit train service is "a specialized shuttle-type service in equipment (railroad or privately owned), dedicated to such service, moving between origin and destination."[5]

Way and through train services are not independent. In most cases, a way train movement precedes or follows a through train movement. However, unit and way train services are mutually exclusive for a given shipment or class of traffic. These relationships are considered in the definition of the variable WTR, which is defined as the ratio of carloads handled in way trains to total carloads originated and terminated. WTR increases with the proportion of carloads handled in non-unit train service. Conversely, WTR decreases as the proportion of carloads moved in unit trains increases. WTR is expected to have a negative sign.

Density

The density variable used in the study (train miles per running track mile) was described earlier. Because increasing densities may reflect line congestion and more train-hours per revenue ton-mile, this variable is expected to have a negative sign. In addition to line congestion, railways may experience yard congestion as facilities are consolidated and traffic grows. In this study, yard density is measured in locomotive-miles per yard track mile (YD). The variable YD is expected to have a negative sign in the model, indicating some yard congestion during the period.

Miles of Road

According to R-1 data, miles of road (or route miles) is the length of all rail lines regardless of the number of tracks. Miles of second, third, and fourth main tracks; crossovers and passing tracks; and side tracks are excluded from this measure. Miles of road can be increased through mergers and decreased through abandonment, consolidation, or sales. Miles of running track can be increased on parts of a carrier's system, while miles of road are decreased on others.

Miles of road and track reflect the miles operated by a given railroad. Because some segments are operated by more than one company, the sum of the miles operated for the industry as a whole may be greater than the sum of the miles owned. However, all activity measures (such as train miles and car miles) reflect train movements over operated segments. Therefore, the track and operational data used in this study are consistent.

According to R-1 data, miles of road decreased from 144,230 in 1987 to 119,684 in 2006 as a result of abandonment, the consolidation of lines after mergers, and the sale of branch lines to local and regional railroads. Over time, traffic has been concentrated on fewer lines. As a result, Class I carriers have been relieved of the time-intensive tasks of picking up and delivering freight on low-density lines. For the most part, these tasks are now performed by local operators with lower overhead costs and more flexible scheduling.

Miles of road should have a negative sign in the model, indicating that the elimination of less productive route miles has improved the efficiency of train operations. However, miles of road may capture several effects. The route miles of some railroads jumped dramatically after mergers. In this respect, MR is an important control variable.

Commodities

The commodity variables are control variables. They do not directly explain operational efficiency. However, they may account for unique industry shipping patterns. Although railroads haul many commodities, 14 Standard Transportation Commodity Code (STCC) groups comprise 99% of the freight tonnage originated by U.S. Class I railroads. A commodity variable is defined for each of the 14 STCC groups shown in Table 1. Each variable represents the commodity's share of total tons handled by a railroad. In addition, a variable is created for STCC group 42: empty trailers and containers. Although these movements are classified as revenue tons, the only revenue cargo is the tare weight of the trailer or container. The 16th variable encompasses all other commodities. To avoid singularity, only 15 of the variables are included in the model.[6]

Accounting for Mergers, Acquisitions, and Other Changes Over Time

Average length of haul (described earlier) accounts for increased efficiencies from longer hauls and reductions in interchange switching—both of which result from end-to-end mergers. However, other effects of mergers (such as network efficiencies) are accounted for through railroad indicator variables.

Table 1: Primary Commodities Transported by U.S. Class I Railroads

Commodity	Percent of Tons Originated
Coal	43.5
Chemicals & allied products	8.6
Farm products	7.6
Non-metallic minerals	7.2
Misc. mixed shipments	6.4
Food & kindred products	5.4
Metals & products	3.2
Metallic ores	3.1
Petroleum & coke	2.8
Stone, clay & glass products	2.7
Waste & scrap materials	2.5
Lumber & wood products	2.2
Pulp, paper & allied products	1.9
Motor vehicles & equipment	1.7

Source: AAR. *Railroad Facts*, 2007 Edition.

Each Class I railroad that existed during the 1987-2006 period is represented by an indicator variable – e.g., KCS. When the observation is for the Kansas City Southern Railway, KCS equals one. Otherwise, KCS equals zero. Additional indicator variables are defined for mergers. For example, the UP system includes three railroads that appear in the database: UP, SP, and CNW. CNW was acquired by UP in 1995. Union Pacific merged with Southern Pacific in 1997. In the analysis, UP-CNW assumes a value of one in 1995, and each year thereafter – but is zero otherwise. Similarly, the variable UP-SP assumes a value of one in 1997, and each year thereafter – but is zero otherwise.

Analogous variables are defined for other mergers or acquisitions. For example, Burlington Northern merged with ATSF in 1996 to form BNSF. CSX and NS acquired parts of Conrail in 1999. In 2002, the Canadian National Railway consolidated ICG, GTW, and other rail lines into the Grand Trunk Corporation (GTC). In the Grand Trunk system, GTC is one if the year is 2002 or later; however, GTC is zero otherwise. The ICG indicator variable assumes a value of one when GTC is one, or when the observation is for the old ICG prior to 2002. The GTW variable works in a similar manner.

Collectively, the railroad indicator variables account for the unique characteristics of each railroad's system and changes in network characteristics resulting from mergers. However, all changes from 1987 to 2006 cannot be captured with aggregate data. Many innovations occurred in train location technology, centralized traffic control, computerization of locomotives, and systems modeling. These "other influences" are captured in the time variable (T), which is an integer that represents the elapsed time in years since 1987. T is expected to have a positive sign, indicating that railroad operational efficiency has been increasing over time.

STATISTICAL MODEL AND RESULTS

The model includes variables of different denominations, and data ranges. A logarithmic form of the regression model is useful for this type of analysis because it allows for easy comparison of effects. The dependent variable is the natural log of revenue ton-miles per train hour. The activity variables, the commodity variables, and time are entered as natural logs—therefore, their estimates are elasticities. The indicator variables (which take values of zero or one) do not affect the slope of the regression surface. Instead, they shift the intercept, and in doing so represent each railroad's unique system factors.

Main Effects

The parameter estimates and standard errors of the operational variables are shown in Table 2, along with the t-values and p-values associated with these estimates. A p-value represents the probability of observing a greater absolute value of *t* if the null hypothesis is true – i.e., the parameter estimate (the partial slope) is not significantly different from zero.

As expected, the signs of ALH, AVCARS, and NET are positive. Moreover, all three variables are highly significant with p-values of less than 0.01. ALH has the strongest positive influence. The effects are as follows: A 1.0% increase in length of haul increases operational efficiency by 0.40%. A 1.0% increase in average cars per train increases operational efficiency by 0.35%. A 1.0% increase in net load increases operational efficiency by 0.37%.

Table 2: Parameter Estimates and P-Values of Operational Variables

Parameter	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	11.31208	1.61563	7.00	<.0001
Log of Average Length of Haul (ALH)	0.40040	0.10949	3.66	0.0004
Log of Average Cars per Train (AVCARS)	0.35368	0.12048	2.94	0.0038
Log of Net Tons per Car (NET)	0.36688	0.11470	3.20	0.0017
Log of Way Train Trips per Car O/T (WTR)	-0.06662	0.02321	-2.87	0.0047
Log of Roadway Congestion Index (TD)	-0.32073	0.08894	-3.61	0.0004
Log of Yard Congestion Index (YD)	-0.15172	0.03319	-4.57	<.0001

As expected, TD and YD have negative signs and are highly significant. A 1.0% increase in train density results in a 0.32% decrease in revenue ton-miles/train hour. A 1.0% increase in yard density results in a 0.15% decrease in RTM/TH. In addition, a 1.0% increase in WTR results in a 0.07% decrease in operational efficiency.

Commodity Effects

As shown in Table 3, most of the commodity share variables (including coal) are not statistically significant. The unexpected negative sign of the coal variable may be the result of strong correlations with NET, AVCARS, and ALH. Coal unit trains have some of the highest net load factors and the most cars per train.

Three of the commodity share variables are statistically significant with positive signs: lumber products, miscellaneous mixed freight (which is shipped in containers), and pulp & paper products. Some lumber products are shipped on bulkhead or spine flatcars with achievable net loads of 100 tons. Pulp/paper products include woodchips (which can be loaded to 100 tons in large hopper cars), and newsprint and other paper products that use boxcars. Because boxcars can be reloaded by shippers, they incur relatively few empty miles. Therefore, fewer train hours are required to reposition these cars.

Several trends and efficiencies explain the positive sign of miscellaneous mixed freight: (1) increased use of double-stack flatcars; (2) increased use of 40-foot (versus 20-foot) marine containers; (3) increased use of larger domestic containers; (4) increased use of intermodal express trains; and (5) high utilization of flatcars. The first three trends increase the revenue tons per intermodal flatcar, while the fourth trend reduces the number of train hours required for land- and mini-bridge movements. In addition, intermodal flatcars can be reloaded with containers by the same shippers at origin and destination. Consequently, railroads incur fewer empty line haul or switching miles.

Table 3: Parameter Estimates and P-Values of Commodity Variables

Parameter	Parameter Estimate	Standard Error	t Value	Pr > |t|
Log of Coal Share	-0.04514	0.08877	-0.51	0.6118
Log of Chemicals Share	-0.07420	0.05837	-1.27	0.2057
Log of Farm Products Share	0.00623	0.04322	0.14	0.8856
Log of Nonmetallic Minerals Share	-0.09931	0.03933	-2.53	0.0126
Log of Food Products Share	-0.11716	0.08314	-1.41	0.1608
Log of Misc. Mixed Freight Share	0.07834	0.02563	3.06	0.0026
Log of Metal Products Share	0.06417	0.04146	1.55	0.1238
Log of Stone/Clay/Glass Products Share	-0.17032	0.05364	-3.18	0.0018
Log of Lumber Products Share	0.13213	0.03463	3.81	0.0002
Log of Petrol./Coke Products Share	-0.04918	0.03556	-1.38	0.1686
Log of Waste Materials Share	-0.06077	0.04550	-1.34	0.1836
Log of Motor Vehicles Share	-0.01584	0.01875	-0.84	0.3996
Log of Pulp & Paper Products Share	0.15435	0.06387	2.42	0.0168
Log of Metallic Ores Share	0.00051707	0.01316	0.04	0.9687
Log of Empty Containers Share	-0.01667	0.00617	-2.70	0.0077

Three of the commodity share variables are statistically significant with negative signs: empty containers; stone, clay, and glass products; and nonmetallic minerals. The first effect (empty containers) is easy to explain. Although the tare weights of containers are classified as revenue freight, their weights are small in comparison with the train hours required to reposition them. The effects of stone, clay, and glass products cannot be easily explained. This STCC group includes a variety of dimensional building materials which are relatively light-loading and move as individual carloads (rather than as unit trains or multi-car shipments). Their dispersed origins and destinations (and small shipment quantities) may result in less efficient train and yard operations—which could partially explain the significant negative sign. Nonmetallic minerals account for most of the railroads' non-revenue ton-miles, including the movement and distribution of crushed stone, riprap, and gravel used in track construction and maintenance. These movements result in train hours but no revenue ton-miles.

Fixed Network and Time-Related Effects

As shown in Table 4, T is highly significant with a positive sign, indicating that operational efficiency has been increasing over time as a result of technological or managerial innovations not reflected in the activity, commodity, or indicator variables. The effects of mergers are primarily reflected in average length of haul and the indicator variables. ALH accounts for the efficiencies of longer hauls and less-frequent interchange, whereas the indicator variables account for network effects such as combining lines, yards, and traffic control systems; eliminating redundancies and circuitous movements to avoid interchange; and broadening the pool of labor and specialized equipment available to the merged railroad. After controlling for ALH, the signs and parameter estimates of the railroad variables suggest that additional benefits have resulted from mergers.

The sign and estimate of each railroad indicator variable is relative to the variable left out of the equation, which is the unmerged UP railroad. As shown in Table 4, the signs of the two UP merger variables (UP-SP and UP-CNW) are positive and their estimates are greater than those of the individual railroads (SP and CNW). Similarly, the parameter estimate for the merged BNSF railroad is greater than the estimates for the BN or ATSF. Likewise, the estimates for the merged CSX and NS systems (CSX-CR and NS-CR) are greater than the estimates for NS, CSX, or Conrail.

Moreover, the estimate for the merged GTC system is greater than the estimates for the ICG and GTW railroads. Collectively, these indicator variables suggest that—after controlling for changes in lengths of haul—mergers have resulted in operational efficiency gains.

Table 4: Parameter Estimates and P-Values for Railroad Operational Model

Parameter	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	11.31208	1.61563	7.00	<.0001
Log of Miles of Road (MR)	-0.30520	0.08238	-3.70	0.0003
Log of Time (T)	0.11898	0.01453	8.19	<.0001
ATSF	-0.29006	0.08981	-3.23	0.0015
BNSF	0.33333	0.13874	2.40	0.0175
BN	-0.30684	0.08173	-3.75	0.0002
UP-SP	0.70667	0.15428	4.58	<.0001
UP-CNW	0.97562	0.20757	4.70	<.0001
SP	-0.70992	0.12679	-5.60	<.0001
CNW	-0.97671	0.19178	-5.09	<.0001
SOO	-1.23953	0.19985	-6.20	<.0001
ICG	-1.45662	0.22820	-6.38	<.0001
GTW	-2.12493	0.33796	-6.29	<.0001
GTC	2.43181	0.38182	6.37	<.0001
KCS	-1.40813	0.25154	-5.60	<.0001
CR	-0.87898	0.14582	-6.03	<.0001
CSX	-0.55094	0.11366	-4.85	<.0001
CSX-CR	0.98397	0.16235	6.06	<.0001
NS-CR	0.91573	0.16292	5.62	<.0001
NS	-0.66424	0.10828	-6.13	<.0001

Another network variable (miles of road) is also statistically significant with a negative sign. There may be several reasons for this result. (1) Miles of road decreased throughout the period, while operational efficiency increased. This suggests that the concentration of traffic on higher-density lines resulted in more efficient train operations. (2) Many of the miles that disappeared were sold to local or regional railroads that consolidate and distribute traffic on low-density lines. Many of these operations are time-intensive, requiring lengthy switching operations at branch-line stations. In effect, the transfer of lines to local operators relieved Class I railroads of some intensive switching operations. The parameter estimate of −0.30 suggests that (when controlling for changes in length of haul and other network efficiencies gained from mergers) a 1.0% reduction in miles of road has improved operational performance by 0.30%.

Model Fit and Properties

This model explains more than 97% of the variation in operational efficiency (Table 5). The R-Square adjusted for number of independent variables is 0.963. The F-test for goodness of fit is highly significant. Moreover, the coefficient of variation is less than 1%, indicating a very high level of precision.

Two statistical problems often occur in panel datasets: (1) nonconstant variance and (2) autocorrelation. Nonconstant variance occurs when variability increases with firm size or output. For example, the variance in revenue ton-miles among railroads may increase with output.

However, the model is essentially one of indexes (e.g., train-miles/running track mile), in which size and activity are scaled. Thus, the null hypothesis is constant variance. The p-value of the chi-square statistic used to test for constant variance is 0.70. A high p-value indicates there is little evidence to reject the assumption of constant variance. The Durbin-Watson statistic (which tests for autocorrelation) is 2.143. The probability of a greater value is 0.194. The probability of a lesser value is 0.806. These results suggest that serial correlation is not an issue. This conclusion is bolstered by an autoregression analysis in which none of the terms introduced to capture serial correlation were statistically significant.

Table 5: Model and Error Sum of Squares and Goodness of Fit Measures

Line	Source of Variation (a)	Degrees of Freedom (b)	Sum of Squares (c)
1	Model	40	20.0285
2	Error	152	0.5938
3	Corrected Total	192	20.6223
	Mean Square Errors	**Source**	**Value**
4	Model	Line 1(c) / Line 1 (b)	0.5007
5	Error	Line 2(c) / Line 2 (b)	0.0039
6	Total	Line 3(c) / Line 3 (b)	0.1074
	Goodness of Fit	**Source**	**Value**
7	R-Square	Line 1 (c) / Line 3 (c)	0.971
8	Adjusted R-Square	1 – [Line 5 (c) / Line 6 (c)]	0.963
9	Root Mean Square Error	SQRT(Line 2 (c))	0.0625
10	Mean Value of ln(RTM/TH)		10.419
11	Coefficient of Variation	Line 9 (c) / Line 10 (c) * 100	0.60%
12	F-Value	Line 4 (c) / Line 5 (c)	128.17
13	Prob. of > F-Value		<.0001

Qualitative Factors Not Considered

Although the statistical results are good, it must be recognized that many qualitative factors which influence railroad performance are not reflected in the model. To some extent, management and market influences may show up in the indicator and time variables. Technological factors that cannot be directly quantified may be reflected in the time trend variable—which is highly significant. While commodity/industry effects are represented in the model, institutional and labor factors (such as regulations and work rules) are not; nor are the dual freight and passenger train operations which occur on some lines. All of these factors may affect the revenue ton-miles per train hour that are attainable on certain parts of the railroad system. Nevertheless, the model explains 97% of the variation in operational performance among railroads over time.

CONCLUSION

This paper has analyzed the effects of key operational factors on railroad performance, while controlling for commodity and network effects. The results suggest that much of the improvement in performance between 1987 and 2006 was attributable to higher net loads, longer trains, and longer hauls. Other gains may have resulted from mergers and the transfer of branch lines to local and regional operators. The extent to which these trends will continue in the future is unclear. Increases in average lengths of haul are likely if land bridge and unit-train shipments grow proportionally. However, large increases in ALH may not be possible without mergers or consolidations, which are unlikely at present. On the other hand, performance may be improved through increases in train size and net load.

According to R-1 data, in 2006 the average train included less than 70 cars. Much longer trains are operationally feasible. To some extent, average cars per train can be increased through greater utilization of unit trains. As more commodities move in unit trains, a corresponding reduction occurs in way train activity. At the same time, net loads can be increased by substituting 286,000-pound and 315,000-pound cars for smaller ones. However, larger cars come with a price. System-wide use of heavier cars may require additional track investments and higher track maintenance costs. For example, Kalay and Guins (1998) estimated that track and bridge costs could increase by more than 20% on some routes if 315,000-pound cars are used instead of 263,000-pound cars. However, some of the infrastructure investments needed to support heavier railcars may have been made in recent years.

With growth in traffic, train densities have increased over time. The concentration of railroad traffic has undoubtedly had many benefits. However, the results of this study suggest that road and yard traffic may be at levels which are impacting system performance. This does not mean that all segments of the railroad system are congested. Since traffic is concentrated in certain corridors, it is likely that congestion is primarily a problem on some high-density lines. However, an assessment of the levels of congestion on particular lines cannot be made from this study.

In conclusion, some of the variables responsible for the impressive performance of Class I railroads during the last 20 years have been identified in this paper. However, if the volume/capacity ratios projected by Cambridge Systematics, Inc. (2007) occur, future railroad performance may be threatened by line and yard congestion. Railroads are adding capacity by constructing additional main tracks in high-density corridors. Whether these investments are adequate to meet the needs of future traffic is a question that is beyond the scope of this paper.

As noted in the introduction, railroad performance is important from a multi-modal perspective. Highway agencies are hoping to reduce the growth of highway congestion by keeping long-haul freight on rail lines instead of highways. If railroad performance deteriorates as a result of congestion, the performance of the surface transportation system (as a whole) will be affected.

Endnotes

1. A signal *aspect* is a visual arrangement of signal lamps and/or positions that conveys information about the status of upcoming blocks when viewed from the direction of an approaching train. The *indication* of a signal is the information conveyed by its aspect. For example, a red-colored light in a certain position (the aspect) means "stop" or "stop and proceed" (the indication). The simplest traffic control system has only two aspects. The signals display only red (stop) or green (proceed). However, a two-aspect system does not provide the capacity necessary for busy rail lines. In North America, three-aspect signaling is the most common type. However, four- and five-aspect systems are sometimes used. As the number of aspects increases, so does the information conveyed by the signal.

2. Block signal systems are widely used in the United States. In such a system, the network is partitioned into segments (or blocks) of uniform lengths based on train stopping distances. The train separation distance is prescribed in number of blocks (e.g., 1, 2, or 3), depending upon the number of signal aspects and other traffic control factors. The number of signal aspects and block lengths are key factors in line capacity and network performance.

3. The R-1 report is the annual report of a Class I railroad to the Surface Transportation Board. In the United States, each Class I railroad is required to report information to the Surface Transportation Board for regulation. The R-1 report includes information on the miles of road and

track owned or operated by a Class I railroad as well as operational statistics such as annual train-miles, car-miles, and ton-miles.

4. In some cases, it is possible to identify suspect R-1 data by using other variables as cross references. For example, the total tons and carloads handled in a year can be calculated from individual commodity subtotals in the QCS report and compared to the reported totals. Yard locomotive miles can be cross-checked using yard switching hours. Yard locomotive miles should be approximately equal to yard locomotive hours multiplied by typical switching speeds (e.g., 6 to 8 mph). Similarly, unit train car miles should be approximately equal to the carloads handled in unit trains times unit train miles. Several outliers that affected the analysis were discovered using these methods and by analyzing data series. The most reasonable carload and tonnage data were used. In some cases, the calculated totals were used instead of the reported totals. For one railroad, yard locomotive miles were zero or a very low number for several years. For this railroad, yard locomotive miles were computed by multiplying the yard locomotive hours by 6 mph. In total, these data substitutions were minimal.

5. This definition is included in the STB's instructions to Schedule 755, on Page 92 of the R-1 Annual Report form.

6. Singularity is a condition that occurs when one of a set of variables can be predicted from the values of the other variables in the equation.

References

American Association of State Highway and Transportation Officials (AASHTO). *Freight Rail Bottom Line Report*. Washington D.C., 2003.

Association of American Railroads (AAR). *Railroad Facts – 2006 Edition,* October, 2007.

Berndt, E.R., A.F. Friedlaender, J.S. Wang Chiang, and C.A. Vellturo. "Cost Effects of Mergers and Deregulation in the U.S. Rail Industry." *Journal of Productivity Analysis* 4, (1993): 127-44.

Bitzan, J. and T.E. Keeler. "Productivity Growth and Some of its Determinants in the Deregulated U.S. Railroad Industry." *Southern Economic Journal* 70, (2003): 232-253.

Cambridge Systematics, Inc. (CSI). *National Rail Freight Infrastructure Capacity and Investment Study*. Cambridge, Massachusetts, September 2007.

Kalay, S. and T. Guins. "Heavy Axle Loads: The Dollars and Sense Case." *Railway Age*, March, 1998.

Lee, T. and C.P. Baumel. "The Cost Structure of the U.S. Railroad Industry Under Deregulation." *Journal of the Transportation Research Forum* 28 (1), (1987): 245-53.

Martland, Carl. D. "Productivity and Prices in the U.S. Rail Industry: Experience from 1965 to 1995 and Prospects for the Future." *Journal of the Transportation Research Forum* 38 (1), (1999): 12-25.

Oum, T.H., W.G. Waters II, amd C. Yu. "A Survey of Productivity and Efficiency Measurement in Rail Transport." *Journal of Transport Economics and Policy* 33(1), (1999): 9-42.

U.S. Surface Transportation Board. *R-1 Class I Railroad Annual Reports to the Surface Transportation Board (R-1)*, 1987-2006.

Wilson, Wesley W., "Cost Savings and Productivity Gains in the Railroad Industry." *Journal of Regulatory Economics* 11 (1), (1997): 21-41.

Denver Tolliver *is director of the Transportation & Logistics graduate program at North Dakota State University and senior research fellow at the Upper Great Plains Transportation Institute. Tolliver earned a baccalaureate degree in geography from Morehead State and the Master of Urban & Regional Planning and Ph.D. from Virginia Tech. He is a long time member of TRF and past president of the Agricultural Transportation Chapter.*

John Bitzan *is an associate professor in the College of Business at North Dakota State University. His research interests include market structure and performance, the impacts of regulatory change, and the determinants of wages. He has a Ph.D. in economics from the University of Wisconsin-Milwaukee.*

Doug Benson *has over 20 years experience in transportation research specializing in the analysis of railroad operations and in the development of computerized systems used for transportation analysis. Current research interests include railroad and multi-modal analyses, the development of a multi-modal research program incorporating detailed waterborne commerce data and the utilization of the Indian Reservation Roads database. He has supported the development of several computer programming systems and has published papers on personal computer and mainframe software systems, railroad database systems, and freight systems.*

Aircraft Leasing and Its Effect on Air Carriers Debt Burdens: A Comparison Over the Past Several Decades

by Richard Gritta and Ellen Lippman

Leasing has always been an important source of finance to carriers in the U.S. airline industry. In the 1960-1970s, many carriers employed a type of lease called a financial lease as an alternative source of funds to acquire aircraft. It had a major advantage over purchasing the aircraft. It was "off-the-balance sheet financing," as the obligations under this type of lease appeared only in the footnotes to carrier balance sheets. Little use was made of short-term lease agreements during this period. The situation has changed radically over the past three decades. In 1976, the Financial Accounting Standards Board issued SFAS No. 13 defining specific criteria for capital leases that required the reporting of these "off-the-balance sheet financing" as both a leasehold asset and a long-term liability recorded under the long-term debt section of the balance sheet. In response, the air carriers substantially altered the way they finance airplanes. Carriers began to lease more and more of their aircraft, but they did so by structuring leases as shorter-term operating leases, which are not reported on companies' balance sheets. By strategically violating the criteria for capital leases, the air carriers once again pushed the leases off the balance sheet. The purpose of this research is to demonstrate the switch in the characteristics of aircraft leasing and to quantify the effects of such leases on air carrier debt burdens. In the process it will be argued that "debt is debt" no matter how it is structured. The paper updates two research studies by the authors to 2008.

INTRODUCTION

Increasingly, airlines are experiencing financial difficulties. Prior to 1978, few airlines declared bankruptcy due to the regulated nature of the airline industry; however, since then, over 155 air carriers have declared bankruptcy and reorganized or ceased operations, and this rate has increased in recent years. Just since the year 2000, more than 50 airlines have declared bankruptcy.[1] The airlines include both small air carriers and some of the larger airlines including United, Delta, and, recently, Japan Airlines. Some airlines have responded to their financial troubles by merging; for instance, Delta and Northwest Airlines merged as did America West and USAir. Air carriers also have responded by attempting to strengthen their operations and financial statements. One mechanism to strengthen a financial statement is through the increased usage of operating leases for aircraft purchases. Prior studies (Gritta 1974; Gritta, Lippman, and Chow 1994) have discussed the increased usage of leases over purchases of aircraft. This study updates that data, specifically identifying changes in leasing activity from the 1990s to 2008, to determine how leasing behavior has changed over time and the effect of such changes on the financial statement. Additionally, proposed changes in the reporting of lease accounting, as offered jointly by the International Accounting Standards Board (IASB) and the Financial Accounting Standards Board (FASB), are discussed and the potential effect of these proposals on the airline industry is considered.

PRIOR RESEARCH

Gritta (1974) reviewed aircraft lease accounting by US domestic airlines. The carriers he studied included American, Eastern, TWA, United, Braniff, Continental, Delta, National, Northeast, Northwest, and Western Air Lines. He found that these airlines leased 317 aircraft, representing

19.2% of their entire fleet. The actual incidence of leases ranged from zero planes leased for Northwest, Delta, and Continental, to 83% leased by Northeast. Accounting standards in effect then did not require capitalization of lease payments. Thus, airlines could structure purchases of aircraft as leases to avoid recording the asset and related liability associated with the purchase. To determine the effect of structuring purchases as leases, Gritta (1974) capitalized the lease payments using a present value of future cash flows methodology. Gritta (1974) capitalized only those long-term leases characterized as financial leases per Vancil's definition of a finance lease.[2] These included leases with rentals that exceeded the purchase price of the equivalent asset, or a contract lease term equal to the useful life of the plane. Operating lease payments were not capitalized.

Table 1 lists two debt ratios from the 1974 study that compared the ratios before and after lease capitalization. The ratios include total debt to equity and long-term debt to total capital (defined as long-term debt and equity).

Table 1: Ratios Before and After Lease Capitalization

	Total Debt/Equity				Long-Term Debt/Total Capital (%)		
	Before	After	% Change		Before	After	% Change
American	2.73	3.31	21.2		53.9	60.3	11.9
Eastern	3.67	4.45	21.3		73.2	77.2	5.5
TWA	3.07	3.53	15.0		63.0	66.8	6.0
United	2.28	2.71	18.9		52.7	58.0	10.1
Braniff	3.41	3.86	13.2		65.8	68.1	3.5
Continental	3.06	3.06	Unchanged		62.0	62.0	Unchanged
Delta	1.89	1.89	Unchanged		38.7	38.7	Unchanged
National	1.60	1.66	3.8		25.7	27.4	6.6
Northwest	0.80	0.80	Unchanged		17.2	17.2	Unchanged
Western	3.70	3.99	7.8		63.2	65.5	3.6

Source: Gritta (1974)

As is evident from Table 1, the ratios for many of the airlines were significantly, and negatively, affected by the capitalization of finance leases.

Subsequent to the Gritta (1974) study, the FASB issued Statement of Financial Accounting Standard (SFAS) No. 13 (1976) that established criteria to classify leases as either capital or operating. For lessees, SFAS No. 13 defines leases as capital leases if they meet one of these four criteria: lease term at least 75% of the useful life of the asset, present value of the lease payments at least 90% of the fair value of the asset at the date of lease inception, a bargain purchase, or a transfer of ownership at the end of the lease. Capital leases are, in substance, a purchase of an asset, and SFAS No. 13 mandates that the asset and related lease obligation be recorded on the balance sheet of the lessee at an amount equal to the present value of the minimum lease payments.[3] Payments for leases classified as operating leases continue to be expensed annually. Thus, as Gritta (1974) advocated, finance leases are now capitalized on the financial statements. However, operating leases are not capitalized, so opportunities for manipulation of the financials continue to exist if firms manage lease terms to characterize leases as operating.

Gritta capitalized lease payments only for those leases he characterized as finance leases. Had current accounting rules been used in the original study, Gritta (1974) would have classified more leases as capital leases than those capitalized in the 1974 study. However, the percentage of non-finance leases in the original study was not significant. Leased aircraft amounted to only 19.2% of the total fleet of 1,651 (317 leased/1,651 total aircraft). Of those 317 leased aircraft, only 41 (13%) were not classified as finance leases, just 2.5% of the total fleet (Gritta 1974, p. 48).

Gritta, Lippman, and Chow (1994) conducted a follow-up study on airline leasing by the major airlines to determine whether the usage of leases had changed post SFAS No. 13 and to calculate the impact, if any, on the financials from airlines' usage of operating leases. Since the 1974 study, several airlines included in the original study had ceased operations (Eastern and Braniff) and several others (National, Western, and Northwest) had merged with other airlines. The follow-up study included the remaining airlines from the 1974 study (American, TWA, United, Continental, and Delta) as well as new major airlines (Alaska Airlines, Southwest, USAir, and America West).

Gritta et. al. (1994) found two important results. One, air carriers had increased their usage of leasing to 56.5% of all aircraft, and the carriers now structured a significant number of their leases as operating leases, something new from the 1974 study. These results are detailed in Table 2.

Table 2: Leases of Planes

Carrier	Percentage of Planes Leased	Percent of Leases Operating
American	61.4	79.1
TWA	65.2	60.9
UAL	45.5	85.1
Southwest	N/A	N/A
Continental	68.6	84.3
Delta	44.2	90.1
Alaska	79.0	N/A
USAir	47.3	80.1
America West	81.2	100.0

Source: Gritta et al. (1994)

Additionally, when operating leases were capitalized, the effects on computed financial ratios were dramatic. As evident in Table 3, typical financing ratios worsened with capitalization. And the decline was more pronounced than in the 1974 study. For instance, in the 1974 study UAL's total debt to equity changed 18.9% with lease capitalization; in the Gritta et al. (1994) study, UAL's ratio changed 103.3%.

Table 3: Ratios Before and After Lease Capitalization

	Total Debt/Equity				Long-Term Debt/Total Capital (%)		
	Before	After	% Change		Before	After	% Change
American	3.27	5.66	73.1		67	82	22.4
TWA	NMF	NMF	NMF		NMF	NMF	NMF
UAL	5.19	10.55	103.3		61	87	42.6
Southwest	1.92	2.96	54.2		60	71	18.3
Continental	-2.56	-4.31	68.4		NMF	NMF	NMF
Delta	2.36	5.12	116.9		60	81	35.0
Alaska	3.26	5.03	54.3		70	81	15.7
USAir	3.90	8.39	115.1		71	87	22.5
America West	-7.67	-13.59	77.2		1.2	1.09	-9.2

Source: Gritta et al. (1994)
NMF – not meaningful figure

CURRENT DATA

The present study identified the major U.S. carriers and reviewed their lease usage in 2008 to determine changes since the prior studies in 1974 and 1995. The mix of airlines used in this study differs from past studies due to bankruptcies and mergers, suggesting continual upheaval in the airline industry. Table 4 documents the changes in air carriers in the current and previous studies.

For the 1974 study, data on leasing were available from Schedule B14 previously filed with the Civil Aeronautics Board (CAB). In 1980, this information was no longer collected by the CAB. Now, some leasing information is available in the 10-K where the airlines identify aircraft as owned or leased, and then classify the leases as either capital or operating. Table 5 provides a summary of leasing information and compares it with the information disclosed in the Gritta et al. study (1994).

Table 4: Airlines in Studies

	Gritta 1974	Gritta et al. 1994	Present Study
American	X	X	X
Eastern	X	c/o	
TWA	X	X	merged
United	X	X	X
Braniff	X	c/o	
Continental	X	X	X
Delta	X	X	X
National	X	merged	
Northeast	X	merged	
Northwest	X	merged	
Western	X	merged	
Alsaka		X	X
Southwest		X	X
USAir		X	X
America West		X	merged
AirTran		X	X
JetBlue		X	X

c/o - ceased operations

Table 5: Owned and Leased Aircraft

Carrier	Total Fleet	Operating Leases	Capital Leases	Planes Leased	Leased Current Study	Leased 1994 Study
AirTran	136	N/A	N/A	100	73.5%	n/a
Alaska	110	35	0	35	31.8%	79.0%
American	665	220	76	296	44.5%	61.6%
Continental	632	466	0	466	73.7%	68.6%
Jet Blue	142	55	4	59	41.5%	n/a
Delta	1,023	258	81	339	33.1%	44.2%
Southwest	537	82	9	91	16.9%	n/a
UAL	689	411	69	480	69.7%	45.5%
US Airways*	413	343	0	343	83.1%	52.5%

*Prior study data adjusted to reflect merger with America West
N/A: not available
Source: Data from carrier annual reports on www.sec.gov Website

Overall, the usage of leasing has declined since the 1994 study from 55.2% in 1994 to 50.8%. But this decline is somewhat misleading. First, for the three air carriers not included in the 1994 study, two of the three air carriers (Southwest and Jet Blue) had lower lease rates. Historically, Southwest conservatively finances its business, using less traditional debt and fewer leases, while Jet Blue has carried high traditional debt burdens since it was founded and relies more on traditional sources apart from leasing. Second, some airlines from the 1994 study (Continental, UAL, US Airways) increased the percentage of planes leased, while other airlines (Alaska, American, Delta) significantly decreased the percentage of leased aircraft. In the case of American and Delta, some leases were cancelled and aircraft returned to the lessors during the reorganizations of these airlines as the airlines sought to cut capacity.

While Table 5 focuses on all leasing activities, Table 6 focuses specifically on operating leasing. The incidence of operating leases is a better indicator of structuring leases to avoid capitalization than is the incidence of all lease activities.

Table 6: Operating Lease Usage

	Percentage of Leases Operating		Percentage of Fleet Operating	
Carrier	**Current Study**	**1994 Study**	**Current Study**	**1994 Study**
Alaska	100.0	N/A	31.8	N/A
American	74.3	78.8	33.1	48.5
AirTran	N/A	N/A	N/A	N/A
Continental	100.0	84.3	73.7	57.8
Delta	76.1	90.0	25.2	39.8
JetBlue	93.2	N/A	38.7	N/A
Southwest	90.1	N/A	15.3	N/A
UAL	85.6	85.1	59.7	38.7
USAir	100.0	84.8	83.1	44.5

Source: Gritta et al. (1994) and computed from data on www.sec.gov Website

The percentage of leases identified as operating is significant. Alaska, Continental, and USAir structure all of their leases as operating leases. This is an increase from the 1994 study. Two other air carriers, Jet Blue and Southwest, structure over 90% of their leases as operating leases. While neither firm was in the 1994 study, such a high percentage of operating leases is significant. However, consistent with what was described in Table 5, Delta and American show a reduction in operating leases as a percentage of total leases, although the decline is not significant for American. For air carriers in the 1994 study, all but Delta and American increased usage of operating leases as a percentage of the total fleet. Continental, UAL, and USAir have significantly increased their usage of operating leases for their fleet. USAir's percentage of fleet structured as operating leases is a high 100.0%.

To determine the effect of lease capitalization upon the financials, 2008 financial data on the air carriers were accumulated. Table 7 details the financial data on the airlines. The liabilities and equity amounts are those recorded on the company's balance sheet, while the capitalized lease amount was computed as a present value of operating lease payments as reported in the company's 10-K. To be consistent with the two prior studies, a simple present value calculation was performed, using the discount rate of 10% to be consistent with the rate used in both the 1974 and 1994 studies. (In actuality, a lower discount rate may be appropriate currently. This would increase the capitalized lease amount, thereby showing a greater decline in financial health.)

Table 7: Total Capital, Including Aircraft Leases (in thousands)

Carrier	Current Liabilities	Long Term Liabilities	Capitalized Leases	Equity
AirTran	712	1,105	1,797	246
Alaska	1,392	2,441	577	595
American	11,071	16,352	5,318	(4,905)
Continental	4,474	8,107	8,756	105
Delta	11,022	33,118	7,883	874
JetBlue	1,081	3,681	1,167	1,261
Southwest	2,806	6,549	1,375	4,953
UAL	7,281	14,645	3,708	(2,465)
USAir	3,966	3,209	4,849	(221)

Source: Carrier 10Ks

The figures for current and long term liabilities were taken directly from carrier 10K balance sheets. The capital leases equivalents were calculated in the same manner as the 1994 study.[4]

It is apparent that the airlines are in trouble. American, UAL, and US Airways have negative equity positions as of December 31, 2008. Delta's positive equity is due to the adoption of fresh-start accounting when prior accumulated losses were removed after emerging from bankruptcy protection in 2007. Southwest, a low cost provider, is the financially healthiest air carrier.

Using the financial information from Table 7, financial ratios were calculated to determine the impact of lease capitalization on debt ratios in Table 8. Total Capital is defined as the sum of long term debt, capitalized leases (where appropriate), and equity (the company's stockholders equity). Stockholder's equity was not adjusted for changes in classification from operating to capital leases. While operating lease expense will disappear with lease capitalization, depreciation and interest expense would increase. Over the life of the lease, these amounts will approximate one another.

Table 8: Change in Ratios with Lease Capitalization

	Long-Term Debt/Total Capital (%)			Total Debt/Net Worth		
	Before	After	% Change	Before	After	% Change
AirTran	81.8	92.2	12.7	7.39	14.69	98.8
Alaska	80.4	83.5	3.9	6.44	7.41	15.1
American	142.8	129.3	-8.06	-5.59nm	-6.67nm	19.3
Continental	98.7	99.4	0.7	119.82	203.21	69.6
Delta	97.4	97.9	0.5	50.50	59.52	17.9
JetBlue	74.5	79.4	6.6	3.78	4.7	24.3
Southwest	56.9	61.5	8.1	1.89	2.17	14.8
UAL	120.2	115.5	-3.9	-8.89nm	-10.40nm	17.0
US Airways	107.4	102.8	4.3	-32.47nm	-54.41nm	67.6

Source: Computed from Table 7
nm = no meaningful figure

For nearly all of the ratios, the airlines appear more risky when operating leases are capitalized. Yet, do these changes necessarily affect the evaluation of the riskiness of the airlines? Table 9 displays the ranking of the airlines' debt ratios before and after lease capitalization. The riskiest ratio is labeled one, second riskiest two, and so on.

Table 9: Ranking of Risk Associated with Air Carriers Before and After Lease Capitalization

	LTD/Total			Debt/Equity	
Carrier	Before	After		Before	After
Air Tran	6	6		3	3
Alaska	7	7		4	4
American	1	1		nm	nm
Continental	4	4		1	1
Delta	5	5		2	2
Jet Blue	8	8		5	5
Southwest	9	9		6	6
UAL	2	2		nm	nm
US Airways	3	3		nm	nm

nm = not meaningful

While the ratios changed, the riskiness of the firms relative to one another remained unchanged with or without the capitalization of the operating leases. Thus, the capitalization of operating leases may not provide significantly useful information, at least when comparing one airline with another. While firms seem to have gone to great lengths to structure more aircraft purchases as leases, and more of these as operating, when comparing one air carrier with another, the capitalization of leases does not affect their relative riskiness as seen through the rankings.

PROPOSAL FOR CHANGES TO LEASES AND THE EFFECTS ON THE AIRLINE INDUSTRY

In the United States, SFAS No. 13, and now ASC 840, has been vilified as one of the worst accounting pronouncements in part due to the "bright lines," the set percentages that must be met in order to warrant capitalized lease recognition. Likewise, the international leasing standard IAS 17 does not have strong support. In response, in March 2009 the International Accounting Standards Board and the FASB issued a discussion paper on leases, "Leases: Preliminary Views" (FASB 2009). In this paper, the boards jointly recommended new lease accounting standards and requested comments about the proposal. The boards continue to work together to craft a new lease accounting standard acceptable to both groups.

The boards believe that a lease represents a right by the lessee to use an asset. "A lease is a contract in which a right to use a specific asset is conveyed for a period of time, in exchange for consideration" (FASB 2010). The boards also believe that a lease includes an obligation to pay for the right to use the asset. The discussion memorandum proposes that all leases, capital or otherwise, should be recorded as assets and liabilities on the lessee's balance sheet. While there exist some differences between the suggested lease accounting by the IASB and the FASB, the IASB and FASB agree on the necessity to record all leases on the lessee's balance sheet. In effect, it has been argued that "debt is debt" no matter how it is structured.

Previously, some leases were in fact purchases of assets and thus were classified as finance leases, where the lease and the related obligation were recorded as an asset and corresponding liability. Other leases were executory contracts, and these were classified as operating leases, with the lease payments recorded as lease expenses. If the proposal is accepted, capitalization of all lease payments for property, plant, and equipment, regardless of prior classification, will occur such that the present value of all lease payments will increase assets and liabilities. Thus, the lease capitalization suggested in this paper will no longer be necessary. Yet, while lease capitalization does impact the reported assets and liabilities, and computed ratios thereof, interestingly, the rankings of these ratios among the airlines were unaffected, suggesting that the public, with or without the aid of lease capitalization, has other financial markers to determine the riskiness of air carriers.

Endnotes

1. Prior to deregulation, the CAB did allow the merger of a weak carrier into a stronger one. The merger of Northeast Airlines, then classified as a major, into Delta in the early 1970s is an example. This was a way around a bankruptcy filing and thus explains the low rate prior to the 1980s. There is also no question that the financial upheaval caused by the events of 9/11 helps to explain the sudden spike in filings under the bankruptcy code since 2001.

2. Vancil defined a financial lease as a contract with non-cancelable payments that, in total, are greater than the cost of the item leased (Vancil and Anthony 1963).

3. In July 2009, all SFAS were combined into the Accounting Standards Codification that became the single source of authoritative accounting principles. Lease accounting is now referenced in ASC 840, but the leasing requirements remain the same as SFAS 13.

4. As noted, operating leases are not directly reflected on the balance sheets of the airlines, but the yearly future obligations of these leases are disclosed in the footnotes to the financial statements. The operating lease disclosures reported in the footnotes include the yearly obligation for each of the next five years (LP_t), and then a total of the remaining obligations (TLP) under those leases beyond the five years. Just as in the 1994 (Gritta et al.) study, the fifth year lease payment (per the carriers' footnotes) was estimated to continue, such that the remaining life of the leases per air carrier differed dependent upon the remaining balance of lease payments and year five's lease payment. It should be noted that prior to the sunset of the CAB, lease forms called B-14s required the exact lease payments for each year of the lease, and there was no aggregating the payments in the future. This makes the current analysis a little less precise than was the case in the 1974 study. The impact of the obligations on the balance sheet was estimated by determining the present value of these non-cancellable operating lease payments in a manner consistent with the 1994 study. The discount rate used (10%) was the same used in both the 1974 and 1994 studies for comparative purposes. The following formula shows the approach used to estimate L_o, the net present value of the operating leases for each carrier:

$$\sum_{t=1}^{n=5} \frac{LP_t}{(1+k)^t} + \sum_{t=n+1}^{a} \frac{RLP}{(1+k)^t} = L_o$$

LP_t are the payments for each of the first five years presented in the carriers' reports, k is the discount rate of 10%, and RLP represents an estimate of the yearly remaining lease payment assumed to be the same as in year five, except for the last year when it is adjusted so that total lease payments since year five equal the total remaining lease payments as reported in the footnote. The remaining years of the lease, a, is computed by dividing the total lease payment after year five (TLP) by year five's lease payments. The present values of the lease obligations were then included with the recorded long-term debt to compute the adjusted long term debt obligations for each airline.

References

Financial Accounting Standards Board. "Accounting for Leases." *Statement of the Financial Accounting Standards Board (SFAs) #13*, 1976.

Financial Accounting Standards Board. Discussion Paper. "Leases: Preliminary Views*,"* 2009. Http://www.fasb.org/draft/DP_Leases.pdf.

Financial Accounting Standards Board. *Summary of Tentative Decisions Reached to Date* (as of February 18, 2010). Http://www.fasb.org/cs/ContentServer?c=FASBContent_C&pagename=FASB%2FFASBContent_C%2FProjectUpdatePage&cid=900000011123.

Gritta, R. "The Impact of the Capitalization of Leases on Financial Analysis." *Financial Analysts' Journal* 30 (2), (1974): 47 - 52.

Gritta, R., G. Chow, and E. Lippman. "The Impact of the Capitalization of Leases on Financial Analysis: An Issue Revisited." *Logistics and Transportation Review* 30 (2), (1994): 189-202.

Vancil, R. and R. Anthony. *Leasing of Industrial Equipment*. McGraw-Hill Book Co., (1963).

Dr. Richard D. Gritta *is professor of finance and transportation at the Robert B. Pamplin Jr., School of Business Administration, University of Portland, Portland, Oregon. He earned his BBA from the University of Notre Dame, his MBA from Indiana University, and his Ph.D. from the University of Maryland. He has published numerous articles in such journals as the Journal of the Transportation Research Forum, Transportation Journal, Logistics and Transportation Review, Credit and Financial Management Review, Financial Management, Handbook of Transport Strategy, Policy and Institution, Journal of Transportation Law, Logistics and Policy, and many others. He was a Senior Fulbright Fellow to the Transportu Katendra at the Warsaw School of Economics, Poland, from February to June of 1996 and a Senior Fulbright Specialist to the Groupe ESC (Toulouse School of Business, France) from August to November of 2007. He is currently associate editor-air transportation for the Journal of the Transportation Research Forum and is the immediate past president of the Transportation Research Forum.*

Dr. Ellen J. Lippman *is an associate professor of accounting in the Pamplin School of Business Administration at the University of Portland. She earned her B.A. from Colorado College, her M.S. in accounting from the University of Wisconsin (Madison), and her Ph.D. in accounting from the University of Oregon. She is also a Certified Public Accountant, specializing in financial accounting and auditing. Dr. Lippman has published in the Journal of Business Ethics, Accounting History, AICPA Professor/Practitioner Cases, Strategic Finance, Journal of Accounting and Business Research, Transportation Law Journal, and Logistics and Transportation Review. Her interests include lease accounting in the airline industry, ethical credit lending, and accounting.*

Rail Competition Changes Since the Staggers Act

by Marvin E. Prater, Ken Casavant, Eric Jessup, Bruce Blanton, Pierre Bahizi, Daniel Nibarger, and Isaac Weingram

Agricultural and other shippers are concerned about the sufficiency in rural areas of transportation capacity, the sufficiency of competition in the transportation system, the reliability of transportation services, and the reasonableness of rates. This paper examines the sufficiency of rail freight competition and the effects of intramodal competition on rail rates.

The paper begins with a review of the importance of rail transportation for U.S. agricultural producers. Specific attention is paid to the nature of competition faced by railroads, especially since deregulation, using the analytical tool of inverse Herfindahl-Hirschman Indices (HHI), by USDA Crop Reporting Districts (CRD). As shown by the inverse HHIs, the overall level of rail competition for grains and oilseeds has generally decreased since the 1985-1992 period, even though rail competition has increased for some CRDs. In addition, revenue to variable cost ratios (R/VC) increased in most of the CRDs analyzed, and the analysis found them related to the number of railroads competing in the CRD.

Competition is then analyzed relative to the revenue per ton, revenue per ton-mile, and the revenue to variable cost ratios (R/VC) associated with the level of competition for six states with the least rail-to-rail competition, and distant from water transportation, with those for four states having more rail-to-rail competition and close proximity to water transportation.

INTRODUCTION

Railroad deregulation encouraged greater reliance on free markets to promote railroad profitability and public benefits (Grimm and Winston 2000). The Staggers Act significantly reduced economic regulation in the railroad industry, which has benefited shippers as well as railroads, but relies on competition to protect shippers and the public.

Since the Staggers Act, the average rate of return on investment for the railroad industry has increased from less than 2.5% during the 1970s to slightly more than 10% during 2006 and 2007 (AAR 2008). The return on equity for the railroad industry, when compared with revenue adequacy standards using STB's Capital Asset Pricing Model (CAPM),[1] has exceeded revenue adequacy[2] levels from 2002 through 2006. In addition, railroad industry earnings above CAPM revenue adequacy standards have widened from 2004 through 2006 (Christensen & Associates 2008). However, when using the STB's historical revenue adequacy standard, the Discounted Cash Flow model (DCF), the railroad industry did not appear to be revenue adequate (Christensen & Associates 2008).

During the first decade after the Staggers Act, the annual benefits to shippers from railroad deregulation amounted to more than $12 billion in 1999 dollars, equivalent to $14.7 billion in 2007 dollars (Grimm and Winston 2000). Shippers have benefitted from 20 years of decreasing rail rates (in terms of inflation-adjusted revenue per ton-mile) and the preservation of rural lines sold or leased to smaller railroad firms. Many of these new short line railroads have been able to operate profitably on rail lines abandoned by the major railroads, and have generally provided more individualized service to shippers (Babcock, Prater, Morrill, and Russell 1995).

As expected, the distribution of benefits has tended to favor grain producers and shippers in regions with more transportation competition (Bitzan, Vachal, VanWechel, and Vinge 2003). In addition, rates have not declined uniformly across commodities, and rates for some commodities are significantly higher than those for other commodities. In particular, from 1987 to 2004, rail rates for grain have increased 9%, but rates have declined for coal, motor vehicles, and miscellaneous mixed shipments (GAO 2006).

Despite the initial success of the Staggers Act, agricultural producers and shippers continue to express concern about decreased rail-to-rail competition, rapidly increasing rail rates, poor rail service, rail capacity constraints, and the fair allocation of rail capacity (NGFA 2005, USDA 2005). Because of these shipper concerns, Congress required USDA and USDOT in Section 6206 of the Food, Conservation, and Energy Act of 2008 (PL 110-246) to submit a report on rural transportation issues. The basis for this paper comes from some of the analysis from that report (USDA/DOT 2010).

Current Congressional debate includes proposals to reform railroad regulation because of decreases in rail competition since 1980, which have increased the amount of railroad market power relative to that of shippers (USDA 2000). Thus, a major purpose of this paper is to measure the level of rail competition for the transportation of grains and oilseeds in each crop reporting district (CRD). Two common ways of measuring changes in competition are the four-firm concentration ratio and the use of a Herfindahl-Hirschman Index (HHI). This paper examines changes in the four-firm concentration ratio in the United States since 1980 and changes in the inverse HHI by CRD since 1985 and finds that the amount of rail-to-rail competition for the shipment of grain and oilseeds has decreased for both measures.

In addition, this paper compares nominal rail rates per ton and per ton-mile as well as revenue-to-variable cost ratios of six states with limited rail-to-rail competition, and distant from barge-loading facilities, with those of four states with more rail-to-rail competition and close proximity to barge-loading facilities.

An affordable and reliable transportation network is necessary to maintain the strength and competitiveness of American agriculture and rural communities. Rail service is a particularly important part of that network for U.S. agriculture, because it is virtually the only cost-effective shipping alternative available for low-value, bulky commodities in rural areas that are distant from water transportation and markets. Therefore, examining the performance of the railroad system in agricultural movements is timely and policy relevant.

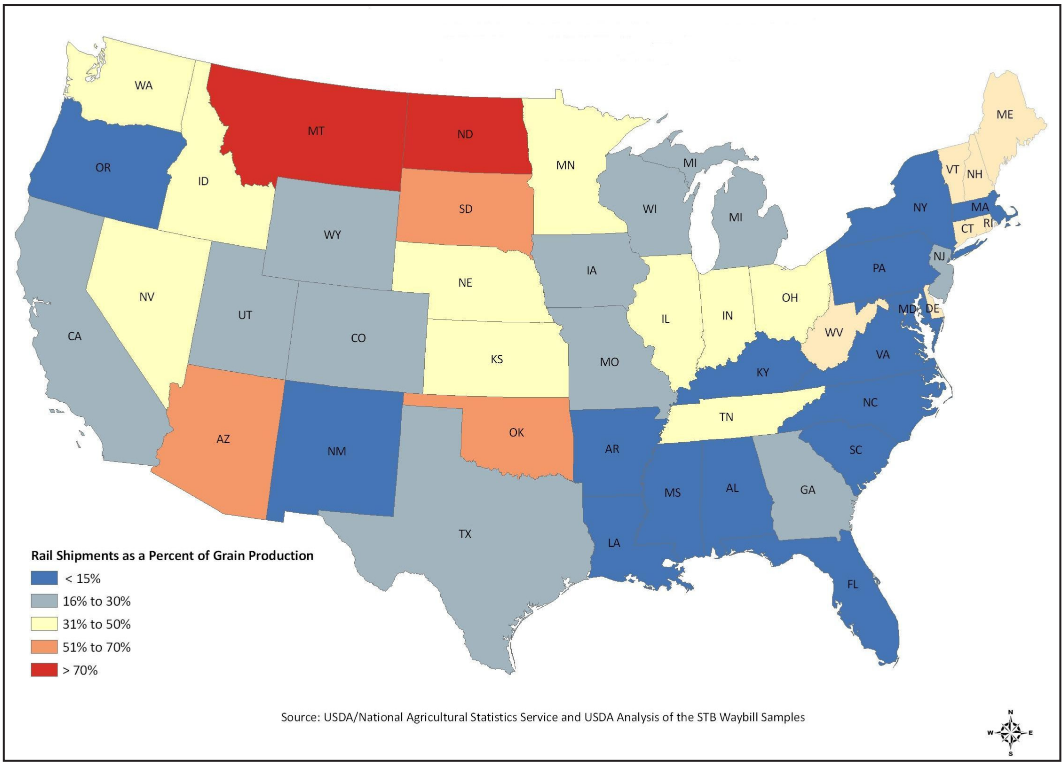
Agricultural producers are dispersed over the entire country, and unlike most other industries, are unable to move their operations at will as their production is tied to the land, and often to a particular climate. Because agricultural production is tied to arable land, producers must be able to transport their products to markets, many of which are located long distances from the farms.

Nine of the 10 top wheat-producing states are more than 150 miles from barge transportation on the Mississippi River, which usually provides the strongest intermodal competition to railroads for the long-distance movement of grain to export ports. Unlike other agricultural shippers in the United States, wheat shippers in much of the Great Plains have no cost-effective transportation alternatives to railroads. The wheat produced in these areas moves long distances to domestic markets for processing and consumption or to coastal ports for export. Shippers in these regions have little direct access to inland waterway transportation, and the distances involved make truck transportation uneconomical. Consequently, wheat is particularly dependent on rail; 66% of all wheat and wheat exports moved by rail during 2007 (USDA 2010).

Agricultural shippers in Montana and North Dakota are particularly dependent on rail transportation because of their distance to inland waterways and the prohibitive distance for the use of trucks. On average, railroads transported more than 70% of the grains and oilseeds originated in Montana and North Dakota during the crop marketing years from 2004 to 2007 (Figure 1). A study by Tolliver and Dybing (2007) revealed that during crop marketing year 2004, railroads transported 78% of North Dakota crops. Another recent study states that Montana ships nearly 100% its wheat by rail, which corresponds with USDA findings for Montana (Cutler, Goldstein, Fauth, Crowley, and Whiteside 2009).

During the crop marketing years 2004–2007, railroads transported more than 50% of the grain production of Arizona, Oklahoma, and South Dakota. During the same time period, rail moved more than 30% of grain and oilseed production in the states of Idaho, Illinois, Indiana, Kansas, Minnesota, Nebraska, Nevada, Ohio, Tennessee, and Washington.

Figure 1: Railroad Shipment/Grain Production Ratio, Average 2004-2007 Marketing Years



LITERATURE REVIEW

Numerous studies discuss railroad industry competition and pricing, providing varying degrees of analysis as to the impact of competition within the agricultural industry. Many of these papers, which are regional in scope, investigate the impact of deregulation after the Staggers Rail Act of 1980, with the majority written in the decade after enactment. The following literature review is not a complete survey of the prior research related to railroad competition, but instead emphasizes the interaction between railroad competition and rail grain transportation prices.

Babcock, Sorenson, Chow, and Klindworth (1985) investigated the impact of the Staggers Act on Kansas agriculture. The study found substantial railroad rate reductions in the four-year period of 1981 through 1984. The pattern of rate changes suggested the presence of both intramodal and intermodal transportation competition. Tariff rates to the Gulf of Mexico during this period dropped 34%, compared with a 64% increase in the four years preceding the Staggers Act.

MacDonald (1987) used regression analysis of the 1983 waybill sample data to examine the rail rates for corn, wheat, and soybeans. Tonnage, distance, and volume of the shipments were inversely related to rates. In addition, increased intramodal competition, as depicted by the reciprocal of the Herfindahl Index, was inversely related to rates, and rail rates increased with distance to waterways.

A later study (MacDonald 1989a) used waybill data from 1981 through 1985 to analyze rail rates and competition for corn, wheat, and soybeans. MacDonald found that when rail service goes from a monopoly to a duopoly, rail rates decline 18%. The addition of a third competing railroad resulted in an additional 11% decrease in rail rates. In addition, shippers located 400 miles from barge access paid rail rates that were 40% higher than those located 100 miles from barge access.

Finally, MacDonald calculated inverse Herfindahl-Hirschman indices for each CRD and concluded that each CRD was characterized by rail oligopolies.

Harbor (2008) found similar results for rail rates for soybeans. Rail rates decreased 10.9% when moving from a monopoly to two-railroad competition in a market 300 miles from a barge-loading facility. Adding a third railroad decreased rates another 6.5%. Furthermore, in the 12-state region studied, the average inverse HHI for corn had dropped to 1.86 in 2004 from 2.30 in 1983. The average inverse HHI for soybeans and wheat decreased from 2.46 in 1983 to 1.90 in 2004 and from 1.85 in 1983 to 1.58 in 2004, respectively.

Chow (1986) studied post-Staggers rail grain rates for the Central Plains region. The analysis indicated an overall reduction in wheat rail rates of 34.5% in the five-year period after enactment of the Staggers Act, with the most significant reductions occurring in movements to the export markets.

Kwon, Babcock, and Sorenson (1994) examined the impacts of the Staggers Act in the latter half of the 1980s and found that railroads practiced differential pricing for intra-Kansas and export shipments of wheat. There were substantial differences in the factors affecting the revenue-to-variable cost ratios for intra-Kansas wheat movements versus that of Kansas export wheat movements. They found revenue-to-variable cost ratios increased steadily from 1986 through 1989, but may have been caused by diminishing export demand.

Fuller, Bessler, MacDonald, and Wohlgenant (1987) found deregulation to have had a significant effect on rail corridors linking Kansas and Texas with Gulf ports and a relatively modest effect on the corridor linking Indiana with East Coast ports. Real rail rates declined \$.37 per bushel in the Kansas corridor and \$.31 per bushel in the Texas corridor during the 1981-1985 period. In the Indiana corridor, estimated real rail rates decreased \$.08 per bushel. Railroad deregulation had little statistically significant effect on real rail rates from Iowa and Illinois to the Gulf ports.

Koo, Tolliver, and Bitzan (1993) examined railroad pricing behavior in North Dakota. The authors stated that the region has unique transportation characteristics that include limited intermodal competition due to great distances to barge-loading facilities and to major domestic and export markets. They found that distance, volume, weight per car, intramodal competition, and intermodal competition had significant negative effects on rail rates.

Thompson, Hauser, and Coughlin (1990) evaluated the pre- and post-Staggers effect of competition on railroad revenue-to-variable cost ratios for export shipments of corn and wheat. The regression results for corn rates were less significant than those for wheat rates. There was a lack of identifiable differences in pre- and post-Staggers pricing, which they felt may be attributable to the close correlation between changes in operating factors, such as shipment size and destination opportunity. They concluded that their results did not indicate a clear effect of the Staggers Rail Act on rail rate competitiveness.

Wilson and Wilson (2001) examined rail rates per ton-mile for barley, corn, sorghum, wheat, and soybeans moved by rail. The explanatory variables include commodity ton-miles, commodity prices, average length of haul, and a non-linear specification of deregulation that allows the effects to phase in over time. The results showed that commodity prices have positive effects on rail rates per ton-mile and length of haul has a strong negative effect. The results indicated a large negative effect on rates per ton-mile from deregulation, which dissipate with time.

RAIL COMPETITION IN AN ERA OF DEREGULATION

This section discusses the deregulation of the railroad industry, the role of competition, and various types of competition in the railroad industry today to establish a foundation for the analysis of changes in railroad competition for grain and oilseed transportation.

Deregulation of the Railroad Industry

The constraints of pervasive economic regulation, although meant to protect shippers from the abuse of railroad market power, resulted in nearly bankrupting the railroad industry as well as increasing shipper costs. Furthermore, federal legislators recognized that industry regulation was expensive for both industry and government, and created market distortions for nearly all regulated markets (Hovenkamp 2005). Congress deregulated railroads in response to arguments that the industry needed greater pricing and operating freedom to avoid more bankruptcies (Meyer 1973).

As the nation deregulated the railroad industry, conflicting goals included the preservation of effective transportation competition, the regulatory protection of captive shippers, deregulation of rail rates when sufficient competition is present, and revenue adequacy of railroad firms.

The concept of adequate competition is so important that competition is mentioned four times, avoidance of undue concentration of market power is mentioned once, and adequate railroad revenues or sound economic conditions is mentioned twice in the 15 Rail Transportation Policy goals of the Staggers Act and ICCTA (Public Law 104-88). The presence of transportation competition was expected to protect most shippers by constraining the use of railroad market power. On the other hand, adequate revenues are necessary for rail service to remain viable and continue providing service.

In cases when rail-to-rail competition is not present, captive shippers need meaningful and cost-effective rate appeals processes to protect against the excessive use of railroad market power (GAO 1999, USDA 2004). Until 2008, the only rail rate appeals used by shippers were Stand-Alone Cost procedures, which cost millions to adjudicate and involve the construction and operation of a hypothetical railroad (GAO 1999, USDA 2004, USDA and USDOT 2010). Whereas a coal case typically involves the construction of a hypothetical railroad between a single origin and a single destination, a rate case for agricultural shippers would typically involve single or multiple origins to multiple destinations (USDA 2004). Agricultural shippers in the Northern Plains appealed rail rates in the McCarty Farms case, but spent millions of dollars and lost on appeal after 16 years (USDA and USDOT 2010). Thus, coal shippers are the main users of the Stand-Alone Cost procedures, which are considered far too expensive and complicated for agricultural shippers to use. Indeed, an agricultural shipper would spend far more filing a Stand-Alone Cost rate appeal than the value of the case (USDA 2004, USDA and USDOT 2010).

Even though Congress had required small rate case appeals procedures as early as 1986, small shippers essentially had no protection until 1996, when the STB first made small rate case appeals procedures available (USDA and USDOT 2010). Small shippers, however, did not use those procedures because they did not perceive them to be cost effective and were concerned about the uncertainty of the process (USDA 2004). The STB held a proceeding regarding small rate case appeals procedures and set new rules for small rate case appeals in 2008. Although chemical shippers (i.e., E.I. Dupont v. CSX Transportation) have successfully used these procedures, so far no agricultural shipper has appealed rail rates under these new appeals procedures (USDA and USDOT 2010).

Role of Competition

Some economists claim that the way to preserve and extend the benefits of deregulation is to increase rail competition (Grimm 2004). Many shipper groups have echoed this conclusion in comments prepared for various proceedings before the Surface Transportation Board. Market-based competition is a fundamental economic policy of the United States (National Commission 1979). Competition requires businesses to become efficient and effective in providing the kinds and quality of goods and services the consumer desires.[3] Competitive markets reduce market distortions and result in the efficient allocation of resources, providing a basis for economic development. "The U.S. economy is an example of how free markets can lead to the creation of wealth, making possible improved living standards and greater prosperity (Lewis 2004)." Furthermore, industries sheltered from competition are less vigorous and successful than industries subject to competition (Porter 1990).

When economic regulation is imposed on an industry, competition is not as important because government protects the consumer and social welfare (Kahn 1998). When an industry is deregulated, however, competition and antitrust enforcement become the major forces protecting the consumer and society from unfair business practices (Kahn 1998). The loss of competition, combined with deregulation, could lead to the unrestrained use of market power (Kahn 1998). This is especially true in highly concentrated industries that possess market power, such as the railroad industry (Kahn 1998). Unrestrained use of railroad market power would likely result in unnecessarily high rail rates and the inability of agricultural producers to reach multiple and competing markets. Because agricultural producers typically receive a price net of transportation, higher rail rates and inability to access a variety of markets result in reduced producer income. The preservation and protection of competition is vital for the economic prosperity of agricultural producers and shippers contending with a mostly deregulated railroad industry.

Effective Competition

In order for competition to be effective, it must be cost competitive. Four types of competition can constrain the use of railroad market power (Tye 1991, STB 1998):

- Intermodal competition from other transportation modes, such as motor carrier, multimodal, and barge transportation
- Intramodal (rail-to-rail) competition among individual, independent railroads
- Geographic competition, in which a producer can haul products to rail loading facilities located on competing railroads or in which a buyer could obtain products from other originating locations
- Product competition, in which a producer can substitute other inputs in the manufacture of a product

Barges, railroads, and trucks not only compete against each other, they also complement each other. Before agricultural products reach the market, they have often been transported by two or more transportation modes. This balance between competition and complementaries provides agricultural shippers with a highly efficient, low-cost system of transportation (USDA 2006).

Decreased rail-to-rail competition among Class I railroads has resulted in an increased ability of railroads to raise rates. The presence of a competing railroad has a noticeable effect on rail rates. Rail rates rise well above incremental costs in regions with one or two railroads and are far removed from navigable rivers (MacDonald 1989b).

Although product and geographic competition can limit railroad pricing in some cases, these forms of competition are less relevant to market dominance today in light of the rapid consolidation of the rail industry (STB 1998).

The average number of route miles operated by each of the Class I railroads in the United States has more than tripled since 1980, resulting in dominance over larger geographic regions by a single Class I railroad. Railroad mergers of the 1960s and 1970s combined smaller rail systems that operated in smaller geographic territories. In the 1980s, newly merged rail systems began to gain dominance within some geographic regions. For instance, in 1960, the average Class I railroad in the United States operated 1,956 route miles. By 1980, this had increased to 4,226 miles, and by 2007, to 13,473 miles (AAR Ten-Year Trends various editions).

As a result, many farmers in the Plains States no longer have a cost-effective option of hauling grain to an elevator served by a competing railroad. In 1980, the ability of a farmer to haul grain to an elevator served by a competing railroad often provided the competition necessary to constrain rail rates. Today, only two Class I railroads are dominant in the western United States and two are dominant in the eastern United States. This decrease in rail-to-rail competition has decreased the effectiveness and the relevance of geographic and product competition (USDA 1998).

RAILROAD CONCENTRATION AND MARKET SHARES

Since the 1920s, many railroads have merged. During the 1960s and 1970s, many of the mergers combined financially weak railroads with stronger firms in the hope of developing a financially stable railroad that was large enough to compete effectively with other transportation modes. After deregulation, the pace of merger activity picked up as railroads strove to increase geographic range, eliminate duplicate lines, reduce costs by increasing the size of the firm, and gain increased market power.

Many at the time of enactment of the Staggers Act did not foresee the extent of the loss of rail-to-rail competition because of rail mergers, which has resulted in increased market power. Many economists, however, anticipated the rationalization of the rail network because the regulated railroad industry was characterized by over-capacity. Consequently, reduction in excess capacity was a logical and expected result of deregulation. The concentration of increased tonnage on fewer track miles has enabled railroads to reap enormous economies of scale. Rail costs have fallen 60% in real terms, and most of these savings have been passed on to shippers (STB 2009).

Today there are two major duopolies—one serving the western United States and the other serving the East. In addition to these four mega railroads, during 2007 there were three smaller Class I railroads serving the central portion of the nation, 33 regional railroads, and 523 local railroads (AAR 2008).

Market Concentration and Share

The top four Class I railroads originated 84% of grain and oilseed traffic in 2007, compared with only 53% in 1980 (see Figures 2 and 3). In addition, the market share of the predecessor railroads compared with the current railroads has changed. Whereas the Burlington Northern and Atchison, Topeka & Santa Fe combined for only 30% of the grain and oilseeds originations in 1980, by 2007 the Burlington Northern Santa Fe (BNSF) had 42% of the market. This compares with a 31% market share held by Chicago & Northwestern, Union Pacific, and Missouri Pacific in 1980 that has decreased to only 19% for Union Pacific (UP) in 2007.

Figure 2: 2007 Railroad Grain Origination Market Share

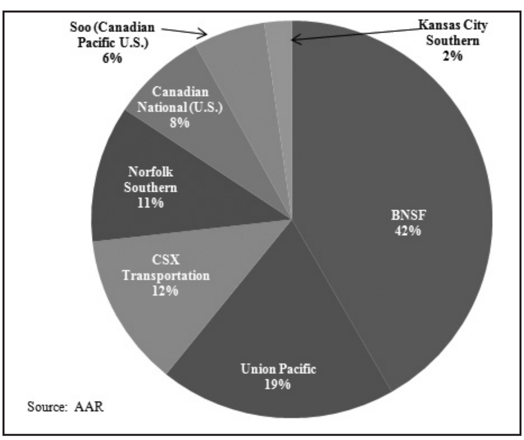
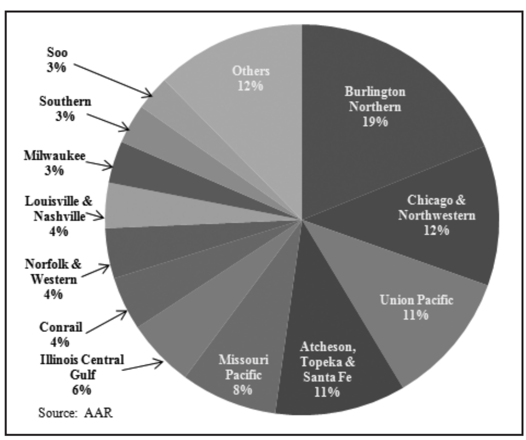


Figure 3: 1980 Railroad Grain Origination Market Share



Railroad concentration and market shares are even higher for specific markets. For instance, the top four Class I railroads transported 94% of the wheat in 2007 compared with only 80% in 1994 (see Figures 4 and 5). The market share for BNSF increased in comparison to its predecessors—54% in 2007 compared with 41% in 1994. UP market share in 2007 was only 20% in 2007 compared with 29% for its predecessors in 1994. The Soo (Canadian Pacific U.S.) market share increased to 11% in 2007 from only 4% in 1994, while the CSX market share increased to 9% compared with only 4% for its predecessors.

Figure 4: 2007 Railroad Wheat Origination Market Share

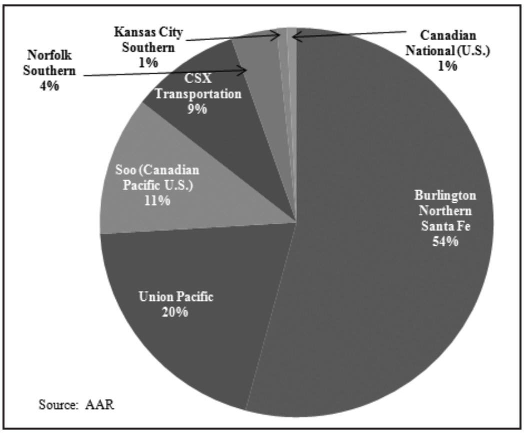
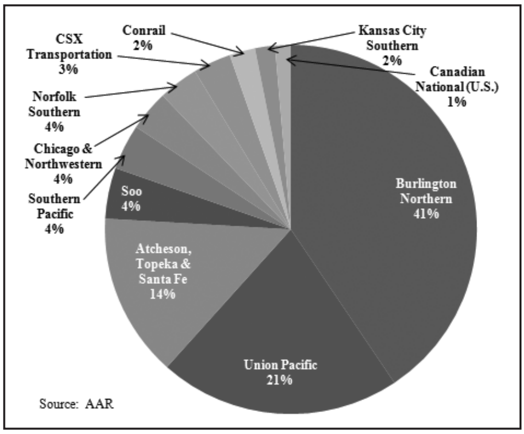


Figure 5: 1994 Railroad Wheat Origination Market Share



However, the level of rail-to-rail competition is not a function of the market concentration of railroads in the nation as a whole. Instead, it is a function of the quality and effectiveness of competitive options in particular markets. It is not only the number of competing railroads to which shippers or receivers have access, but also the effectiveness of competition from the other transportation modes.

INVERSE HERFINDAHL-HIRSCHMAN ANALYSIS OF RAIL-TO-RAIL COMPETITION

The Herfindahl-Hirschman Index (HHI) is a commonly accepted measure of market concentration. It estimates the ability of a firm to use market power. An HHI value, however, does not measure the actual use of market power. The HHI takes into account the relative size and distribution of the firms in a market and approaches zero when a market consists of a large number of firms of relatively equal size. It increases both as the number of firms in the market decreases and as the disparity in size among those firms increases.

Markets in which the HHI is between 1,000 and 1,800 are considered moderately concentrated and those in which the HHI exceeds 1,800 are considered concentrated. The maximum value of the HHI is 10,000, which occurs when one firm has a monopoly in the market. According to the U.S. Justice Department and the Federal Trade Commission (1997), transactions that increase the HHI more than 100 points in concentrated markets raise antitrust concerns under the Horizontal Merger Guidelines.

An inverse HHI, calculated by dividing 10,000 by the HHI, is often used to measure railroad concentration. The advantage of an inverse HHI is that it is easier to visualize the number of equivalent railroads with equal market shares that are competing in the market.

An inverse HHI is always one or greater. An inverse HHI of 1.00 is the equivalent of only one railroad competing in the movement of a commodity. An inverse HHI of 2.00 is the equivalent of two railroads competing, with each railroad moving half the tonnage. An inverse HHI of 3.00 is the equivalent of three railroads competing in the market, with each railroad moving a third of the tonnage.

The value of an inverse HHI also can be barely above 1.01 even when multiple railroads are competing in a CRD, depending on the relative market share of each. A market with two railroads, one of which carries 95% of the traffic, has an inverse HHI of 1.10. The value of the inverse HHI for a market with two railroads can range from 1.01 to 2.00. Likewise, the inverse HHI for a market with three competing railroads can range from 1.01 to 3.00.

Key Differences in This Analysis

This study has two major differences from the two recent studies by GAO (2006) and Laurits R. Christensen Associates, Inc. (2008) that used HHI to analyze railroad concentration and rates in markets.

Only the tariff rates have been used for revenue calculations because the STB has no jurisdiction over contract rates. The STB has jurisdiction only on tariff rates having a revenue-to-variable cost ratio of 180% or more. In addition, the actual contract rail rates are not shown in the Confidential Waybill Sample. Therefore, tariff rates were separated from contract rates using a "Contract Flag" field that STB provided to supplement the Confidential Waybill Sample. Due to contract confidentiality concerns, STB has provided the Unmasked Confidential Waybill Sample, which contains the actual contract rates, for only very few studies (i.e., Christensen and Associates [2008], GAO [2006]). Many other studies examining railroad revenues and rates were not able to remove the bogus contract rates because the contract field flag is not available in the Public Waybill Sample.

When calculating the tonnages for use in the inverse HHIs, data from all movements were used. Other studies used similar methods to calculate the inverse HHIs.

The second major difference in this study is that tonnages originated on smaller railroads connecting to only one other railroad were considered as part of the connecting railroad. This gives a more accurate portrayal of actual market share controlled by each railroad. When smaller railroads connected to two or more railroads, no attempt was made to assign the smaller railroads' volumes to a particular railroad. This is because little industry data are available regarding which railroads have contractual interchange commitments that may or may not strictly limit their ability to interchange with other railroads.

This study split the period from 1985 to 2007 into three multi-year time periods (periods were chosen to reflect structural changes in the industry) rather than using data for single years. This was done to obtain more CRDs having more than 30 observations, below which no results were reported for the CRD. The three periods include:

- Period 1: 1985–1992, an eight-year period representing the early years of deregulation, and including some important railroad mergers.
- Period 2: 1993–2002, 10 years that saw many mergers and the formation of the Eastern and Western railroad duopolies. Important operational issues arose during the implementation of these mergers.
- Period 3: 2003–2007, five years in which capacity constraints on the rail system first appeared, when the early retirement of engineers and conductors caused operational problems, and disruptions caused by storms were unusually severe. Major increases in rail rates due to capacity constraints and high fuel costs also occurred during this period.

An inverse HHI for originated tonnage by CRD was calculated and mapped for four major commodity groups:

- Grain and oilseeds
- Grain products including dried distillers grains with solubles (DDGS)
- Food products excluding grain products and DDGS
- Fertilizers

Analysis of Inverse HHI and Revenue-to-Variable Cost Ratio

As rail-to-rail competition decreases in a CRD, the market power of the railroads increases. A decrease in competition could result in higher rail rates and could give railroads the market power to change service terms. The revenue-to-variable cost ratio is an indicator of that market power.

The first part of this study used inverse HHIs to measure the degree of rail-to-rail competition in each CRD. The absolute value of the inverse HHIs and the degree of change are both important to an understanding of competitive status.

Inverse HHI Analysis

A markedly decreasing trend in rail-to-rail competition for shippers of the four product groups was found; many of the CRDs having higher inverse HHIs during Period 1 (1985 to 1992) moved to lower inverse HHIs by Period 3 (2003 to 2007) (see Table 1). The level of rail-to-rail competition for grain and oilseeds shippers decreased in 109 CRDs, and only 38 CRDs had an increase in rail-to-rail competition. The percentage of CRDs having a decrease in the level of rail-to-rail competition for the other three product groups ranged from 67% to 76%, which is similar to that of grains and oilseeds.

Revenue-to-variable cost ratios (R/VC) for all product groups shifted into the higher R/VC ranges. For grains and oilseeds, 108 CRDs (83%) had an increase in the R/VC ratio, but only 22 (17%) had a decrease. The percentage of CRDs experiencing increases in the R/VC ranged from 77% for food products to 83% for grain and oilseeds. The R/VC ratios were calculated by dividing the freight revenue field in the Confidential Waybill Sample by the variable cost field.

Table 1: Summary of HHIs and R/VCs for Four Commodity Groups Analyzed

		Change in HHI			Change in R/VC		No. of CRDs HHI =1	
Commodity Group		>0	Same	<0	>0	<0	Period 1	Period 3
Grains & oilseeds	No.	38	9	109	108	22	20	25
	%	24	6	70	83	17	10	15
Grain Products	No.	23	2	59	48	13	14	13
	%	27	2	70	79	21	11	13
Food Products	No.	25	11	113	84	25	20	40
	%	17	7	76	77	23	11	25
Fertilizers	No.	12	5	35	19	5	11	12
	%	23	10	67	79	21	13	21

Source: USDA analysis of STB Waybill Samples

The percentage of CRDs in which a railroad had a monopoly (inverse HHI equal to 1.00) for all product groups increased between Period 1 and Period 3. For grain and oilseeds, the number of CRDs with a rail monopoly increased from 20 (10% of the total CRDs) in Period 1, to 25 (15%) in Period 3. Eleven CRDs in this group of 20 had a change in the inverse HHI and nine had no change. Only two CRDs had an increase in competition since Period 1; one had an increase of 0.09 and the other had an increase of 0.80.[4]

R/VC Ratio Analysis

Increased competition is expected to result in lower rail rates. The percentage of CRDs having average R/VC ratios below 180 increases as the level of rail competition increases during periods 1 and 3 (Table 2). For example, during Period 3, only 50% of the CRDs that were served by a rail monopoly had average R/VC ratios below 180. In contrast, during the same period, 93% of the CRDs had average R/VC ratios below 180 when more than four strong railroads were competing. The finding that increased competition results in lower rail rates is consistent with the conclusions of studies by MacDonald (1989) and Harbor (2008).

Table 2: Grain and Oilseeds, Changes in R/VC Ratios by Inverse HHI[a]

		Number of Crop Reporting Districts			
Inverse HHI Range	**Revenue-to-Variable Cost (R/VC) Range**	**R/VC Period 1**	**Percent of HHI Range**	**R/VC Period 3**	**Percent of HHI Range**
1.00	< 100	0	0%	0	0%
	> 100 and <= 180	6	60%	5	50%
	> 180 and <= 240	4	40%	5	50%
	> 240 and <= 300	0	0%	0	0%
	> 300	0	0%	0	0%
> 1.00 and <= 2.00	< 100	8	12%	2	4%
	> 100 and <- 180	53	83%	43	78%
	> 180 and <= 240	3	5%	10	18%
	> 2.40 and <= 3.00	0	0%	0	0%
	> 3.00	0	0%	0	0%
> 2.00 and <= 3.00	< 1.0	7	15%	4	10%
	> 1.0 and <= 1.80	37	80%	29	75%
	> 1.80 and <= 2.40	2	5%	6	15%
	> 2.40 and <= 3.00	0	0%	0	0%
	> 3.00	0	0%	0	0%
> 3.00 and <= 4.00	< 1.0	3	11%	1	4%
	> 1.0 and <= 1.80	22	85%	20	87%
	> 1.80 and <= 2.40	1	4%	2	9%
	> 2.40 and <= 3.00	0	0%	0	0%
	> 3.00	0	0%	0	0%
> 4.00 and <= 5.75	< 1.0	1	6%	0	0%
	> 1.0 and <= 1.80	16	94%	13	93%
	> 1.80 and <= 2.40	0	0%	1	7%
	> 2.40 and <= 3.00	0	0%	0	0%
	> 3.00	0	0%	0	0%
Total Number of CRDs		163		141	
Maximum R/VC		198.62		228.56	
Minimum R/VC		65.17		68.98	

Calculated on tariff rail rates only when more than 30 observations in a CRD.
Source: Surface Transportation Board, Confidential Waybill Sample.

The number of railroads serving the grain and oilseeds market for each CRD is shown in Figures 6 and 7. Regions that changed to a rail monopoly since Period 1 include parts of Arkansas, Louisiana, Michigan, Mississippi, Ohio, Nebraska, Oregon, South Dakota, Tennessee, Virginia, Washington, and Wyoming. Many of the regions having inverse HHIs less than 2.0 are in areas of the country important in the production of grain and oilseeds and distant from barge-loading facilities.

Figure 6: Inverse HHI for Grain and Oilseed Shipments by Rail, 2003-2007

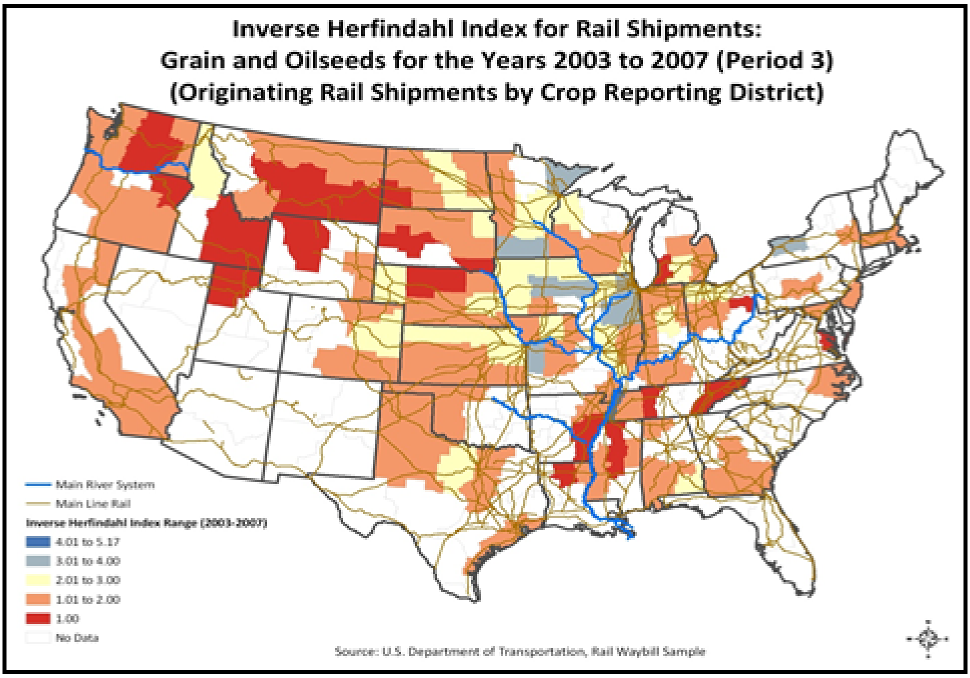


Figure 8 shows the changes in the inverse HHI by CRD. Major grain production regions that have gained rail-to-rail competition since Period 1 include northeast Minnesota, central and eastern Iowa, and the Dallas/Fort Worth region of Texas. CRDs that have lost the equivalent of 4.25 to 2.58 competing railroads include west central Missouri, western Tennessee, north central Indiana, parts of Ohio, and a portion of Texas. CRDs that have lost the equivalent of 1.41 to 2.58 competing railroads since Period 1 include parts of Arkansas, Colorado, Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Missouri, Nebraska, Ohio, Oklahoma, and Texas. All of these states were in the top 20 U.S. grain- and oilseed-producing states during 2007.

Figure 7: Inverse HHI for Grain and Oilseed Shipments by Rail, 1985-1992

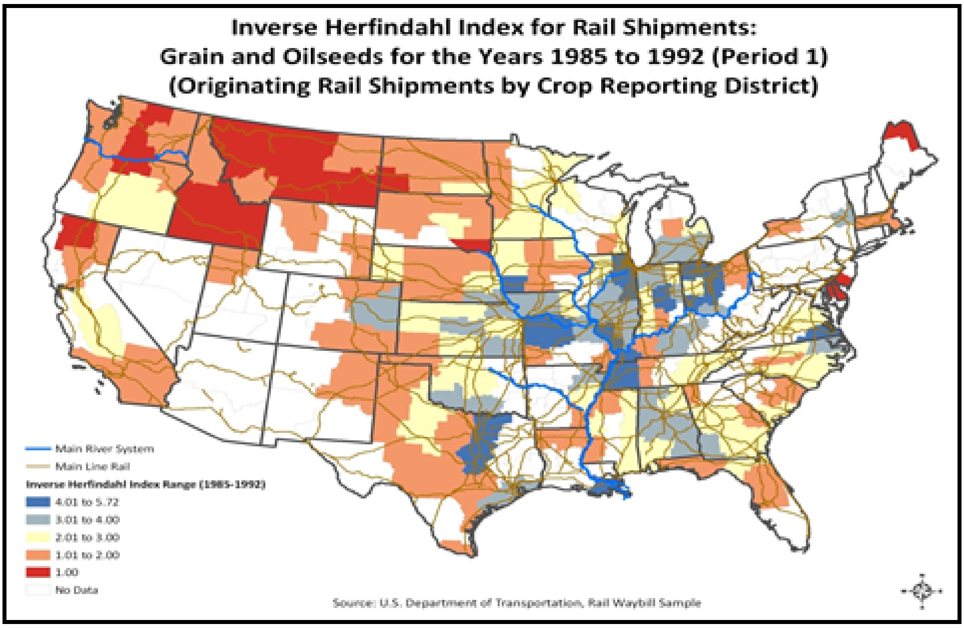
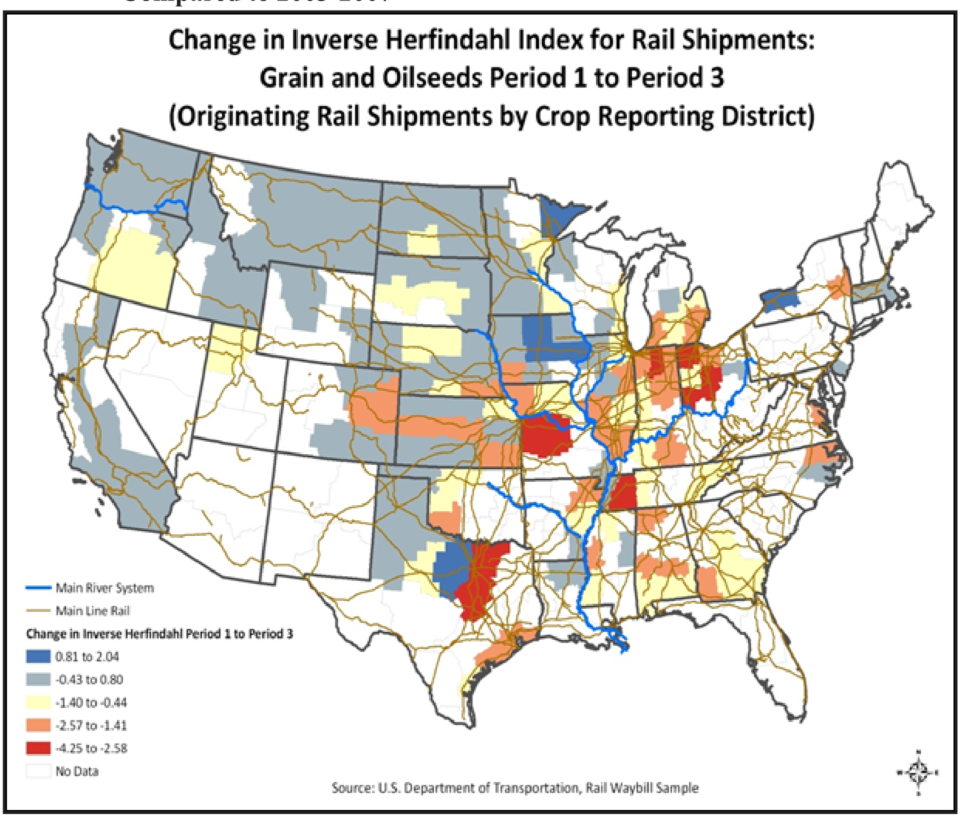


Figure 8: Change in Inverse HHI for Grain and Oilseed Shipments by Rail, 1985-1992 Compared to 2003-2007



Comparison of Rail-to-Rail Competition and Distance-to-Water Transportation by State

The study then examined annual statewide tariff rail rates from 1988 through 2007 for a group of six states with limited rail-to-rail competition and varying distances from barge-loading facilities, and a group of four states with more rail-to-rail competition and closer to barge loading facilities.

States in the first group—with less rail-to-rail competition and varying distances from barge-loading facilities—include Montana, North Dakota, South Dakota, Nebraska, Kansas, and Colorado. The average distance to barge-loading facilities from the middle of these states ranges from 200 to 850 miles. For states showing a range of distances to water, the shorter distances are to facilities on the Missouri, Arkansas, Snake, or Illinois Rivers; the longer distances are to facilities on the Mississippi or Ohio Rivers (Table 3). Barge movements on the Missouri and Arkansas Rivers have fewer cost efficiencies compared with rail transportation; barge movements on the Mississippi, Ohio, and Illinois rivers do realize cost efficiencies compared with rail.

All these states produce large amounts of grain and oilseeds. For instance, Nebraska is ranked third in the United States in grain and oilseed production, Kansas sixth, South Dakota seventh, North Dakota ninth, Colorado fourteenth, and Montana eighteenth.

Grain producers in Montana and North Dakota have raised issues for many years about high rail rates. Rates in Montana and North Dakota have often been higher than for South Dakota, Nebraska, and Kansas grain, even though the grain travels shorter distances over the same track to reach Pacific Northwest markets. Montana and North Dakota have appropriated funds to study grain and oilseed rail rates and to appeal those rates to the STB.

The states with more rail-to-rail competition and proximity to barge-loading facilities are Illinois, Indiana, Iowa, and Missouri. These states, also, are major grain and oilseed producers; Iowa is ranked first in the United States, Illinois second, Indiana fifth, and Missouri tenth. All of these states border the Mississippi or Ohio rivers, and the Illinois River runs through Illinois. The average distance from the middle of these states to barge-loading facilities is from 50 to 150 miles.

The levels of intermodal and intramodal competition, average length of haul, and time are among the factors expected to affect revenue per ton and revenue per ton-mile. The levels of intramodal and intermodal competition would be among the factors expected to affect the R/VC ratio.

In 1988, Montana and North Dakota paid the highest nominal (not adjusted for inflation) rail rates in the nation to move grain and oilseeds (see Table 3).[5] Montana grain shippers paid $25.41 per ton and North Dakota $22.61. Kansas shippers paid only $11.69 and Nebraska $17.59. The average rates for states with more transportation competition ranged from $9.06 to $12.12 per ton, but the average length of haul for those states was only 538 miles, compared with 994 miles for the states having less transportation competition.

By 2007, however, four states paid more to ship grain than shippers in Montana ($27.70): Nebraska paid $30.07 per ton, South Dakota $29.64, North Dakota $28.89, and Iowa $28.28. The average length of haul for the states paying more than Montana ranged from 1,150 to 1,710 miles (STB Waybill samples). Montana rates per ton had increased 8.3% and North Dakota 21.7% since 1988, but the rate increase for the other eight states shown in Table 3 ranged from 30.4% (Colorado) to 67.1% (Iowa). The greater shipping distances since 1988 for Iowa shippers contributed to the unusually large increase in their rate per ton (the average length of haul increased from 497 miles to 1,240 miles). By 2007, the average length of haul for the states having less transportation competition had increased to 1,198 miles, compared with only 845 miles for states having more transportation competition (STB Waybill samples).

Nominal tariff rates per ton-mile show that states lacking transportation competition do not always pay higher rates than states having more transportation competition; both transportation competition and length of haul are expected to be negatively related to the rate per ton-mile. In 1988, Montana (987 miles) and North Dakota (1,150 miles) paid the highest nominal tariff rates per ton-mile, but Illinois (460 miles), Indiana (561 miles), Kansas (612 miles), and Missouri (635 miles) paid the next highest rates (see Table 3). By 2007, Indiana (674 miles) paid the highest tariff rates

Table 3: Grain and Oilseeds, Comparisons of Nominal Tariff Rail Revenue per Ton and Ton-mile and R/VC by State (in $/ton)

State	Ave. Miles to Water Trans.	Avg. Haul (miles)		Revenue per Ton ($)			Revenue per ton-mile (cents)			Revenue/VC Ratio		
		1988	2007	1988	2007	Change	1988	2007	Change	1988	2007	Change
Lower levels of rail competition and distance from water transportation:												
Montana	400	987	930	24.51	27.70	2.29	2.58	2.90	0.32	186	187	1
North Dakota	410	884	1150	22.61	28.89	6.28	2.56	2.46	-0.10	166	191	25
South Dakota	200 - 340	1194	1710	18.41	29.64	11.23	1.54	1.95	0.41	117	151	34
Nebraska	250 - 530	1166	1488	17.59	30.07	12.48	1.51	2.10	0.59	108	148	40
Kansas	220 - 460	614	852	11.69	22.92	11.23	1.91	2.79	0.88	117	176	59
Colorado	500 - 850	1120	1056	18.34	26.34	8.00	1.64	2.82	1.18	125	167	42
Higher levels of rail competition and closer to water transportation:												
Illinois	50 - 90	461	806	9.06	16.82	7.76	1.97	2.27	0.30	115	151	36
Indiana	120	562	674	11.79	19.64	7.85	2.10	2.93	0.83	132	151	19
Iowa	150	496	1174	9.30	28.28	18.98	1.87	2.28	0.41	134	171	37
Missouri	125	634	727	12.12	19.73	7.61	1.91	2.73	0.82	108	162	54

Source: Surface Transportation Board, Confidential Waybill Samples.

per ton-mile, followed by Montana (930 miles), Colorado (1,056 miles), and Kansas (852 miles) (STB Waybill samples). The states having the least increase in tariff rates per ton-mile include North Dakota (with a 0.10 cent decrease), Illinois, South Dakota, Iowa, and Montana. Colorado, Kansas, and Indiana had the steepest increases.

An analysis of R/VC ratios based on tariff rates, which indicate the profitability of a movement for the railroads, shows mixed results relative to the amount of transportation competition. In 1988, Montana and North Dakota grain shippers had the highest R/VC ratios, at 186 and 166. The R/VC ratio for Montana was nearly 40% higher than it was for Iowa, which had the third highest R/VC rate among the 10 states selected for comparison (see Table 3). In 2007, North Dakota and Montana grain shippers still paid the highest R/VC ratio, and Kansas, Iowa, and Colorado paid the next highest ratios. In 2007, however, the Montana R/VC ratio was only 10% higher than Iowa's. The R/VC ratio for Montana increased 0.5% and North Dakota's 13% between 1988 and 2007. The R/VC ratio for Missouri increased 33%, that of Kansas 34%, and that of Nebraska 27%.

The use of statewide averages may have masked the relationship between transportation competition and R/VC. Prior studies by McDonald (1989a) and Harbor (2008), which are based upon individual waybills, show a relationship between rail-to-rail and intermodal competition and rail rates.

SUMMARY AND CONCLUSIONS

An affordable and reliable transportation network is necessary to maintain the strength and competitiveness of American agriculture and rural communities. Agricultural commodities are often produced in large quantities at locations distant from domestic and international markets, making rail a natural and preferred choice of transportation. Truck transportation is not cost-effective for many agricultural shippers, who are often located long distances from markets, and barge transportation is not an option for most. Rail is the only cost-effective transportation mode broadly available for many agricultural producers. Railroads transport nearly all of the grains and oilseeds produced in Montana, more than 70% of that produced in North Dakota, and more than 50% of that produced in Arizona, Oklahoma, and South Dakota.

Rail deregulation encouraged greater reliance on free markets to promote railroad profitability and public benefits, but relied on competition to protect shippers and the public. The extent of the loss of rail-to-rail competition due to railroad mergers, and the associated increase in market power, was not foreseen by many when the Staggers Act was passed. However, the abandonment of rail lines was a predictable outcome of railroad deregulation. Railroads under regulation were burdened by significant excess capacity. Deregulation permitted mergers and line abandonments, which eliminated overcapacity as a problem for railroads, and greatly increased railroad market power and profitability on remaining lines.

The preservation of competition as a substitute for regulation is vital for the economic prosperity of agricultural producers and shippers contending with a deregulated railroad industry. In deregulating the rail industry, however, Congress recognized that intermodal competition had the potential to be as effective as rail-to-rail competition in restraining the exercise of market power.

Railroad concentration for grains and oilseeds shipments has increased substantially since 1980 due to railroad consolidation and rail abandonment. The four-firm concentration ratio for U.S. railroad grain and oilseed traffic increased from 53% in 1980 to 84% in 2007. Market concentration is even greater for some individual commodities, such as wheat. The four-firm concentration ratio for U.S. railroad wheat traffic increased from 80% in 1994 to 94% in 2007.

Inverse HHI analysis shows the level of rail-to-rail competition for grains and oilseeds decreased significantly between 1985 and 2007. The number of competing lines decreased in 70% of the CRDs and increased in only 24%. The CRDs served by only one railroad increased from 10% in 1985 to 15% by 2007. As rail-to-rail competition fell, rail rates rose. The ratio of revenue to variable costs increased in 83% of the CRDs, but declined in only 17%.

The percentage of CRDs having average R/VC ratios below 180 increases as the level of rail competition increases. During period 3, only 50% of the CRDs that were served by a rail monopoly had average R/VC ratios below 180. In contrast, during the same period, 93% of the CRDs had average R/VC ratios below 180 when more than four strong railroads were competing.

Many grain- and oilseed-producing regions that are distant from barge-loading facilities changed to rail monopolies after deregulation. Many areas with less rail-to-rail competition are in regions important in the production of grain and oilseeds and are distant from barge-loading facilities.

Since the early 1990s, portions of west central Missouri, western Tennessee, north central Indiana, parts of Ohio, and a portion of Texas have lost the equivalent of 4.25 to 2.58 competing railroads. Parts of Arkansas, Colorado, Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Missouri, Nebraska, Ohio, Oklahoma, and Texas have lost the equivalent of 1.41 to 2.58 competing railroads. All were among the top-20 grain and oilseed producing states in 2007.

This study also examined annual statewide tariff rail rates from 1988 through 2007 for a group of six states with limited rail-to-rail competition and varying distances from barge-loading facilities, and a group of four states with more rail-to-rail competition and closer to barge loading facilities. States in the first group include Montana, North Dakota, South Dakota, Nebraska, Kansas, and Colorado. States in the second group include Illinois, Indiana, Iowa, and Missouri.

The levels of intramodal and intermodal competition, as well as length of haul, affect rail rates per ton. In 1988, Montana and North Dakota shippers paid the highest nominal tariff rail rates in the nation to move grain and oilseeds. By 2007, however, Nebraska, South Dakota, North Dakota, and Iowa all paid more to ship grain than Montana. The average length of haul for the states paying more than Montana ranged from 1,150 to 1,710 miles.

Nominal tariff rates per ton-mile show that states lacking rail-to-rail competition do not necessarily pay higher rates than states with more transportation competition; both the level of transportation competition and length of haul are expected to be negatively related to the rate per ton-mile. In addition, data analyzed at the state level can mask relationships that may be more apparent in analyses done at the CRD level. Such possibilities warrant further research and examination.

An analysis of R/VC ratios based on tariff rates, which indicate the profitability of a movement for the railroads, showed that Montana and North Dakota grain shippers had the highest R/VC ratios during 1988 and 2007.

Due to data limitations and time constraints, the authors were unable to do further analyses to examine more fully the relationship between rail-to-rail competition and R/VC ratios. Analyses that are more exhaustive are required, such as the R/VC ratios presented in this study are an average of the R/VC ratios for movements by tariff rates only. It is possible that some contract rail rates, which were not available for this analysis, are less than, are equal to or exceed the tariff rates in particular CRDs.

Endnotes

1. At the time of the Christensen study, the STB used the CAPM standard to evaluate the revenue adequacy of the railroad industry. On January 28, 2009, STB adopted a new measure that is the simple average of the CAPM and a multi-stage discounted cash flow method of estimating revenue adequacy.

2. Simply stated, revenue adequacy is the ability of a firm to attract adequate capital in competitive capital markets. Regulated firms are considered revenue adequate when the return on owners equity exceeds the calculated cost of equity capital.

3. USDA defines "efficient" as being cost-efficient; "effective" is the production of a product or service having the features and quality that consumers want.

4. Not all CRDs are represented in both Period 1 and Period 3. Thus, the numbers do not always tally.

5. We did not adjust for inflation because we are comparing relative rates state to state and not trying to show whether real rates have gone up or down for each state. We are looking at whether there are rate differences between states having more transportation competition and those having less transportation competition at the given points in time.

References

Association of American Railroads (AAR). *The Grain Book*, 1981, 1995.

Association of American Railroads (AAR). *The Rail Transportation of Grain,* 2008.

Association of American Railroads (AAR). *Railroad Facts*. 2008.

Association of American Railroads (AAR). "Economics and Policy Department." *Railroads and Grain*, 2009.

Association of American Railroads (AAR). *Railroad Ten-Year Trends.* Various editions.

Babcock, Michael W., L. Orlo Sorenson, Ming H. Chow, and Keith Klindworth. "Impact of Staggers Rail Act on Agriculture: A Kansas Case Study." *Proceedings of the Transportation Research Forum* 26(1), (1985): 364-372.

Babcock, Michael W., Marvin E. Prater, John Morrill, and Eugene R. Russell. "Short-Line Railroads' Performance." *Transportation Quarterly* 49(2), (1995): 73-86.

Bitzan, John, Kimberly Vachal, Tamara VanWechel, and Dan Vinge. *The Differential Effects of Rail Rate Deregulation: U.S. Corn, Wheat, and Soybean Markets.* Upper Great Plains Transportation Institute, June 2003.

Christensen, Laurits R. and Associates, Inc. *A Study of Competition in the U.S. Freight Railroad Industry and Analysis of Proposals That Might Enhance Competition.* Surface Transportation Board (2008): 8-54.

Chow, Ming H. "Interrail Competition in Rail Grain Rates on the Central Plains." *Proceedings of the Transportation Research Forum* 27(1), (1986): 164-171.

Cutler, John, Andrew Goldstein, G. W. Fauth III, Thomas Crowley, and Terry Whiteside. *Railroad Rates and Services Provided to Montana Shippers: A Report Prepared for the State of Montana.* February 2009.

Fuller, Stephen, David Bessler, James MacDonald, and Michael Wohlgenant. "Effects of Deregulation on Export-Grain Rail Rates in the Plains and Corn Belt." *Proceedings of the Transportation Research Forum* 28(1), (1987): 160-167.

Government Accountability Office (GAO). *Railroad Regulation: Current Issues Associated With the Rate Relief Process.* GAO/RCED-99-46, 1999.

Government Accountability Office (GAO). *Freight Railroads: Industry Health Has Improved, But Concerns About Competition and Capacity Should Be Addressed.* GAO-07-94, pg. 14, 2006.

Grimm, Curtis M. Testimony before the U.S. House of Representatives, Committee on Transportation and Infrastructure, Subcommittee on Railroads, March 31, 2004.

Grimm, Curtis M., and Clifford Winston. "Competition in the Deregulated Railroad Industry: Sources, Effects and Policy Issues." Sam Pelzman and Clifford Winston eds. *Deregulation of Network Industries.* Washington, D.C.: Brookings Institution Press (2000).

Harbor, Anetra L. "Competition in the U.S. Railroad Industry: Implications for Corn, Soybeans, and Wheat Shipments." Presentation by Marvin E. Prater at the 2008 Transportation Research Forum in Fort Worth, Texas, March 17, 2008.

Hovenkamp, Herbert. *The Antitrust Enterprise: Principle and Execution.* Harvard University Press, Cambridge, MA, (2005): 237-238.

Kahn, Alfred. STB Ex Parte No. 575, *Review of Rail Access and Competition Issues.* April 2, 1998: 73.

Koo, Won W., Denver D. Tolliver, and John Bitzan. "Railroad Pricing in Captive Markets: An Empirical Study of North Dakota." *Logistics and Transportation Review* 29(2), (1993): 123-137.

Kwon, Y.W., Michael W. Babcock, and Orlo L. Sorenson. "Railroad Differential Pricing in Unregulated Transportation Markets: A Kansas Case Study." *The Logistics and Transportation Review* 30(3), (1994): 223-244.

Lewis, William W. *The Power of Productivity, Wealth, Poverty, and the Threat to Global Stability.* University of Chicago Press, Chicago, IL, (2004): 91.

MacDonald, James M. "Competition and Rail Rates for the Shipment of Corn, Soybeans, and Wheat." *Rand Journal of Economics* 18(1), (1987): 151-163.

MacDonald, James M. *Effects of Railroad Deregulation on Grain Transportation.* U.S. Department of Agriculture, ERS Technical Bulletin No. 1759, Washington, D.C., 1989a.

MacDonald, James M. "Railroad Deregulation, Innovation, and Competition: Effects of the Staggers Act on Grain Transportation." *Journal of Law and Economics* 32, (1989b): 63-96.

Meyer, John R., Chair of Task Force, et. al. *Improving Railroad Productivity: Final Report of the Task Force on Railroad Productivity.* A Report to the National Commission on Productivity and the Council of Economic Advisors, Washington, D.C.,1973.

National Commission for the Review of Antitrust Laws and Procedures. *Report to the President and the Attorney General.* Washington, D.C., (1979): 177.

National Grain and Feed Association (NGFA). STB Ex Parte No. 658, *The 25th Anniversary of the Staggers Rail Act of 1980: A Review and Look Ahead.* October 19, 2005.

Porter, Michael. *The Competitive Advantage of Nations*. The Free Press, New York, NY, (1990): 117-120, 225-238, 416, 708.

Public Law 104-88, §10101 (1), (3), (4), (5), (6), and (12).

Surface Transportation Board (STB). Confidential Waybill Sample. (1985-2007).

Surface Transportation Board (STB). STB Ex Parte No. 627: *Market Dominance Determinations—Product and Geographic Competition*. Decision 29540, December 10, 1998.

Surface Transportation Board (STB). *Study of Railroad Rates: 1985-2007*. January 2009.

Thompson, S. R., R. J. Hauser, and B. A. Coughlin. "The Competitiveness of Rail Rates for Export-Bound Grain." *The Logistics and Transportation Review* 26(1), (1990): 35-53.

Tolliver, Denver and Alan Dybing. "Impacts of Transportation Infrastructure on the Economy of North Dakota." A report to the North Dakota Legislative Council by the Upper Great Plains Transportation Institute. 2007. p. 37.

Tye, William B. *The Transition to Deregulation: Developing Economic Standards for Public Policies*. Quorum Books, Westport, CT, (1991): 151.

U.S. Department of Agriculture (USDA). STB Ex Parte No. 627, *Market Dominance Determinations—Product and Geographic Competition*. Http://www.ams.usda.gov/AMSv1.0/getfile?dDocName=STELDEV3020178&acct=atpub, May 29, 1998.

U.S. Department of Agriculture (USDA). STB Ex Parte No. 582, *Public Views on Major Rail Consolidations*. Http://www.ams.usda.gov/AMSv1.0/getfile?dDocName=STELDEV3020172&acct=atpub, February 29, 2000.

U.S. Department of Agriculture (USDA). STB Ex Parte No. 646, *Rail Rate Challenges in Small Cases*. Http://www.ams.usda.gov/AMSv1.0/getfile?dDocName=STELDEV3020187&acct=atpub, July 21, 2004.

U.S. Department of Agriculture (USDA). STB Ex Parte No. 658, *The 25th Anniversary of the Staggers Rail Act of 1980: A Review and Look Ahead*. October 19, 2005.

U.S. Department of Agriculture (USDA), National Agricultural Statistical Service (NASS). *Crop Production*. 2005-2008.

U.S. Department of Agriculture (USDA). *Transportation of U.S. Grains: A Modal Share Analysis*. September 2006.

U.S. Department of Agriculture (USDA). *Updated data for Transportation of U.S. Grains: A Modal Share Analysis*. To be released in 2010.

U.S. Department of Agriculture and the U.S. Department of Transportation (USDA/DOT). *Study of Rural Transportation Issues*. Http://www.ams.usda.gov/AMSv1.0/RuralTransportationStudy, April 2010.

U.S. Department of Justice and the Federal Trade Commission. *Horizontal Merger Guidelines*. Washington, D.C.: (1997) §1.51.

Wilson, Wesley W. and William W. Wilson. "Deregulation, Rate Incentives, and Efficiency in the Railroad Market." B. Starr McMullen ed. *Transportation After Deregulation*. New York: Elsevier (2001): 1-24.

***Marvin Prater** is a transportation economist specializing in railroad transportation and rural infrastructure issues and is employed by the Transportation Services Division, Agricultural Marketing Service, U.S. Department of Agriculture. Marvin earned a B.S. degree in horticulture, a master's in business administration, and a Ph.D. in economics from Kansas State University. Marvin has been a member of the Transportation Research Forum since 1997, presented papers at the annual forum, and authored papers for the* Journal of the Transportation Research Forum *and the U.S. Department of Agriculture.*

***Ken Casavant** is a professor in the School of Economic Sciences and a nationally renowned transportation economist, after spending 35 years at Washington State University. He has received numerous teaching, research and service awards throughout his tenure at Washington State University, including recipient of the "Lifetime Achievement Award" from the Upper Great Plains Transportation Institute 2006, the "Washington State University Sahlin Faculty Excellence Award" 2004, and being named "Distinguished Scholar" by the Western Agricultural Economics Association in 2003. Ken also serves as the faculty athletic representative to the President for Washington State University.*

***Eric Jessup** is an associate professor in the School of Economic Sciences at Washington State University. His research focus and area of specialty is transportation economics and freight systems modeling. He is also co-principal investigator for the multi-year, statewide freight research and implementation project for the state of Washington, the Strategic Freight Transportation Analysis (SFTA) study.*

***Bruce Blanton** currently serves as Director of the Transportation Services Division with the U.S. Department of Agriculture's Agricultural Marketing Service. Over the years he has served in various positions in both government and industry. Past positions he has held at USDA include: deputy assistant secretary for Congressional Relations, special assistant to the Farm Service Agency administrator, special assistant to the under secretary for International Affairs and Commodity Programs, and confidential assistant to the deputy secretary. Blanton also worked as a senior analyst for agriculture for the United States Senate Committee on the Budget. In the private sector Blanton served as: executive director of the National Renderers Association, vice president of rendering sales for Moyer Packing Company, and as an economic research associate for the American Farm Bureau Federation. A native of the Southwest, Blanton holds a bachelor's degree in agricultural business management from New Mexico State University and a master's degree in agricultural economics from the University of Missouri-Columbia.*

***Pierre Bahizi** is an economist with the Transportation Services Division at the U.S. Department of Agriculture. Prior to working for USDA, Pierre was an economist in the Division of Consumer Expenditure Surveys at the Bureau of Labor Statistics. He has also worked in the Travel and Tours Department at the Smithsonian Institution. Pierre graduated from the University of Texas at Arlington with a B.A. in economics.*

***Daniel Nibarger** is an economist with the U.S. Department of Agriculture, Foreign Agricultural Service. Daniel is from Gardner, Kansas, and attended Kansas State University in Manhattan, Kansas. At Kansas State, Daniel earned a B.S. degree in economics with an international economics overlay, and an M.S. in economics with an emphasis in international trade. During his graduate coursework, Daniel spent a summer in Wuhan, Hubei Province, China, teaching English language courses at Huazhong University. After completing his M.S., Daniel earned a Legislative Fellowship with the State of Kansas Legislature. He has also served as a labor economist with the Kansas Department of Labor, researching ways to project employment and unemployment. Daniel has also served as an economist for the Transportation Services Division of the U.S. Department of*

Agriculture. Daniel has presented his research at the 2006 Western Economics Association-International annual meeting, the 2010 Transportation Research Forum, and the 2009 Caribbean Conference for Agricultural Trade.

Isaac Weingram *is a research assistant at the Federal Reserve Bank of Boston. Previously, Isaac worked as an agricultural economist assistant in the Transportation Services Division of the U.S. Department of Agriculture. He graduated from Washington University in Saint Louis with an A.B. in mathematics and economics.*

Transatlantic Port Issues

by Mary R. Brooks and Larissa M. van der Lugt

This paper examines differences between Northwest Europe and Eastern North America with respect to port commercial activities, port policy, port hinterland access and competition, port governance, and port security. While the volume of traffic on the transatlantic trade route has grown, its share of world trade with Europe has declined slightly. The authors explore the relevant contextual issues the ports on the trade route have experienced in the last decade, and using the Baltazar and Brooks (2001) Environment-Strategy-Structure framework, examine the structure and strategies followed by ports as they seek to deal with the changing environment (context) on all routes, not just the transatlantic.

INTRODUCTION

The port industry has undergone significant change in the last 25 years in large measure due to two significant revolutionary changes in global trade and transportation. The first of these was the widespread implementation of containerization. The year 2006 marked 50 years since Malcolm McLean first shipped "boxes" (really truck chassis) on the Ideal-X from New Jersey to Houston and irrevocably changed the way most manufactured goods are transported (Levinson 2006). Containerization was a key enabler of the wave of globalization that has occurred over the last quarter century. The need for terminal investment to support this wave meant many ports, desirous of participating in the burgeoning traffic growth that containerization promised, searched for ways to finance and build such investment. The second revolution was in the governance and management of ports. In the 1980s and 1990s, governments, around the world faced with increasingly unmanageable budgetary deficits, were seeking ways to reduce deficits and provide port services more efficiently while, concurrently, encouraging ports to be self-sufficient if not generators of revenue for government. The era of new public management and port reform began (discussed in Brooks and Cullinane 2007). The restructuring of port activities often included a move towards privatization or corporatization of ports, but most frequently led to a growth in the use of concessions to achieve these objectives. The most common concessions were for container terminal activities. The confluence of these two profoundly changed the cargo movement business and the role that ports play in this business, with traditional servicing of tanker and dry bulk activities remaining a core business activity.

This paper is about the port industry servicing the North Atlantic trade route. We use the generic Environment-Strategy-Structure configuration, as expanded by Baltazar and Brooks (2007), as a framework for its descriptive analysis of the port industry on both sides of the North Atlantic trade route. Core to this concept is that the performance of ports is related to the fit between environment (context) and, the structure and strategy of the port. While this paper will not focus on the output measure of performance, relating the context to the structure and strategy of ports provides the ability to explore the differences between ports on the U.S. East Coast of North America and those in Northwest Europe.

Using Eurostat data, the paper begins by examining the trade that underpins port demand on the transatlantic and how that trade is evolving; where possible, given data limitations, it examines the ports that service that trade. The focus then turns to the specific sector of port activity most influenced by the globalization activities of the last 25 years—containerization. Here, it is clear that the transatlantic trade is not the key trade route in global trade development, but that other factors influence the prospects for key Atlantic ports, and these will be discussed. As is the case for ports globally, port hinterland accessibility influenced by port congestion and security have become the critical context issues for ports in the region, so the next section will examine these issues more closely and what the Atlantic ports are doing to address them. As the ability of port authorities to respond to globalization, competitive pressures, and congestion is a product of the governance of the port, the following section focuses on the changing role of the state in port governance first in Europe and then in North America, leading to a discussion of current port governance, a structural issue. Finally, the paper will close with an examination of port management strategies in response to these structure and context changes. Throughout the paper, the authors explore the developments in the specific regional contextual issues the ports on the trade route have experienced in the last decade and examine the differences in management structures and resultant port management strategies between both sides of the Transatlantic maritime trade.

THE EVOLUTION OF TRANSATLANTIC TRADE

The transatlantic trade (defined as Eureopean Union–United States and Canada)[1] is the smallest of the major global container trades in manufactured goods. While initially much more dominant, as one of three global trade lanes in the early 1990s, it has lost its position as the Asian trades have grown to dominate. In the early 1990s, trade on three routes—the transatlantic, the transpacific, and the Asia–Europe—accounted for 50% of global container trade. With the rise of Asia, the transatlantic share of global container traffic fell to less than 5% by 2007 (Drewry Shipping Consultants 2008).

Europe–North America has long been one of the major trade lanes in the world's movement of goods. Total seaborne external trade of the 27 members of the European Union (EU-27) with North America accounts for about 6% of the total seaborne external trade of the EU (Table 1). While the absolute volume of traffic has grown over the period from 1999 to 2008 by 15.0%, North America's share of EU-27 trade with the world has declined as the EU-27's trade with Asia has grown, particularly in the 2000-2006 period.

Table 1: Europe – Canada/U.S. Trade Development 1999-2008 (Seaborne Trade in 000 Tons)

	EU-27 with U.S. and Canada			EU-27 with the World			U.S. Canada Share
Year	**Import**	**Export**	**Total**	**Import**	**Export**	**Total**	**%**
1999	86,670	77,464	164,135	1,865,300	572,069	2,437,368	6.73%
2000	90,918	90,954	181,872	2,001,493	613,417	2,614,910	6.96%
2001	85,636	86,546	172,182	2,061,248	569,370	2,630,618	6.55%
2002	77,507	91,847	169,354	2,062,969	610,481	2,673,450	6.33%
2003	79,539	90,508	170,046	2,174,356	642,697	2,817,053	6.04%
2004	81,500	96,405	177,905	2,340,770	670,171	3,010,941	5.91%
2005	79,234	98,382	177,616	2,407,780	718,723	3,126,503	5.68%
2006	81,346	91,581	172,927	2,468,422	728,593	3,197,015	5.41%
2007	89,512	85,632	175,144	2,551,411	746,568	3,297,979	5.31%
2008	107,654	81,093	188,747	2,529,735	789,239	3,318,974	5.69%

Source: Eurostat (2009)

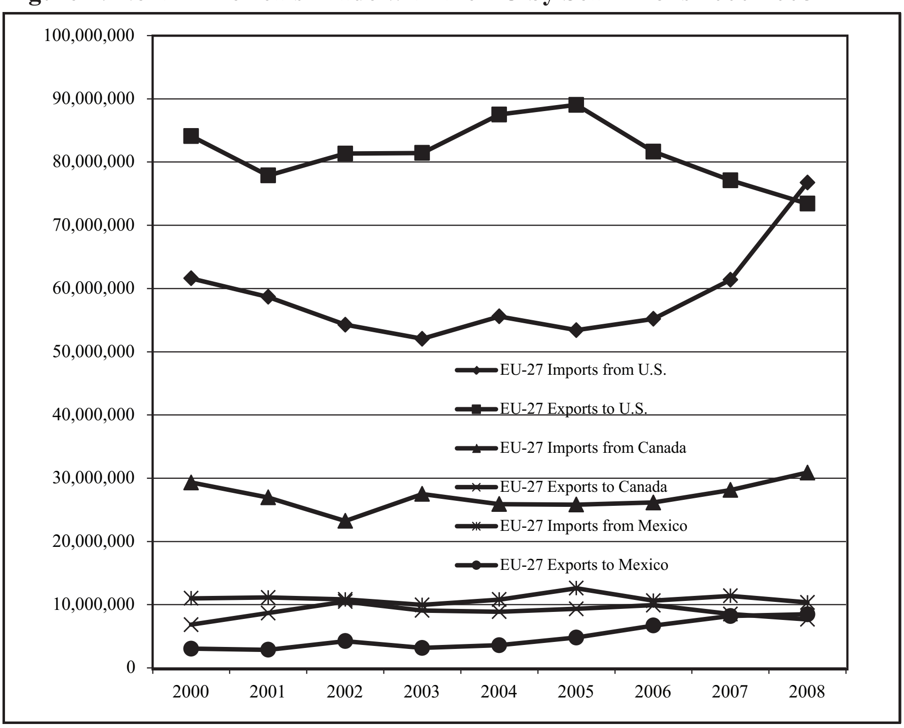
The figures in Table 1 do not show the decline in port traffic due to the global economic crises. Before September 2008, port volumes were softening and after that date, a significant and substantial decline in port traffic occurred; in 2009, most major ports in the region reported double-digit declines in traffic. As of January 2010, port volume levels have not recovered to their 2007 highs, although some recovery has occurred in some ports. The full impact of the global economic crisis on future port traffic is not yet clear.

Including Mexico and decomposing Europe's trade with Canada, Mexico, and the United States (Figure 1), the most stable have been the smaller volume of European exports to Canada and the trade with Mexico, while European exports to the United States have declined significantly. Since the recession of 2001-02, European imports from North America have risen, and most notably from the United States. Mexico's participation in trade with the E.U. has been smaller than Canada's but EU exports to Mexico have been rising in recent years and were similar to Canada's in 2008. Mexico's participation in transatlantic trade has been mostly through transshipment on the U.S. east coast or in the Caribbean.

In terms of trade in manufactured goods and component parts, the Europe–North America maritime container trade is still substantial in Europe's total maritime trade. However, the share has declined over the last two decades as the rise of China has altered global trading patterns, especially in the container trades. In the container trades, the transatlantic trade is the smallest of the four main trades, with a diminishing share over time as it is the slowest growing of the four; the fastest growing is Asia–Europe (13.0% 2006 over 2005) followed by transpacific (10.5%), intra-Asian trade (9.7%), and then transatlantic at 4.1% (Watson 2007). Comparing 2007 traffic over 2006, there was a retrenchment in both transpacific and intra-Asia trades, perhaps reflecting the softening of market conditions over 2007, while the Asia–Europe and transatlantic continued to grow, albeit at less than 5% (Drewry Shipping Consultants 2008). The transatlantic container trade, slow-growing and stable, reflects a mature trade route between developed countries.

For both the European and the North-American seaports, the transatlantic trade is a substantial one, although not dominating their traffic. Based on the individual statistics as published by the ports through their Web sites, it can be concluded that, for the Northwest European ports, the intra-European trades have the largest share. In most cases, the Asian trades follow. Antwerp and Bremerhaven are ports where the transatlantic trade dominates with about a 15% share, according to data available from individual ports' Web sites. On the North American side of the Atlantic Ocean, the traffic pattern is less clear. While consolidated calendar year 2007 data is available by port

Figure 1: North America's Trade with the EU by Sea in Tons 2000-2008



Source: Eurostat (2009)

(foreign and domestic) from the Data Navigation Center of the U.S. Army Corps of Engineers, the 2006 data are more detailed. However, that detail is by commodity (U.S. Army Corps of Engineers 2008), and does not distinguish whether the foreign origin or destination is Europe, Canada, or Asia. Therefore, the role of transatlantic trade for U.S. East Coast ports cannot be determined from existing public data sources. Likewise, the Canadian statistics available do not afford the opportunity to examine waterborne trade (as has been done using Eurostat statistics).

Trade Developments Within Europe

Western European markets continue to mature. The total market volume in Europe's most important countries and in traditional market sectors, such as consumer goods and automotive products, are showing moderate growth rates, contrasting with the booming growth in these markets in the 1970s and 1980s. The largest segment in Europe's maritime trade is liquid bulk. In the European Union, there are around 116 refineries, many of them located in seaports (Notteboom 2009). The Port of Rotterdam has the largest share in the transshipment of liquid bulk with a throughput of 194 million tons in 2007 (Port of Rotterdam Web site). Other large liquid bulk ports are Marseille (France), Bergen Ports (Norway), Le Havre (France), and Wilhelmshafen (Germany), followed by ports in Italy (Trieste, Augusta) and UK (Tees & Hartlepool, Milford Haven and Southampton) (ITMMA 2009). The liquid bulk trades are highly imbalanced with large import flows and relatively small export flows. The second largest segment is dry bulk, mainly consisting of iron ore and coal. These products are inputs for steel plants and electricity generating stations and are therefore highly captive to particular ports. The five most important ports are Rotterdam (Netherlands), Hamburg (Germany), Dunkerque (France), Amsterdam (Netherlands), and Antwerp (Belgium), followed by ports in the UK (Immingham), Spain (Gijon), and Italy (Taranto) (ITMMA 2009). Just like liquid bulk, dry bulk shows a significant imbalance with large import volumes and smaller export ones. Although containerized trades have a minor share in terms of volume, their impact on port systems is substantial in terms of growth rates and infrastructure demands. While demand in Western Europe might have matured, changing production and supply chain patterns have induced strong growth in container flows in and out of Europe. Globalization, outsourcing of production to low cost countries, and postponement in manufacturing and assembly have led to increasing containerized flows both intra-Europe and between Europe and other continents. In some ports, this meant that growth rates above 10% were achieved between 2004 and 2008.

The developing countries in Central and Eastern Europe have shown substantial economic growth over the last decade. The center of gravity for economic trade, therefore, is shifting from west to east, and likely to result in different trade patterns in the future. It is already clear that the northern German and Baltic ports already have benefitted from this shift.

Trade Developments in North America

The shipment of energy is the largest single trade handled at U.S. ports, with Philadelphia, New York, Portland, Baltimore, and Boston hosting 92.3% of tanker calls in the North Atlantic region of the United States in 2007, and the first two of these hosting 76.5% of the tanker calls. In the South Atlantic port range, five ports handled 72.4% of the tanker calls in 2007, but the business was more evenly spread (U.S. Maritime Administration 2009b). For Atlantic Canada, energy (including coking coal) was the largest single import and export, accounting for 92.0% of imports and 65.2% of exports for the four Atlantic Canadian provinces in 2005 (Brooks et al. 2009). In 2006, Port Hawkesbury in the Strait of Canso, Saint John, and Quebec City were the primary export ports for energy on the Canadian East Coast, while energy imports were serviced at a slightly wider range of ports, including Come-By-Chance and Sydney (Statistics Canada 2009).

The recent downturn in trade in dry bulk has affected ports on the North Atlantic. It is too soon to know the impact in terms of port statistics, but a significant volume of traffic at Canadian ports remains commodities. The largest destination (49.1%) for the Canadian overseas outbound volume in 2006 was Asia followed by Europe (28.5%). Total inbound shipments from overseas ports declined 5.2% in 2006, with energy shipments (of crude petroleum) from ports in both Europe and Africa, accounting for the majority of the decline. Of the tonnage arriving at Canadian ports from overseas, the largest share originated in Europe (30.7%) (Statistics Canada 2009: 15). Using another data source, it is perhaps interesting to note that for the 9.5% of 2007 Canadian traffic by value bound for West Europe and East Europe, air provides significant competition to the marine mode (Table 2) at the top end of the product value range. However, marine is the primary mode from a volume perspective.

Table 2: Canada's Exports by Origin, Destination and Mode, 2007 (C$000)

Total Exports Country of Destination	Eastern Provinces	Western Provinces	Total 2007	Main Modes Used (% of total value)
United States	244,748	109,461	354,210	Road (49), Rail (20)
Other Countries	56,521	37,834	94,355	Marine (64), Air (27)
West Europe	30,536	8,514	39,050	Marine (58), Air (38)
Asia	10,569	20,676	31,245	Marine (83), Air (15)
Latin America	7,466	4,121	11,587	Marine (44), Road (23)
Middle East	3,023	1,778	4,800	Marine (63), Air (27)
East Europe	2,004	847	2,851	Marine (51), Air (34)
Oceania	1,437	771	2,208	Marine (50), Air (36)
Africa	1,457	1,120	2,578	Marine (59), Air (23)
Other	30	6	36	Marine (66), Air (32)
Total	**293,263**	**145,667**	**438,931**	

Source: Transport Canada (2009), Table EC-12.

Development of Container Port Systems

The rising tide of container shipments over the past 25 years has been one of the most dominant trends noted at both European and North American ports. By 2007, three North European ports were in the top 15 of global container ports while no East Coast U.S. or Canadian ports could make that claim. The top 50 container ports included six North European ports—Rotterdam (ranked 6), Hamburg (9), Antwerp (14), Bremen-Bremerhaven (22), Felixstowe (30), and Le Havre (39)—but only two U.S. East Coast ports—New York New Jersey (19) and the consolidated ports of Georgia (41). Virginia had been ranked 48th in 2004 but disappeared from the 2007 rankings (Journal of Commerce 2008).

Table 3: Top 20 Container Lines and Service Patterns (2009 rank)

Rank (1)	Company	Major Ports Called (2)
1	A P Moller-Maersk	A, B, F, G, LH, M, NY, R, V
2	Mediterranean Shipping Co. (MSC)	A, F, G, LH, M, NY, R, V
3	CMA-CGM Group	A, B, F, G, LH, M, NY, R, V
4	Evergreen Group	A, B, F, G, LH, NY, R, V
5	APL	A, B, F, G, LH, M, NY, R, V
6	Cosco Container Lines	A, B, F, G, LH, NY, R, V
7	Hapag-Lloyd Containerline	A, G, H, LH, M, NY, R, V
8	China Shipping Container Lines	A, F, G, H, LH, NY, R, V
9	NYK	A, B, G, H, LH, M, NY, R, V
10	Hanjin	A, F, G, LH, NY, R, V
11	OOCL	A, B, G, H, LH, M, NY, R, V
12	MOL	A, F, G, NY, R, V
13	Hamburg Sud Group	A, B, F, G, LH, NY, R, V
14	K Line	A, B, F, G, H, LH, NY, R, V
15	Yang Ming	A, B, F, G, NY, R, V
16	CSAV Group	A, F, G, NY, R, V
17	Hyundai Merchant Marine (HMM)	A, B, F, G, LH, NY, R, V
18	Zim	A, F, G, H, LH, NY, R, V
19	Pacific International Lines (PIL)	A, R
20	United Arab Shipping	A, B, F, G, LH, NY, R, V

Notes: 1. Rank is based on TEUs of deployed shipboard capacity at the end of June 2009 as cited by AXS-Alphaliner. (http://www.axs-alphaliner.com/top100/index.php).

2. Port key: A=Antwerp, B=Bremen Ports, F=Felixstowe, G=Georgia Ports, H=Port of Halifax, LH= Le Havre, M=Port of Montreal, NY=NewYork New Jersey, R=Rotterdam, V=Virginia Ports. It was difficult to confirm Hamburg information and so Hamburg has not been included.

Source: Created by the authors based on information supplied by the port Web sites or port directories and any not listed on port Web sites were cross-referenced to the shipping line Web sites.

Most of the major global liner shipping companies (in the top 20) provide marine container services to the United States and to Europe, but not all provide services to Canada directly (Table 3). That said, most are concerned with servicing the Asian trade to the key ports and have Europe–Asia and transpacific services. The transatlantic service may be an extension of Asia–Europe services or part of a Suez service that calls only Mediterranean ports. As not all lines or ports provide detailed routing information, a more definitive conclusion is not possible.

Northwest European Container Port System

About 130 seaports in Europe handle containers; of these, 40 accommodate intercontinental container services (Notteboom 2009). The Northwest European container port range consists of both relatively large load center ports and a substantial number of regional ports. In the Mediterranean, a few purely transshipment ports exist, whereas pure transshipment ports are absent in Northwest

Europe. The emergence of the transshipment ports in the Mediterranean began in the latter half of the 1990s with Gioia Tauro, Malta, Algeciras, and Port Said growing strongly as the hub and spoke philosophy was adopted following the Asian ports' lead. Such growth was also certainly related to the rapidly increasing container flows between Asia and Europe. The already existing deepwater load center ports in Belgium, the Netherlands, and Germany, fulfilling the transshipment function, can explain the absence of new transshipment ports in Northwest Europe. Table 4 lists the 15 major ports in Europe with their throughput volumes and growth. Growth has been phenomenal at all Northwest European ports except Felixstowe, which struggled with its infrastructure investment plans during the period.

One of the characteristics of the ports in Northwest Europe is that they are located in relative proximity and, therefore, Europeans think of continental access in terms of gateway regions rather then individual ports (Notteboom 2009). Figure 2 shows Europe's multi-port gateway regions and how they are connected to their hinterlands. The contestable hinterlands of the Northwest European ports are substantial. Rotterdam, Antwerp, Hamburg, and, to a lesser extent, Bremen all have overlapping hinterlands in the central part of Europe, particularly Germany, Austria, and Switzerland. In Northeast Europe, ports like Hamburg and Bremen have benefitted from economic growth in central and eastern Europe, and become bigger ports as a result. Ports in the Baltic, like Saint Petersburg, Tallin, and Klaipeda, have shown steady growth figures over the last decade.

Table 4: Top 15 European Container Ports in 2008 and Growth Since 1995

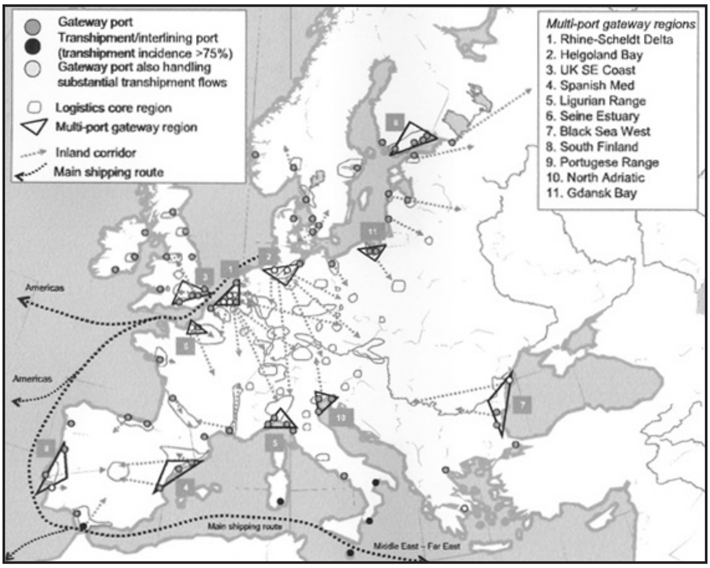
Rank 2008	European ports	1995 (000 TEU)	2008 (000 TEU)	Growth 1995 - 2008
1	***Rotterdam***	***4,787***	***10,784***	***125.3%***
2	***Hamburg***	***2,890***	***9,737***	***236.9%***
3	***Antwerp***	***2,329***	***8,664***	***272.0%***
4	***Bremen***	***1,518***	***5,448***	***258.9%***
5	Valencia	672	3,597	435.3%
6	Gioia Tauro (transshipment port)	1,155	3,468	200.3%
7	Algeciras (transshipment port)	970	3,324	242.7%
8	*Felixstowe*	*1,924*	*3,200*	*66.3%*
9	Barcelona	683	2,569	276.1%
10	*Le Havre*	*689*	*2,500*	*262.8%*
11	Marsaxlokk (transshipment port)	528	2,337	342.6%
12	*Zeebrugge*	*515*	*2,210*	*329.1%*
13	Genoa	965	1,767	83.1%
14	*Southampton*	*600*	*1,710*	*185.0%*
15	Piraeus	615	431	-29.9%
	Total top 15	**20,840**	**61,746**	196.3%
	Total Europe	**33,280**	**90,710**	172.6%
	Share Rotterdam	14%	12%	-14.3%
	Share top 3	30%	32%	6.7%
	Share top 15	63%	69%	9.5%

Note: The bolded and italicized ports are those located in Northwest Europe.
Source: Notteboom (2009).

The U.S. East Coast Port System

With the development of containerization by Malcolm McLean out of New Jersey, the Port of New York New Jersey had an early start in the race to dominate container port development in North America. Other ports on the eastern seaboard of North America were not far behind with new facilities being developed along the seaboard in the 1960s. With the development of stack train technology by American President Lines for rail service out of Los Angeles and Long Beach in 1984, the trade patterns shifted with a large volume of U.S. consumer demand being serviced via land bridge, that is goods transported out of Asia and over the U.S. West Coast ports and onward by rail to U.S. Midwest and East Coast markets.

Figure 2: Europe's Multi-port Gateway Regions



Source: Notteboom (2009)

In more recent years, the Panama Canal has become a means of accessing the eastern markets with all-water routes from Asia through the Panama Canal to Gulf and East Coast ports to more local truck-accessible destinations. By 1997, 20 U.S. container ports handled almost all of the U.S. container traffic (more than 95%), with 10 container ports on the U.S. East Coast handling about two-fifths of the total (Table 5). Growth over the last 12 years has been particularly strong, driven by consumer demand fueled by relatively easy credit, but there has been little change to the east–west coast split in terms of share of traffic, and four ports—New York New Jersey, Savannah, Norfolk, and Philadelphia—all acquired more than a doubling of volumes. The pattern of growth in container terminal throughput in the United States is on a scale similar to that of Europe.

Table 5: Top U.S. East Coast Container Ports in 2008 and Growth Since 1997

Rank 2008	U.S. Custom Ports on the East Coast	1997 TEUs	2008 TEUs	12-Year Growth
3	New York, NY	1,738,613	3,955,689	127.5%
4	Savannah, GA	530,261	2,106,437	297.2%
5	Norfolk, VA	770,790	1,584,632	105.6%
8	Charleston, SC	955,620	1,325,628	38.7%
11	Port Everglades, FL	454,504	680,841	49.8%
12	Miami, FL	623,492	669,493	7.4%
13	Baltimore, MD	260,553	430,331	65.2%
16	Philadelphia, PA	90,517	218,055	140.9%
18	Wilmington, DE	104,200	186,918	79.4%
20	Jacksonville, FL	199,438	158,452	-20.6%
	Total TEUs All Container Traffic	14,860,367	28,308,784	90.5%
	Total TEUs Top 20 Container Ports	14,208,159	27,254,527	91.8%
	Top 20 Share of Total Traffic	95.6%	96.3%	
	Total Top 20 EC Container Ports	5,727,987	11,316,478	97.6%
	East Coast Share of Top 20	40.3%	41.5%	

Source: Maritime Administration (2009a).

Approximately 75% of the U.S. population resides in the eastern part of the United States, and its needs have been filled with a three-coast supply chain strategy: land bridge from the U.S. (and Canadian) West Coast, Panama to Gulf and Southeast Coast ports, and transatlantic (including Suez) from Europe (and Asia). The advent of cross-docking and transload operations for containers have altered the trade flows, and thus the importance and development of land-side networks. Corridor development has been a dynamic process as the private sector railroads have competed fiercely with long-haul trucking to grow business and the largest global corporations engage in streamlining their global supply chain networks. The mature and stable transatlantic route has steadily grown throughout the period.

HINTERLAND CONGESTION AND ACCESSIBILITY

In North-America

This has been a critical issue for U.S. East Coast ports like New York New Jersey and Virginia ports but has been less of an issue for the Canadian ports of Halifax and Montreal, where inland connections have been well-developed from the very beginning and on-dock rail access was designed into the initial construction of the container terminals in the 1960s. Port access, however, has been such a problem for many U.S. ports that a major study of landside access was completed in the early 1990s (Transportation Research Board 1993). Examining the nature of the port access problems faced by U.S. ports (including appropriate strategies for dealing with it), the report identified four main areas to be addressed—physical impediments, land use policies, regulatory constraints, and institutional issues. In the intervening years, U.S. ports have looked to their local state and local governments for assistance in dealing with growing congestion in the immediate vicinity of ports and on the corridors connecting ports to the major hinterlands, in part because port policy in the

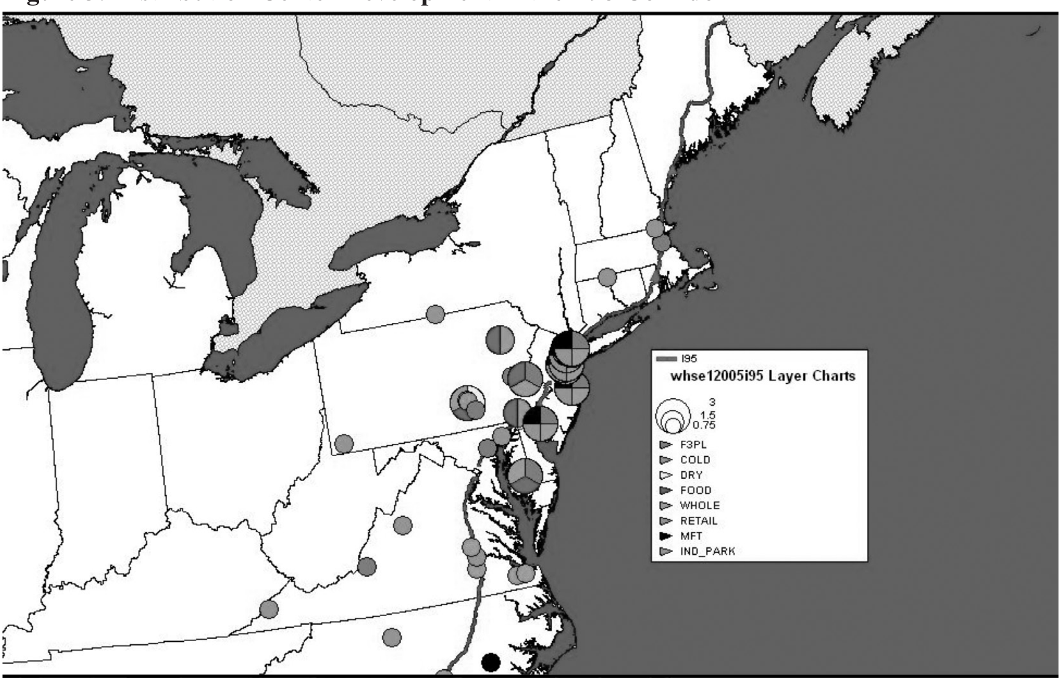
United Statesis decentralized (this will be discussed later) and in part because there is a history of local financial support for port activities.

Hinterland congestion has been a factor on the East Coast as truck has competed with car for scarce road capacity in the I-95 corridor. Distribution centers in the mid-Atlantic states (Figure 3) have been developed by major cargo interests to support efficient global networks as the volume of traffic has grown.

More recently, investments intended to streamline access have occurred in the Port of New York New Jersey and in the Heartland Corridor in an effort to address hinterland congestion issues. The Heartland Corridor rail line expansion (completion due in 2010) is the most visionary of the East Coast projects, and investments of US$309 million will improve the rail corridor from Virginia to Ohio in support the Port of Norfolk's hinterland development to the U.S. Midwest (Federal Highway Administration 2007).

In Canada, the Canadian government introduced a Gateways and Corridors Strategy in 2006 to fund infrastructure improvements, recognizing the critical nature of this infrastructure deficit. While the funding strategy began with investments on the West Coast, each region of the country was allocated a share of funds in the 2007 budget, which committed C$2.1 billion to gateways and corridors and provided a national gateway and trade corridor policy framework to guide the investment decisions (Government of Canada 2006). The exact allocation of these funds in support of the Atlantic Gateway but has not yet been determined.

Figure 3: Distribution Center Development in the I-95 Corridor



Source: Brooks, Hodgson and Frost (2006), Figure 2.4 p. 14.

Canada and the United States have not been alone in their efforts to address hinterland congestion arising from the burgeoning growth in container trade and the growing population that competes for the use of trade corridors. A recent study by the Joint Transport Research Centre (2008) examined the issue across Europe and North America, as part of a larger study, concluding it to be one of the most significant issues facing ports today and one that lacks consensus on the appropriate role for ports in addressing the challenge.

Congestion Issues in Europe

In Europe, hinterland congestion is a huge problem for almost all large load center ports and also for some of the regional ports located adjacent to large cities or within larger, densely populated areas. One of the factors behind this is that the evolution of both ports and cities in Europe has a very long history, but that an integrated approach on this only has emerged over the last few decades. Cities have expanded, putting much more pressure on space and infrastructure, and, on the other hand, ports have expanded, resulting in much more traffic to and from the port.

In a response to increased focus on hinterland accessibility and increasing port and hinterland congestion, we see new logistics concepts arising in Europe. Port and terminal operators, also large container carriers, develop terminal networks in the hinterland through which they transport their cargoes to final destinations. ECT in Rotterdam is developing an extended gate concept, with terminals in Venlo and Amsterdam (Netherlands), Duisburg (Germany), and Willebroek (Belgium) that serve as the gate of the deep-sea terminal. ECT, therefore, controls the transshipment including the inland transport up to the inland terminal. Rotterdam is also developing container transit points, purely for the purpose of avoiding congestion on the port's internal and access roads. Containers will be transported by barge from the deep-sea terminal to the transit point and from there by road or barge. The transit points will be developed in joint effort with port service providers and port users and aimed at cargo for the region.

Since January 2009, Eurogate in Hamburg also offers their customers transshipment service; they are developing an inland terminal network, comprised of a substantial number of mainly rail but also barge terminals, at which they offer a complete spectrum of container services, including warehousing of full containers, empty depot storage, container repair and maintenance, and sale and hire of containers.

As an issue, hinterland accessibility and congestion have reached top policy levels in Europe, and even the wider base of the Organisation for Economic Cooperation and Development and the UN Economic Commission for Europe. Whereas, in the beginning, the focus was rather on the hardware side, current interest in the issue has moved past infrastructure towards information systems, and institutional and organizational changes to enhance hinterland accessibility of ports. This is a common contextual factor that ports, not just in Northwest Europe and the U.S. East Coast, now face.

PORT SECURITY

North American Perspective

Prior to the terrorist attacks of 2001, the security efforts of port authorities in North America were mainly targeted towards preventing criminal activities like theft within the jurisdiction of the port. After the attacks, U.S. port authorities and their Canadian counterparts refocused their safety and security agendas towards prevention of incidents and emergency response.

In 2002, the United States passed the Maritime Transportation Security Act of 2002 (MTSA), with effect on July 1, 2004, the same day chosen by the International Maritime Organisation to effect the International Ship and Port Facility Security Code or ISPS Code. However, as many ports operate as landlords, leasing terminals and facilities to private companies, the MTSA limits the security oversight that U.S. port authorities may impose on their tenants. Security compliance, therefore, is primarily a matter between the U.S. government and the private terminal operating companies that lease port premises.

The United States has developed a significant security regime that has had extra-territorial impact. For example, the MTSA instituted rules, such as 96- and 24-hour rules, to reduce the vulnerability of the American container supply chain (GAO 2005). The 96-hour rule requires that all vessels that will call at U.S. ports provide the U.S. government with advance notice of arrival

96 hours before that arrival is expected, thereby allowing the U.S. government to assess the threat posed by the vessel. The 24-hour rule requires that non-vessel operating common carriers and liner shipping companies provide the U.S. government with 24-hour notice of a container being loaded onto a vessel in a foreign port, thereby allowing the U.S. government to assess the threat posed by the container, its contents, or the individuals who packed it. Both of these rules maximize the advantages of the Automated Targeting System used by U.S. Customs and Border Protection, a part of the Department of Homeland Security. The third plank in the prevention platform is the Transport Workers Identity Credential (TWIC) as required by the *Security and Accountability For Every (SAFE) Port Act of 2006*. Now more than one million workers have been credentialed.

These security requirements have been mirrored in Canada where possible so that goods traveling to the United States through Canadian ports are not disadvantaged. The Government of Canada has placed a high priority on ensuring that port competitiveness is maintained; in December 2001, the first U.S. customs officers were placed in the ports of Halifax, Montreal, and Vancouver as part of this effort. Security is an area where the United States has led in concept if not always in execution.

European Perspective

Following the security initiatives undertaken by the United States, the European Union fully agreed with the content of the ISPS code and transposed the associated rules into Community Law, with effect July 1, 2004 (Council Regulation (EC) No. 725/2004). Key to Regulation 725/2004 is that all operational ships and port facilities should have international security certificates issued by the Government as a proof of sufficient compliance with the ISPS Code (Dekker and Stevens 2007). In practice, this means that port facilities and shipping companies have to implement technical as well as organizational measures, such as access and entrance protocols, with their supporting electronic systems, container scan facilities, and additional specialized security personnel. Ports have to identify restricted areas and monitor them carefully to prevent unauthorized access, and implement measures to prevent weapons, dangerous substances, and devices being taken onto ships or into port facilities. Dekker and Stevens (2007) show that, by the beginning of 2004, already 22% of European ports fully complied with the ISPS Code and that the average compliance of all European ports was about 70%. Thus, only a relatively small effort was needed to fully implement the ISPS code-based Regulation 725/2004, as of July 1, 2004. More important, Regulation 725/2004 extends Code provisions to intra-European and domestic maritime trades.

The two major U.S. maritime cargo security initiatives, the Container Security Initiative (CSI) and the 24-hour Advance Notification Rule (24-hour rule), also have had their impacts on European port strategies, the latter being implemented as a compulsory rule. In the CSI program, foreign ports are asked to pre-screen the containers that are loaded onto vessels that will call at U.S. ports. In execution, the CSI program is dependent on the support of the European port authority. In 2007, there were 23 European CSI ports (including Rotterdam, Antwerp, and Hamburg in Northwest Europe) and the United States and Europe agreed to jointly work on the establishment of fast customs clearance processes at both sides of the maritime leg, meeting joint security requirements.

Following Council Regulation 725/2004 that focused on ships and port facilities, the European Commission developed a policy framework that minimizes security risks throughout the entire transport chain. A first step in this direction has been the endorsement of a revised Community Customs Code, which was set to be fully in force in 2009. Regulation 648/2005, which is similar to the U.S. C-TPAT program and the 24-hour rule, sets up a common EU secure custom system based on the electronic exchange of advance information between traders and customs authorities on all goods entering or leaving the European Union. Within this regulation, the possibility for transport operators to obtain a so-called Authorized Economic Operator (AEO) status is introduced. The criteria for obtaining this AEO status are: 1) compliance record with custom requirements,

2) satisfactory system of managing commercial records, 3) financial solvency, and 4) appropriate safety and security standards.

Since September 2001, Europe has mainly followed the U.S. security lead and complied with its intents by developing similar regulations, some divergence has occurred lately. The SAFE Port Act, endorsed in 2007 to be fully implemented in 2012, has led to some negative reactions by non-U.S. port stakeholders. Apart from the reciprocity issue, the SAFE Port Act requires every foreign port with U.S.-bound containers to do 100% container scanning with the appropriate scanning systems. The main criticisms are: a) the unavoidable separation of U.S.-destined containers requires more space, personnel, and operations; b) it is not a risk management approach; 3) it is not clear who will pay for the costs; and 4) it is not clear who will control the commercially sensitive data that are involved with the operations (Pallis and Vaggelas 2007). Ports, responsive to their stakeholders' concerns, have been less enthusiastic about blindly following the U.S. lead.

PORT POLICY, THE ROLE OF THE STATE IN PORT DEVELOPMENT

European Perspective

Specific to the European situation is that transport and port policy is not only a national issue but also an issue that is increasingly dealt with at a pan-European level. Over the last two decades, there has been a shift in the European Commission's role from advisor to interventionist in port policy development.

European Transport Policy. European transport policy started with reforms intended to phase in liberalization of the transport market in order to complete its internal market.[2] In the 2001 transport White Paper (Commission of The European Communities 2001a), the Commission's focus shifted from opening up the transport sector to making the transport system sustainable by decoupling economic growth from transport growth, and by working towards a more integrated European transport network. The primary mechanism for creating an integrated multimodal European transport network is the TEN-T project, which consists of a multitude of discrete projects to build an EU-wide transport network. In 2006, a mid-term review of the progress of the transport policy's implementation took place (Commission of the European Communities 2006). The main conclusions were that, in the first place, the policy recommendations of the 2001 White Paper were still valid, but that a broader more flexible approach is needed. The strict starting point of decoupling economic growth from transport growth was slightly relaxed, and the focus on modal shift has been replaced by a policy aimed at strengthening all modes and integrated transport systems; "co-modality" has become the new language. The midterm review furthermore concluded that a new European port policy was needed, further striving for a "level playing field" for ports and providing for the possibility of investments in ports that strengthen the ports and further integrate them into the European transport corridors and network.

The notion that continental maritime transport or short sea shipping can play a role in shifting cargo from road to more sustainable transport modes has resulted in the development of the "Motorways of the Sea" program. This has put the European ports more clearly on the agenda of the European Transport Policy. One of the concrete objectives, as mentioned in Commission of the European Communities (2007), is the integration of the European seaports into the TEN-T program.

Ports are now recognized as key components of freight logistics networks. The key obstacle to the integration of ports into the freight logistics networks is inadequate inland connections, in particular for rail. The 2009 Green Paper underlines that the future TEN-T infrastructure development policy should give particular attention to the appropriate development of a port's infrastructure and more efficient hinterland connections and to the removal of bottlenecks on major transport corridors (Commission of the European Communities 2009).

European Port Policy. Current European port policy traces back to the 1997 Green Paper (Commission of the European Communities 1997), the first significant policy document for European ports. It initiated a substantial discussion of a key element in European port policy—integration of the European ports in the European transport networks, market access to ports, and financing of and charging for port assets and services. Before 1997, attempts had been made by the European Commission to analyze these issues, but substantial progress had not been forthcoming.

Based on the Green Paper and subsequent discussions, a first port package on market access to port services was proposed in 2001 (Commission of the European Communities 2001b). It resulted in a Europe-wide strike among port workers and was finally rejected by the European parliament. In October 2004, the Commission launched a second proposal that again addressed market access as well as the issue of inter-port competition (Commission of the European Communities 2004). Transparency and state aid guidelines for the financing of port-related investments were expected to result in a level playing field for port services. Although there is common understanding that more transparency in port financing is needed, and that state aid to ports should be restricted, it has been very difficult to translate this into agreed regulation and guidelines. Studies into the financing structures of ports with the aim of identifying state aid that harms the level playing field of ports document that there exists a huge diversity in financing and charging systems of European seaports (Institute of Shipping Economics and Logistics 2006).

In Europe, port reforms started with a gradual increase in the role of the private sector in port services, with ownership and control of the assets remaining with the public sector. The public sector continued to subsidize port-related investments, resulting in a diverse system of public subsidies for port-related investments. Moreover, institutionally European ports differ enormously; even within countries, ports in close proximity can have different governance and ownership structures with port authorities. Therefore, a one-size-fits-all model cannot and will not work. Therefore, it was not surprising that the second port package was also rejected by the European Parliament and withdrawn by the Commission in 2006. The lesson learned was that a more flexible approach to port governance is needed, based on adequate consultation of all stakeholders involved. In October 2007, the European Commission presented its latest communication on ports policy, announcing a number of measures and soft law instruments but with few legislative changes proposed. The principles of the communication are 1) hinterland connections are a highly determining factor in port performance and deserve more attention; 2) expanding capacity is needed but should be done while respecting the environment mostly; 3) ports need a level playing field with clarity for investors, operators and users; 4) to integrate ports well in its civil environment structural dialogues are needed between ports and their adjacent cities; 5) work in ports should comply with societal needs in terms of labor safety and health (Commission of the European Communities 2007).

The real change in the European port policy has been the recent greater involvement of stakeholders in the dialogue, resulting in an understanding that hard legislation will not work due to the diversity in governance and financing and charging systems in ports, and that a soft law approach will better match the specific situation of ports. Common understanding is that a successful ports policy supports competition both within and between ports. This comprises clear rules for public contributions to port related investments, transparent access to port services and creating development potential for competitive services based on high quality labor. Furthermore, the ports policy should, on a structural basis, safeguard the balance between development needs and environmental constraints.

An interesting and yet not answered question is to what extent EU port policy really can influence processes in the Member States and whether it enhances the harmonization of port governance structures throughout Europe (Verhoeven 2009).

North American Perspective

Unlike Europe, there is no effort to have a supranational or continental port policy in North America. Canada and the United States have very divergent approaches to transport policy, and port policy in particular. The closest alignment in policy exists for rail and air modes, but even here coordination of policies is very limited. Nowhere is the divergence in thinking more evident than in the policies that govern maritime transport and ports.

Maritime Transport Policy. U.S. shipping policy is premised on the concept that shipping is a national strategic priority and that the defense of the United States is paramount. Ships carrying goods between two ports in the United States operate within a closed market; the goods must use a ship flagged in the United States, owned by U.S. citizens, crewed by U.S. nationals, and built in a U.S. shipyard. On the other hand, Canadian coasting trade (e.g., cabotage) rules require that the ship carrying domestic cargo must be crewed by Canadians and registered in Canada, or be a duty-paid foreign flag vessel operated under a waiver granted by the Canadian Transportation Agency. Vessels operating under waiver carry only about 4% of Canadian domestic traffic (Brooks 2009). For international traffic, however, both countries have open markets, although international containers, once landed in a U.S. port, are deemed to be domestic containers for purposes of considering the application of cabotage rules.

Port Policy. Again, port policy differs widely between the two countries. Ports in Canada are federally regulated and port policy for the largest ports—those seen to be of national economic importance (known as Canada Port Authorities or CPAs)—is prescribed in the Canada Marine Act. CPAs are required to be financially self-sufficient and are commercialized, non-recourse government agencies. This port policy was reviewed in 2004, and some changes were subsequently made to the regulations governing port access to financing for capital projects (Brooks 2007). On the other hand, there is no national port policy in the United States (Fawcett 2007). Ports are predominantly publicly owned, managed locally or regionally, are highly competitive, and subject to considerable local political influence.

State Aid. As already noted, Canadian ports must be financially self-sufficient; however, they are now able to access Gateway and Corridor funds (previously discussed) if they can present a business case as to why these investments will support Canada's international trade objectives. In the United States, there are limited national resources available for addressing port infrastructure investments, except for that available under such programs as the American Recovery and Reinvestment Act of 2009 (the economic stimulus package) or because inland connections can be improved through the use of funds from the Highway Trust Fund (raised by a gasoline tax). State aid is just that—supporting funds made available by the state or municipality in which the port resides. State aid for U.S. ports, therefore, may take the form of direct local government support, cross-subsidies from other local revenue-generating infrastructure (e.g., toll bridges, tunnels, or airports) or indirectly via federal subsidies for harbor dredging or federal subsidies in the form of tax-exempt bonds. Port state aid has long been a contentious point for Canadian ports that the playing field for inter-port competition is not level.

DIFFERENCES IN PORT GOVERNANCE

European Perspective

Port Governance. Over the last two decades, port governance has changed tremendously in Europe. Not only has the division of ownership and functions between public and private changed but also the legal and institutional position of the port authority organizations, leading to new and different management forms. Although the common objective of the reform programs has been to introduce more private sector involvement in ports in order to diminish the lack of financial transparency and involvement of national governments, and to give port authorities more autonomy, the resulting processes and outcomes have been very different.

During the 1990s, port reform in the United Kingdom often involved outright sale, even of regulatory functions (Baird 2000). This resulted in a number of private ports, whereby a private company has become both owner and operator of the port. On the European continent, port reform programs have resulted in so-called landlord ports, whereby a public or semi-public body is responsible for the development and management of the port area and nautical access, and private companies are responsible for port operations and services. Although we can say that, on the continent, port management is still dominated by public entities (Verhoeven 2009), the form in which this happens differs significantly by country. In the northern countries (Baltic, Germany, Benelux), we see municipal port authorities, which in some cases have been corporatized (Rotterdam, Bremerhaven, Gothenburg). In the more southern countries like France, port authorities remain tied to the national government. Although a lot of changes have been made, the reform process in many countries is still not complete.

The changing institutional and market environment has impacted the responsibilities and activities of port authorities (PAs). Competitive factors, such as port facilities, adequate draft and available land, are not sufficient for satisfying the requirements of port users. Additional competitive factors, such as a high-quality labor market, good hinterland access, and information and communications infrastructure, are in the collective interest of all firms in the port cluster, but are generally not provided without an active role for the PA (De Langen and Van der Lugt 2007). Thus, PAs are increasingly under pressure to invest in such competitive factors. This requires an institutional structure that ensures they can do so. For Rotterdam, this has resulted in corporatization of the port authority. The Port of Rotterdam acts as a corporate entity, but its shares are held by both municipal and national governments. This gives the Port of Rotterdam the freedom to develop new capabilities and acquire the financial means needed for making investments and developing new activities. In Amsterdam, discussions are also ongoing about whether or not to corporatize. Whether such governance reform will result in the desired outcomes remains to be seen.

North American Perspective

The framework of port governance in the United States is complex and very fragmented with a web of public and private organizations involved in management at national, regional, and local levels, each with differing priorities, requirements, and procedures (Newman and Walder 2003; Fawcett 2007). The multitude of jurisdictional forms has led to intense competition among ports and within ports in the United States. Furthermore, U.S. ports are heavily dependent on government (loans, grants and taxes) and tax-exempt revenue bonds for their revenue, and have been argued to be highly inefficient (Helling and Poister 2000). Unlike many other countries, there has not been a reform of port policy in the United States.

In Canada, on the other hand, there was significant and substantial reform of the port system in the 1990s, and the plan for reform was laid out in the National Marine Policy (Transport Canada 1995). The Government of Canada established three models of governance for ports in the country (Brooks 2007). The most important ports (19 in total), in terms of being vital to Canada's international trading role, were deemed to be Canadian Port Authorities, and mandated to be financially self-sufficient in their activities. Remote ports (27 as of May 31, 2009), where the federal government believed it had a public service obligation, were retained as federally owned and controlled. All other ports were de-proclaimed (seen as surplus to national needs) or divested to local or regional interests or to other departments or governments (such as the Department of Fisheries where the primary purpose was to serve the fishery not commercial trading interests) (Transport Canada 2009).

PORT STRATEGIES AND MANAGEMENT

European Perspective

In Europe, the changing market environments not only impacted the institutional structures of ports and port authorities but it has also led to reconsideration by port authorities of their role and strategies. A first strategic question many port authorities in northwest Europe have to deal with is: how do we manage new port expansions? Up to 2008, many of the leading ports in Northwest Europe faced scarcity in capacity. Development plans for up to 30 million TEU in capacity expansion existed in the Le Havre–Hamburg port range, but implementation horizons are very long and even enlarged by discussions on the societal impact of the expansion plans. Getting these plans through the approval process is a matter of smart port authority negotiation with national governments, strong stakeholder interactions, continuous support for the "license to operate" and, in many cases, implementing concrete environmental plans in combination with the physical expansion plans. One example is the Maasvlakte 2 project in Rotterdam; it was only approved to start building after dedicating a substantial part of the reclamation for "nature" and taking specific environment protection measures. The whole decision making process around Maasvlakte 2 was done with intensive involvement of all kinds of stakeholders. Port authorities are also developing specific concession strategies to better manage and control their port land area. They are considering putting in performance requirements and also sustainability requirements.

Another strategic issue port authorities are facing is how to ensure hinterland accessibility (De Langen 2008). As mentioned earlier, hinterland access has become a major issue in port performance and thus of interest to the port authority. In ensuring hinterland accessibility, port authorities are limited in what they can do as operations are done by private companies. Some consider investing in inland terminals (e.g., Port of Barcelona) and hinterland connections (see the previous discussion on Rotterdam), while others establish port community systems to support seamless supply chains or to set access rules on the ports' infrastructure. We conclude that Northwest European ports are very focused on developing hinterland strategies, carefully looking at the position they can take to enhance the performance of the ports' hinterland network.

In their strategy development, Northwest European port authorities increasingly consider whether to act on their own or to develop cooperation with other ports in proximity. Among various Northwest European ports, there are different cooperation initiatives. For example, cooperation was initiated on port expansion between the Port of Rotterdam and Zeeland seaports. Rotterdam and Amsterdam jointly developed a port community system called PortBase. Twice a year the management of the ports of Hamburg, Amsterdam, Rotterdam, Antwerp and Le Havre meet to discuss ongoing issues of mutual concern. The most far-reaching cooperative initiative has been between Copenhagen (Denmark) and Malmo (Sweden) that eventually resulted in a full merger between the two ports. The changed commercial and institutional environment certainly had its influence on these developments: port authorities are facing ever larger customers limiting their own power, and, from the societal side, more and more pressure is put on the performance of ports in terms of sustainability. On the other hand, institutionally, port authorities increasingly have more possibilities to widen their scope of operations.

FINANCING AND CHARGING REGIMES

There is considerable variety in financing and charging regimes in Europe, often related to the governance model that is in place. A study from ESPO into the various port financing and charging regimes in Europe (ESPO 2004) already illustrated this wide variety. At the moment this still holds. Although there have been many attempts at the European level to avoid state aids to ports and to achieve a level playing field in this issue, financial flows from the various governments to ports still occur and there are substantial differences in the financing systems by each country. Some ports are completely self-sustaining, including their sea channels and passage ways; examples are ports in Ireland and the Baltic. Some ports have a division between inside and outside the legally-defined port area, where the state is responsible for the investments in port-related infrastructure outside the port area, and the port authority has this responsibility inside the port perimeter. Examples can be found in Germany, Belgium, and the Netherlands. In some ports there is more structural financial support from the state into ports (e.g., France, Italy).

Following the landlord model, concessions are often awarded by port authorities with considerable variety in concession terms (Notteboom and Verhoeven 2009). There are a wide variety of factors contributing to the structure of concession agreements in Europe, and North America is no different.

Also charging regimes for port services differ widely. Port dues in European ports, in general, are tariff-based, with some involvement of the local or national government. However, depending on the port, there may be room for customer-negotiated rates. In some ports, the government sets charges while, in others, the port authority management controls the charges.

In North America, the situation is similar. The commercialization of Canadian ports through the Canada Marine Act, 1998, gave ports the responsibility to set their own charging regimes and the obligation to live with the financial impact. While many ports initially kept the structure and complexity of port tariffs, there is now a move to make these more commercially-oriented, both in "look and feel" and with favorable line reductions for business activities the port desires to promote, such as feeder traffic or traffic of a particular type. Market-driven tariffs are gradually becoming the norm.

In the United States, the public ownership of ports puts charging mechanisms in local hands. There is no national governance mechanism for ports and no standard approach to port charging. Like Europe, in some ports, government sets charges while, in others, the port authority management controls the charges. Some ports are more market-oriented and others less so.

In summary, there is no standard model for either financing or charging but, as with the world-wide trend towards concessions, it is always location-specific and dependent on the governance regime in place.

CONCLUSIONS

Although the northern transatlantic trade does not dominate the global maritime trade network, it is a substantial trade for both North American and Northwest European ports and shows a relatively low but stable share of global traffic. The port systems on both sides of the Atlantic differ mainly in terms of the number of ports that service each geographic region and how they are governed. They also differ in terms of port management strategies for dealing with the key environment (context) issues they face and the structures imposed by governance regulation, to use the eEnvironment–Structure–Strategy approach by Baltazar and Brooks (2007).

First, the U.S. and Canadian east coasts have a relatively small number of ports in comparison with Europe, where ports are larger and closer together. While this implies that the contestability of their hinterlands is more critical in the European context than in the North American context, this is not necessarily true. Both sets of ports compete as parts of supply chains and these may be continental or intercontinental in both cases. In the environment faced by ports, hinterland boundaries are set by economics rather than by politics and regulation, but the future success of port authority strategies may be dampened by lack of progress on harmonizing port governance in Europe (structure) and by border irritants in the case of North America (environment).

Second, there is the issue of how ports respond to the environment issue of security. Ports in North America and Europe began by adopting the U.S. and IMO security standards. However, as U.S. requirements continue to grow ever more stringent, the SAFEPort Act's requirement for 100% scanning has, for many European ports, become an unacceptable addition to the security regime.

Whether this strategy of acceptance will become one of resistance is unclear, but there is some pushback already in evidence.

In Europe, the ports have found that their environment is best managed through the development of a strategy that goes beyond the provision of a safe and sustainable port area and nautical access. The growing importance of uncongested and efficient hinterland accessibility has resulted in the development of coordination and cooperation strategies within the hinterland networks and between ports in proximity.

In the North American case, the reach of the freight railroads has meant most container ports have and serve a continental hinterland, and compete fiercely for it. There is little strategic cooperation among ports, as the issue of being in proximity is less the case in North America. Also, they are much less likely to be active participants or leaders in coordination strategies as their governance, in the case of Canada, limits their strategic product scope; in the case of the United States, their local political masters have not demonstrated an interest in expanding public service obligations.

Finally, the finance and charging strategies are location, and hence context-specific and structure-specific.

In conclusion, the two main issues that are increasingly relevant in the external context (environment) affecting the structure/strategy relationship of the transatlantic ports are port congestion in relation to hinterland accessibility and port security. Port congestion appears to be a bigger issue in Northwest Europe than in North America, perhaps as a function of the population density. Neither region has a streamlined approach to port governance, and this could well limit the development of appropriate strategies by port authorities, particularly in the North American context.

Endnotes

1. While geographically North America is considered to include Mexico, and this was further supported by the negotiation of the North American Free Trade Agreement, the transatlantic trade has always been defined to not include Gulf of Mexico or Caribbean traffic, and so Mexico is by usage excluded. Unless specifically included, assume it is excluded from this paper.

2. Sea transport was mentioned in the *Treaty of Rome* in the late 1950s, but was not part of Europe's Common Transport Policy, mainly because EC's focus was more on continental transport.

References

Baird, Alfred J. "Port Privatisation: Objectives, Extent, Process and the U.K. Experience." *International Journal of Maritime Economics* 2 (3), (2000): 177-194**.**

Baltazar, R. and M. R. Brooks. "Port Governance, Devolution and the Matching Framework: A Configuration Theory Approach." Mary R. Brooks and Kevin Cullinane eds. Devolution, Port Performance and Port Governance, *Research in Transport Economics* 17, (2007): 379-404.

Brooks, M. R., R. McCalla, A. A. Pallis, and L. Vanderlugt. "Coordination and Cooperation in Strategic Port Management: The Case of Atlantic Canada's Ports." Presentation to the 2009 Annual Meeting of the International Association of Maritime Economists, Copenhagen, June 2009.

Brooks, M. R. and K. Cullinane. "Introduction." Mary R. Brooks and Kevin Cullinane eds. Devolution, Port Performance and Port Governance, *Research in Transport Economics* 17, (2007): 3-28.

Brooks, M. R. “Port Devolution and Governance in Canada.” Mary R. Brooks and Kevin Cullinane eds. Devolution, Port Performance and Port Governance, *Research in Transport Economics* 17, (2007): 237-258.

Brooks, Mary R., J.R.F. Hodgson and J. D. Frost. Short Sea Shipping on the East Coast of North America: An Analysis of Opportunities and Issues. Halifax: Dalhousie University. Project ACG-TPMI-AH08, *Transport Canada*, 2006. Http://management.dal.ca/Research/ShortSea.php

Brooks, Mary R. “Liberalization in Maritime Transport,” Paper for the International Transport Forum 2009 on Transport for a Global Economy: Challenges and Opportunities in a Downturn. Joint Transport Research Centre of the Organisation for Economic Cooperation and Development, Leipzig, May 26-29, 2009. Http://www.internationaltransportforum.org/2009/workshops/pdf/ws1-Brooks.pdf

Commission of the European Communities. “Communication on a European Ports Policy.” COM616 final, 2007.

Commission of The European Communities. “Green Paper – TEN-T: A Policy Review – Towards a Better Integrated Trans-European Transport Network at the Service of the Common Transport Policy.” Com 44, February 2009.

Commission of the European Communities. “Proposal for a Directive of the European Parliament and of the Council on Market Access to Port Services.” COM 654 final, 2004.

Commission of the European Communities. Reinforcing Quality Service in Sea Ports: A Key for European Transport. COM 0035 final, 2001b.

Commission of the European Communities. “Communication from the Commission to the Council and the European Parliament. Keep Europe Moving –Sustainable Mobility for Our Continent.” Mid-term Review of the European Commission’s 2001 Transport White Paper [COM(2006)314]. Office for Official Publications of the European Communities, Luxembourg, 2006.

Commission of the European Communities. “Green Paper on Sea Ports and Maritime Infrastructure.” COM (97) 678 final, 1997.

Commission of the European Communities. “White Paper. European Transport Policy for 2010: Time to Decide” [COM(2001)370]. Luxembourg: Office for Official Publications of the European Communities. 2001a.

De Langen, P.W. “Ensuring Hinterland Access: The Role of Port Authorities.” OECD/ITF Discussion Paper, No 2008-11. Http://www.internationaltransportforum.org/jtrc/DiscussionPapers/DP200811.pdf (2008)

De Langen, P.W. and L.M. Van der Lugt. “Governance Structure of Ports in The Netherlands.” Mary R. Brooks and Kevin Cullinane eds. Devolution, Port Performance and Port Governance, *Research in Transport Economics* 17, (2007): 207-236.

Dekker, S. and H. Stevens. “Maritime Security in the European Union-Empirical Findings on Financial Implications for Port Facilities.” *Maritime Policy and Management* 34(5), (2007): 485-499.

Drewry Shipping Consultants. “Annual Container Market Review and Forecast 2008/09.” Http://www.drewry.co.uk/publications/view_publication.php?id=333 (2008).

ESPO. "Factual Report on the European Port Sector (2004-2005)." 2004. Http://www.espo.be

Eurostat. Comext External Trade Database. 2009. Http://epp.eurostat.ec.europa.eu/newxtweb/defaultquery.do

Fawcett, James A. "Port Governance and Privatization in the United States: Public Ownership and Private Operation." Mary R. Brooks and Kevin Cullinane eds. Devolution, Port Performance and Port Governance, *Research in Transport Economics* 17, (2007): 207-236.

Federal Highway Administration. *PPP Case Studies: Heartland Corridor*. Http://www.fhwa.dot.gov/PPP/heartland.htm (2007).

Government Accountability Office. *Homeland Security: Summary of Challenges Faced in Targeting Oceangoing Cargo Containers for Inspection*, GAO-04-557T. Washington, D.C.: Government Accountability Office. 2005.

Government of Canada. *Budget 2007: Chapter 5*. Ottawa: Government of Canada, October. Http://www.budget.gc.ca/2007/bp/bpc5be.html (2006).

Helling, A. and T.H. Poister. "U.S. Maritime Ports: Trends, Policy Implications, and Research Needs." *Economic Development Quarterly* 14, (2000): 300-317.

Institute of Shipping Economics and Logistics. "Public Financing and Charging Practices of Seaports in the EU." Http://ec.europa.eu/transport/maritime/studies/doc/2006_06_eu_seaports_study.pdf (2006).

Joint Transport Research Centre of the Organisation for Economic Cooperation and Development and the European Council of Ministers of Transport. "Port Competition and Hinterland Connections: Summary and Conclusions" (Discussion Paper 2008-19). Paris, April 10-11, 2008. Http://www.internationaltransportforum.org/jtrc/DiscussionPapers/DP200819.pdf (2008)

Journal of Commerce. "The JoC's Top 50 World Container Ports." 31 July (2008): 28-36.

Levinson, M. *The Box: How The Shipping Container Made The World Smaller and The World Economy Bigger*. Princeton University Press, Princeton, NJ, 2006.

Maritime Administration. "U.S. Vessel Calls by Type of Vessel." Http://www.marad.dot.gov/ports_landing_page/ports_landing_page.htm (2009b).

Maritime Administration. "U.S. Waterborne Foreign Container Trade by U.S. Custom Ports." Http://www.marad.dot.gov/ports_landing_page/ports_landing_page.htm, (2009a).

Newman, D. and J.H. Walder. "Federal Ports Policy." *Maritime Policy and Management* 30 (2), (2003): 151-163.

Notteboom, T. "Economic Analysis of the European Seaport System." Report for the European Sea Ports Organisation, Antwerp: ITMMA, University of Antwerp (2009).

Notteboom, Theo and Patrick Verhoeven. "The Awarding of Seaport Terminals to Private Operators: Current Practices and Viewpoints in European Ports." (Paper 2-68). E-Proceedings of the Annual Meeting of the International Association of Maritime Economists, Copenhagen, June 26 (2009).

Pallis, A.A. and G. K. Vaggelas. "Port and Maritime Security: A Critical Analysis of Contemporary EU Policies." Proceedings of the International Symposium on Maritime Safety, Security and Environmental Protection, Athens, Greece (2007).

Statistics Canada. “Shipping in Canada, 2006.” Ottawa: Statistics Canada. Cat. No.54-205-X. Http://dsp-psd.pwgsc.gc.ca/collection_2009/statcan/54-205-X/54-205-x2006000-eng.pdf (2009).

Transport Canada. “National Marine Policy 1995.” December (1995).

Transport Canada. “Port Programs Transfer Inventory (as of May 31, 2009).” Http://www.tc.gc.ca/programs/ports/transferinventory.htm, (2009).

Transportation Research Board. “TRB Special Report 238: Landside Access to U.S. Ports.” Washington, D.C.: National Academy Press (1993).

U.S. Army Corps of Engineers, “Waterborne Commerce of the United States 2006.” Http://www.ndc.iwr.usace.army.mil/wcsc/pdf/wcusatl06.pdf (2008).

Verhoeven, P. “European Ports Policy.” *Maritime Policy and Management* 36 (1), (2009): 79-101.

Watson, David. “The Atlantic Gateway and the Suez Opportunity.” Halifax: Speech to the Atlantic Provinces Transportation Forum, 12 October 2007.

Mary R. Brooks *is the William A. Black Chair of Commerce at Dalhousie University, Halifax, Canada. She is the founder and chair of the Port Performance Research Network, a network of more than 50 scholars interested in port governance and port performance issues. Her research interests include transportation and global supply chain management, and she is particularly focused on relationships between the buyers and sellers of transportation services. Dr. Brooks received her undergraduate degree from McGill University, her MBA from Dalhousie University, and her Ph.D. in Maritime Studies from the University of Wales in 1983.*

Larissa van der Lugt *has 14 years' research and project management experience in ports, transport and logistics. At the Department of Port Economics at Erasmus she is a senior researcher in the fields of port and maritime economics, port management and port-related logistics development, while working on her Ph.D. Before joining the economics faculty in 2003, she worked as a transport economist in a commercial BV of Erasmus, with the Dutch Ministry of Transport, and with a consulting and engineering company. In all these functions she worked, mainly as project leader, on research projects in port development analyses, logistics chain analyses, economic impact studies, and implementation projects for intermodal transport concepts.*

Bichou, Khalid. Port Operations, Planning and Logistics. *Lloyd's Practical Shipping Guides. London: Informa, 2009. ISBN 978-1-84311-805-3.*

Port Operations, Planning and Logistics

by Mary R. Brooks

By its title, the book promises a reader a solid tome on a subject of considerable practitioner interest. The marketing material suggests that this coverage will be "a comprehensive and detailed analysis" drawing on "established multi-disciplinary academic expertise." The table of contents confirms that we will be introduced to a wide variety of port topics of interest to the practicing port manager. For the most part, they are all here. In fact, the major contribution of this book is its comprehensive coverage and its multi-disciplinary approach. The author is to be lauded for the ambitious coverage he seeks to offer in this book.

Where this book excels are the two topics in which the author has done considerable research and consulting—benchmarking and port security. In these chapters, the text goes from the bigger picture to greater depth, and provides citations to the work of others, drilling down to the level expected by a seasoned academic reader. We also see evidence of why a port manager might choose one alternative over another, which approaches are more useful for a particular situation, and other such useful information for practitioners. Two potential audiences for the book will be engaged in these chapters.

One topic that is of keen interest to both managers and governments and not sufficiently dealt with is the topic of concessions, arising from financial securitization of terminal operations. Chapter 4 includes this topic under a short section entitled port financing and private sector participation. As global terminal operators consolidated their grasp on the management of container terminal operations over the past decade and more, the issue of concessions (or leases in North America) became extremely important. Port authorities did not want to give too much, terminal operators wanted to be sure their risks were manageable and remunerative, and governments wanted to be sure that they did not lose control of strategic assets, that the assets owned by taxpayers were being appropriately managed, and that return to the public purse was acceptable. A few examples of concessions are supplied, and the figures are informative, but the accompanying text barely scratches the surface of this very important topic. This is one topic where more would have been better, and more explanation is mandatory if a second edition is written.

As for port strategies, however, we are provided with the generic rather than the creative or innovative. It would have been of interest to this particular reader to have a well-developed discussion of how ports approach their long-term thinking on acquiring new business, retaining old, and deciding on whether to follow co-operative, competitive, or co-opetitive strategies. This, too, is a suggestion for the second edition.

The author, Khalid Bichou, is a well-known and widely recognized port consultant; a book of his should be of considerable interest to both practitioners and scholars. This would have been a better publication if the author had focused on those areas where his depth of knowledge could have been shared with the reader and his understanding of why particular decisions are made in certain ports and not others.

Where this book struggles is in defining its target audience. It reads like a textbook, with lists and illustrative figures, but fails as a textbook because it does not adequately explain many of the tables or figures; there is no glossary of words or acronyms or additional readings for each chapter. Furthermore, the list of additional readings is mixed with the bibliography and is not topic-specific. Students must not be its target audience. It is written, in part, like a scholarly tome but it does not meet the level of sourcing and citation expected of a serious academic work. So many tables, thoughts, and ideas are not adequately attributed to those who originally developed the ideas. From a scholarly perspective, the depth of coverage is uneven. Given these two off-target audiences, we must assume that its target audience is only the practitioner, hence its appearance in the Practical Shipping Guides list.

In summary, this book delivers a comprehensive, high-level overview but at a price that narrows its target audience. Although it is written like a textbook, the cost being charged by the publisher (US$441 not including duties or taxes) is sufficiently prohibitive that it will not be adopted as a text except for a rare course targeting industry executives. I can only assume that, given its price, the target audiences are library directors who will feel they must buy the book, scholars in the field of port economics whose libraries might pay the price or practitioners whose port coffers are rich enough to pay the price. While the first does not have the expertise to evaluate the value of the book beforehand, the second will be interested in the theoretical and practical content, but the last will not necessarily follow the economic models that it occasionally presents. Therefore, while the content will be of considerable interest to a limited number of practitioners, it will not provide them with the detailed value-added they seek for the price. In other words, this book appears to miss its possible audiences, mostly because of a poor pricing decision by the publisher. It is written for the student but only affordable by a few practitioners. Given this, it seems unlikely there will be demand for a second edition. That is a shame given the need for the book in the textbook market.

Dr. Mary R. Brooks *is the William A. Black Chair of Commerce at Dalhousie University. She has an undergraduate degree from McGill University, an M.B.A. in international business (Dalhousie University), and a Ph.D. in maritime studies (University of Wales). She is currently a Professor of Marketing and Transportation at Dalhousie University. Since 1979 she has held various teaching and research appointments at Dalhousie University, including research associate positions with the Dalhousie Ocean Studies Programme, the Canadian Marine Transportation Centre, and the Oceans Institute of Canada. Her latest book investigating trade in transportation services in North America was published by Edward Elgar in July 2008. Past service includes president of the Canadian Transportation Research Forum (1990-91).*

Schrag, Zachary M. The Great Society Subway: A History of the Washington Metro. *Baltimore: The John Hopkins University Press, 2006. ISBN 9780801882463*.

The Great Society Subway

by Joshua Schank

For transportation junkies in Washington, D.C.—and we are a small but devoted community—there probably could not be a more exciting topic for a book than the Washington Metro. It is the mobile laboratory many of us use every day while we think big transportation thoughts, and in many ways a defining characteristic of our region. It is one piece of transportation infrastructure that we cannot help thinking about and wondering about constantly, with constant questions about its history and operation.

Fortunately, Zachary Schrag has written a book that answers all of those questions and then some. Unfortunately, the book is a thorough documentation of the history of the Metro rather than a thoroughly compelling story line and thus will fail to hold the average reader. But for the true junkies, this book is a must-read.

The author demonstrates a clear affinity for not only Washington, a city where he grew up, but transportation issues as a whole. The book provides a clear, detailed, and relatively unbiased history of one of the largest post-war infrastructure projects in the nation (there is a bit of a pro mass transit, anti-highway undertone, but that should be expected in a book about Metro). It neatly summarizes many of the overall currents in U.S. policy (and in transportation specifically) that flow through the half-century covered by the text, in a way that is informative but not preachy. And it appears to be meticulously researched; it is obvious that the author devoted serious time and effort towards a book that became a labor of love.

The book begins slowly and carefully, providing the reader with much information that he or she needs to know in order to grasp many of the events depicted in later chapters. For example, the first chapter, which covers key historical events and trends in the city prior to 1955, is essential reading. But even the most detail-oriented Washingtonian may find the initial discussions of plans in Chapter 2 slow going. The maps provided offer interesting insights into what could have been, but far too much time is spent discussing the back and forth on various concepts and plans.

Things really start to pick up in what is the most fascinating part of the book from a political perspective – the discussion of the drama and maneuvering surrounding the Three Sisters Bridge. This bridge, which was to link Virginia and D.C. on a proposed spur from existing I-66, created a stir among residents concerned about the traffic it would funnel into Georgetown as well as the destruction of park land and views. The story of how a single Congressman —D.C. Appropriations Subcommittee Chair William H. Natcher—held up funding for what would become Metro in order to ensure this bridge was built is an extraordinarily compelling and untold story that alone makes this book worth reading.

From there, with the building of the Metro assured, the book really takes off and provides dozens of fascinating tidbits that denizens of D.C. will swallow with glee. Most everyone in D.C. and many tourists have heard of the myth that Georgetown fought to keep the Metro out of its neighborhood out of xenophobia; this book tells the real (and of course, more complicated) story. And what rider among us hasn't wondered about the naming of the stations, or the design of the Metro map? Schrag answers all of these questions and more.

The main drawback is that the book was clearly written as a history dissertation. This means that although it shows exceptional research and scholarship, it remains inaccessible to the average reader who just might have an interest in Metro. The book begins very slowly and gets enmeshed in arcane historical elements for a few chapters before finally beginning a compelling narrative. If it had been written in a more entertaining manner, it could have been an extremely popular and informative read in this region and nationwide. But as it is written it will instead serve only as the definitive history of the Washington Metro. One wonders if it might have been possible to accomplish both of those things.

Nonetheless, if you are interested in transportation, and particularly if you live or work or even occasionally visit the Washington region, this book is probably worthy of the substantial time investment it will take to get through it. And if you make your living in transportation, particularly in transportation policy, this is an excellent case study about how things get done and an absolute must-read.

Joshua Schank *is director of transportation research at the Bipartisan Policy Center, Washington, D.C.. He previously worked as a consultant with Parsons Brinkerhoff and as the transportation policy advisor to then Senator Hillary Clinton, working on the most recent reauthorization of the surface transportation bill (SAFETEA-LU). He has also worked as an analyst at the U.S. Department of Transportation Office of the Inspector General and as a transportation planner at the Metropolitan Transportation Authority in New York City. He has served as president of the Washington, D.C., Chapter of the Transportation Research Forum, and is now TRF's Public Relations V.P. Joshua has a Ph.D. in urban planning from Columbia University, a master of city planning from the Massachusetts Institute of Technology, and a B.A. in urban studies from Columbia University. He has published numerous articles on transportation policy and planning, and his first book, All Roads Lead to Congress: The $300 Billion Fight over Highway Funding, was published in October 2007.*

Sloboda, Brian W., ed. Transportation Statistics. *Ft. Lauderdale, FL: J. Ross Publishing, Inc., 2009. ISBN 978-1-60427-008-2.*

Transportation Statistics

by Bhuiyan Monwar Alam and Mohammed Abdulkaleem

This book, edited by Brian Sloboda, should be helpful to analysts using a wide range of data sources and methods to address transportation issues and appropriate remedies. It also provides a snapshot of various technologies that have been applied in statistical analysis of transportation data in the past two decades with the help of computers and discusses the significance of these developments for areas of research in transportation data.

In the first chapter, titled "Macroscopic Road Safety Modeling: A State Space Approach Applied to Three Belgian Regions," Van Den Bossche, Vanhoof, Wets, and Brijs discuss a state space approach which takes an account of different types of road lanes in networks such as motorways, provincial roads, and local roads, and their many properties associated with differences in road risks. The chapter develops a multivariate state space time series model based on data with its components and regression parameter estimates.

The second chapter, titled "Traffic Safety Study: Empirical Bayes or Hierarchical Bayes" by Miranda-Moreno and Fu, discusses the contrast between empirical bayes and full bayes in model parameters and accident risk estimates. The chapter suggests that full bayes are more flexible than empirical bayes for hot spot identification when working with limited observations; however, both empirical and full bays produce the same results when the amount of data is large.

Al-Deek and Emam, in the third chapter, explore the findings of a study of four random models: Weibull, Exponential, Lognormal, and Normal, and develop the best fit random model (which turns out to be Lognormal) that can be used for travel time of certain freeways. The authors argue that the existing methods are insensitive to the traveler's outlook on travel time and that the new method is sensitive to geographical location and congestion level.

The fourth chapter, by Hess and Polak, provides an overview of parking policy and its importance in travel demand management. Based on research on several city centers in the United Kingdom using a mixed multinomial logit (MMNL) model—which has the capacity to accommodate randomness and variety in travelers' taste—the authors conclude that variety in tastes is an important factor that can lead to different conclusions in terms of accessibility, time, and fines for illegal parking. The chapter also discusses technical issues related to heterogeneity, which is very significant in applying the MMNL model.

Cools, Moons, and Wets, in the fifth chapter titled "Modeling Daily Traffic Counts: Analyzing the Effects of Holidays," consider different modeling approaches to analyze daily traffic counts using the effects of holidays, which is an emerging area of research. The authors focus on two major issues. First, they use exponential smoothing and ARMA models for forecasting daily traffic counts and correlating successive traffic counts. The chapter provides a firm base to forecast future traffic counts. Second, the regression model presupposes that the daily traffic counts can be explained by other variables and focuses on a thorough check of holiday and day of week effects. Analysis is performed on the data collected from single inductive loop detectors, which conclude that weekly cycles influence the variability of traffic counts. The Box-Tiao model approach demonstrates that during holidays, the daily traffic flows are significantly lower. This model appears to perform reasonably well when daily traffic flow forecasts are required.

Chapter Six, by Metaxatos, focuses on the problems faced by transportation planners estimating and validating house trip generation rates from a small-scale household travel survey. It addresses three major problems: i) unusual observations, which are identified by traditional methods, ii) small number of observations, which are identified by classification and regression tree analysis, and

iii) missing observations, which are addressed using row column decomposition analysis. The author demonstrates these three problems using a small scale household travel survey and argues that the same procedures can be easily implemented with resources available to transportation analysts in Metropolitan Planning Organizations (MPOs).

There is significant progress in activity based micro-simulation of traditional travel demand. Mohammadian, Auld, and Yagi discuss these recent developments in Chapter Seven. They describe the traditional four-step travel demand models that make up the majority of models used in practice. They also describe activity based models, provide an overview of statistical procedures adopted in various models, and indicate further improvements in model techniques, validation, and transferability.

In Chapter Eight, which examines several simulation methods used in a variety of empirical analyses, Wang and Kockleman, suggest the use of pseudo-random Monte-Carlo (PMC) over quasi-Monte Carlo (QMC) and hybrid methods even when observations are correlated because of their better coverage.

The aftermath of the closure of the Glion Tunnel for repairs is discussed in Chapter Nine. Issues such as loss in tourism economy due to closures are addressed by Harvey's structural time series model. It uses indigenous variables, monthly series of tourism overnights, exogenous variables, and transportation data with monthly frequency. The chapter shows the differences in behavior before and after the closing of the tunnel. Scaglione uses the Chow test to identify break points and to determine approximate moments that become significant with behavioral changes. Morris provides an application of generalized linear models with maximum quasi-likelihood estimation to motor vehicle fatalities. The chapter deals with motor vehicle fatality risk as it varies in complex ways with interacting human, vehicle, and environmental factors. It explains the use of generalized linear models and quasi-likelihood estimation in analyzing motorcycle fatality rates (where use of helmets is required by law) while controlling for two climatic measures (using annual number of days of high heat and inches of precipitation as statistical proxies for residual variation in motorcyclist activities).

The research by Liu and Deng explores travel patterns during weekends. Applying a holistic approach, the researchers examine unique characteristics of weekend travel, evaluate travel survey data, and develop specifications for a statewide travel demand forecast model that can be included in existing long-range transportation planning (LRTP) processes at both metropolitan and state levels. Jin and Horowitz, on the other hand, discuss time-of-day models that deal with time at which travel occurs throughout the day. They present a study in time-of-day choice modeling for long distance trips with focus on transferability of models. The results indicate that departure time choice for long, occasional, and exceptional trips is more complicated than for urban short trips. Jin and Fricker, in the last chapter of the book, examine uncertainty as it applies to statewide travel-demand forecasting and land use models. The study investigates the sensitivity of Vehicle Miles Traveled (VMT) outputs to model parameters and input data using Indiana's statewide INTRLUDE model system.

This book provides a knowledgeable application of statistical and econometric methods to transportation data used in ongoing research. It gives priority to applied processes and describes how these methods can be successfully used and interpreted. It also provides awareness into the variety of data collection sources used in statistical analysis, interprets results, and analyzes the implication of those results. In terms of merit, this book stands alongside some other books on the subject matter – *Transportation Statistics and Micro-Simulation* by Spiegelman, Park, and Rilett, and *Statistical and Econometric Methods for Transportation Data Analysis* by Washington, Karlaftis, and Mannering. It is worth reading this book for a better understanding of transportation data analysis using statistical techniques that are specifically applicable to transportation.

Bhuiyan Monwar Alam *is an assistant professor of urban planning in the Department of Geography & Planning at the University of Toledo, Ohio. His current research interests focus on transportation statistics and spatial analysis, travel behavior, intelligent transportation systems, and GIS applications in planning. He is a transportation engineer and planner.*

Mohammad Abdulkaleem *is a graduate student in the Department of Civil & Environmental Engineering at the University of Toledo, Ohio. His research interests are in intelligent transportation systems and GIS applications in transportation planning.*

Transportation Research Forum

Statement of Purpose

The Transportation Research Forum is an independent organization of transportation professionals. Its purpose is to provide an impartial meeting ground for carriers, shippers, government officials, consultants, university researchers, suppliers, and others seeking an exchange of information and ideas related to both passenger and freight transportation. The Forum provides pertinent and timely information to those who conduct research and those who use and benefit from research.

The exchange of information and ideas is accomplished through international, national, and local TRF meetings and by publication of professional papers related to numerous transportation topics.

The TRF encompasses all modes of transport and the entire range of disciplines relevant to transportation, including:

Economics	Urban Transportation and Planning
Marketing and Pricing	Government Policy
Financial Controls and Analysis	Equipment Supply
Labor and Employee Relations	Regulation
Carrier Management	Safety
Organization and Planning	Environment and Energy
Technology and Engineering	Intermodal Transportation
Transportation and Supply Chain Management	

History and Organization

A small group of transportation researchers in New York started the Transportation Research Forum in March 1958. Monthly luncheon meetings were established at that time and still continue. The first organizing meeting of the American Transportation Research Forum was held in St. Louis, Missouri, in December 1960. The New York Transportation Research Forum sponsored the meeting and became the founding chapter of the ATRF. The Lake Erie, Washington D.C., and Chicago chapters were organized soon after and were later joined by chapters in other cities around the United States. TRF currently has about 300 members.

With the expansion of the organization in Canada, the name was shortened to Transportation Research Forum. The Canadian Transportation Forum now has approximately 300 members.

TRF organizations have also been established in Australia and Israel. In addition, an International Chapter was organized for TRF members interested particularly in international transportation and transportation in countries other than the United States and Canada.

Interest in specific transportation-related areas has recently encouraged some members of TRF to form other special interest chapters, which do not have geographical boundaries – Agricultural and Rural Transportation, High-Speed Ground Transportation, and Aviation. TRF members may belong to as many geographical and special interest chapters as they wish.

A student membership category is provided for undergraduate and graduate students who are interested in the field of transportation. Student members receive the same publications and services as other TRF members.

Annual Meetings

In addition to monthly meetings of the local chapters, national meetings have been held every year since TRF's first meeting in 1960. Annual meetings generally last three days with 25 to 35 sessions. They are held in various locations in the United States and Canada, usually in the spring. The Canadian TRF also holds an annual meeting, usually in the spring.

Each year at its annual meeting the TRF presents an award for the best graduate student paper. Recognition is also given by TRF annually to an individual for Distinguished Transportation Research and to the best paper in agriculture and rural transportation.

Annual TRF meetings generally include the following features:

- Members are addressed by prominent speakers from government, industry, and academia.
- Speakers typically summarize (not read) their papers, then discuss the principal points with the members.
- Members are encouraged to participate actively in any session; sufficient time is allotted for discussion of each paper.
- Some sessions are organized as debates or panel discussions.

Past Presidents

2009	Richard Gritta
2008	Kenneth Button
2007	John (Jack) V. Wells
2006	Anthony M. Pagano
2005	Scott E. Tarry
2004	John S. Strong
2003	C. Gregory Bereskin
2002	Martin Dresner
2001	Mark R. Dayton
2000	Richard S. Golaszewski
1999	Aaron J. Gellman
1998	Richard Beilock
1997	Robert Harrison
1996	Clinton V. Oster, Jr.
1995	Paul K. Gessner
1994	Russell B. Capelle, Jr.
1993	Louis A. LeBlanc
1992	Stephen J. Thompson
1991	Joanne F. Casey
1990	Frederick J. Beier
1989	Thomas N. Harvey
1988	Joedy W. Cambridge
1987	Frederick C. Dunbar
1986	Carl D. Martland
1985	Allan D. Schuster
1984	Douglas McKelvey
1983	William B. Tye
1982	Michael S. Bronzini
1981	Jay A. Smith, Jr.
1980	Samual E. Eastman
1979	Lana R. Batts
1978	Carl J. Liba
1977	Gerald Kraft
1976	Edward Morlok
1975	Barry A. Brune
1974	Richard Shackson
1973	Harvey M. Romoff
1972	Arthur Todd
1971	Herbert E. Bixler
1970	Paul H. Banner
1969	George W. Wilson
1968	Donald P. MacKinnon
1967	David L. Glickman
1966	Edward Margolin
1965	Robert A. Bandeen
1964	Gayton Germane
1963	Herbert O. Whitten
1962	Herbert O. Whitten
1961	Herbert O. Whitten
1960	Herbert O. Whitten
1959	John Ingram (TRF of NY)
1958	Herbert O. Whitten (TRF of NY)

Recipients of the TRF Distinguished Transportation Researcher Award

2010	Clifford Winston
2009	Daniel McFadden
2008	Steven A. Morrison
2007	José A. Gomez-Ibanez
2006	Tae H. Oum
2005	Kenneth Button
2004	Kenneth Small
2000	Michael E. Levine
1998	Edward K. Morlok
1997	Carl D. Martland
1996	Benjamin J. Allen
1995	D. Philip Locklin
1994	Martin T. Farris
1993	C. Phillip Baumel
1992	Alan A. Altshuler
1990	George W. Wilson, Ph.D.
1989	Sir Alan Arthur Walters, B. Sci., Hon. D. Soc. Sci.
1988	Karl M. Ruppenthal, Ph.D.
1987	William S. Vickrey, Ph.D.
1986	William J. Harris, D. Sci., Hon. D. Eng.
1985	John R. Meyer, Ph.D.
1984	Alfred E. Kahn, Ph.D.
1982	W. Edwards Deming, Ph.D.
1979	James C. Nelson, Ph.D.
1978	James R. Nelson, Ph.D.
1977	Lewis K. Sillcox, D. Sci., D. Eng., LL.D., D.H.L.

Recipients of the Herbert O. Whitten TRF Service Award

2010	Captain James R. Carman
2009	John (Jack) Wells
2008	Gordon MacDougall
2007	Robert T. Beard
2006	Gene C. Griffin
2005	Michael W. Babcock
2004	Jack S. Ventura
2000	Paul K. Gessner
1998	Arthur W. Todd
1996	Stephen J. Thompson
1994	Samuel Ewer Eastman
1991	Carl D. Martland, C.E.
1990	Herbert O. Whitten, M.B.A.

Past Editors of JTRF

Wayne K. Talley (1998-2000)
K. Eric Wolfe (1993-1998)
Kevin H. Horn (1993-1998)
Anthony M. Pagano (1987-1993)
Richard Beilock (1987-1993)

Recipients of the TRF Best Paper Award

2010 Tony Diana, *" Predicting Arrival Delays: An Application of Spatial Analysis"*

2009 Tae H. Oum, Jia Yan, and Chunyan Yu, *"Ownership Forms Matter for Airport Efficiency: A Stochastic Frontier Investigation of Worldwide Airports"*

2008 C. Gregory Bereskin, *"Railroad Cost Curves Over Thirty Years – What Can They Tell Us?"*

2007 Rob Konings, Ben-Jaap Pielage, Johan Visser, Bart Wiegmans, *"Container Ports and Their Hinterland: Multimodal Access and Satellite Terminals within the Extended Gateway Concept"*

2006 Ian Savage, *Trespassing on the Railroad.*

2005 Patrick DeCorla-Souza, *A New Public-Private Partnership Model for Road Pricing Implementation.*

2004 Ian Savage and Shannon Mok, *Why Has Safety Improved at Rail-Highway Crossings*?

2000 Ian Savage, *Management Objectives and the Causes of Mass Transportation.*

1999 *C. Phillip Baumel, Jean-Philippe Gervais, Marty J. McVey, and Takehiro Misawa, Evaluating the Logistics Economic Impacts of Extending Six Hundred Foot Locks on the Upper Mississippi River: A Linear Programming Approach*.

1998 Carl Scheraga and Patricia M. Poli, *Assessing the Relative Efficiency and Quality of Motor Carrier Maintenance Strategies: An Application of Data Entry Envelopment Analysis*.

1997 Wayne K. Talley and Ann Schwarz-Miller, *Motor Bus Deregulation and Racial/Gender Effects: Earnings and Employment*.

1996 Todd Litman, *Full Cost Accounting for Transportation Decision Making: Estimates, Implications and Applications*.

1995 Leon N. Moses and Ian Savage, *A Cost-Benefit Analysis of United States Motor Carrier Safety Programs*.

1994 Brian Shaffer and James Freeman, *Variation in Producer Responses to Automobile Fuel Economy Mandates*.

1993 W. Bruce Allen and Dong Liu, *Service Quality and Motor Carrier Costs: An Empirical Analysis*.

1992 Victor E. Eusebio, Stephen J. Rindom, Ali Abderrezak, and John Jay Rosacker, *Rail Branch Lines at Risk: An Application of the Exponential Survival Model on Kansas Duration Data*.

1991 Patrick Little, Joseph M. Sussman, and Carl Martland, *Alternative Freight Car Maintenance Policies with Attractive Reliability/Cost Relationships*.

1990 James A. Kling, Curtis M. Grimm, and Thomas M. Corsi, *Strategies of Challenging Airlines at Hub-Dominated Airports*.

1989 Cathy A. Hamlett, Sherry Brennan, and C. Phillip Baumel, *Local Rural Roads: A Policy Analysis*.

1988 John R. Brander, B. A. Cook, and John Rowcroft, *Auctioning Runway Slots: Long Term Implications.*

1987 *Clinton V. Oster, Jr. and C. Kurt Zorn, Deregulation's Impact on Airline Safety.*

1986 *Chulam Sarwar and Dale G. Anderson, Impact of the Staggers Act on Variability and Uncertainty of Farm Product Prices*.

1985 Brian D. Adam and Dale G. Anderson, *Implications of the Staggers Rail Act of 1980 for Level and Variability of Country Elevator Bid Prices*.

Donald J. Harmatuck, *Back Haul Pricing: Mathematical Programming and Game Theory Approaches to Allocating Joint Costs*.

Jeffrey L. Jordan, Dieudonne Mann, S. E. Pressia, and C. Thai, *Managing the Transportation of Perishable Commodities: Temperature Control and Stacking Patterns*.

1984 K. Eric Wolfe, *An Examination of Risk Costs Associated with the Movement of Hazardous Materials*.

1983 Denver D. Tolliver, *Economics of Density in Railroad Cost-Finding: Applications to Rail Form A*.

1982 Jeffrey Beaulieu, Robert J. Hauser, and C. Phillip Baumel, *Inland Waterway User Taxes: Their Impacts on Corn, Wheat and Soybean Flows and Transport Costs*.

Journal of the Transportation Research Forum
Guidelines for Manuscript Submission

1. Submit manuscripts by e-mail or hardcopy to either General Editor at:

Michael W. Babcock
Co-General Editor, JTRF
Department of Economics
Kansas State University
Manhattan, KS 66506 U.S.A.
Phone: (785) 532-4571
Fax: (785) 532-6919
Email: mwb@ksu.edu

Kofi Obeng
Co-General Editor, JTRF
Dept. of Economics & Transportation/Logistics
School of Business and Economics
North Carolina A&T University
Greensboro, NC 27411 U.S.A.
Phone: (336) 334-7231
Fax: (336) 334-7093
E-mail: obengk@ncat.edu

2. E-mail submissions must be in MS WORD.

3. Final manuscript and abstract (hard or e-mail copy) must be received as soon as possible.

4. The text of manuscripts is to be double-spaced. Use one side of each page only. Articles are limited to a maximum length of 30 pages; industry issue papers and comments are limited to a maximum of 15 pages.

5. The manuscript should have a title page which includes the names, affiliations, address (mailing and e-mail) and phone numbers of all authors. Brief biographical sketches for all authors should be included with the manuscript.

6. The abstract should briefly describe the contents, procedures and results of the manuscript, not its motivation, and should not exceed 100 words.

7. Endnotes are to be used rather than footnotes, used sparingly and placed at the end of the manuscript. Do **NOT** use the endnote feature of the word processing software.

8. *The Chicago Manual of Style* is to be used for endnotes and references. At the end of the manuscript, complete references are listed alphabetically (not by number) by author surname, government agency, or association name. Use the following examples.

   Book:

   Jones, Robert T. *Improving Airline Transportation*. General Publishing Company, Washington, D.C., 2002.

   Chapter in a Book:

   Bresnahan, T.F. "Empirical Studies of Industries with Market Power." R. Schmalensee and R. Williq eds. *Handbook of Industrial Organization*. Amsterdam: North Holland (1989): 1011-1057.

   Journal Article:

   Kane, Louis R. "Grain Transport Rate Projections." *Transportation Journal 2*, (1999): 20-40.

   Journal Article with Multiple Authors:

   Chow, G., R. Gritta, and E. Leung. "A Multiple Discriminant Analysis Approach to Gauging Air Carrier Bankruptcy Propensities: The AIRSCORE Model." *Journal of the Transportation Research Forum* 31 (2), (1991): 371-377.

Ph.D Dissertation:

Jessup, E.L. "Transportation Optimization Marketing for Commodity Flow, Private Shipper Costs, and Highway Infrastructure, Impact Analysis." Dissertation (Ph.D). Washington State University, 1998.

Government Document:

U.S. Senate. Committee on Foreign Relations, *Investigations of Mexican Affairs*. 2 vols. 66th Cong., 2nd session, 1919-20.

9. Tables, figures, graphs, and charts should be titled and numbered by an Arabic numeral (e.g., Figure 2). All figures and charts should be submitted as **separate** files. Any graphics or photos must be 300 dpi and submitted as a separate .tif or .eps file.

10. Headings should adhere to the following style. Place the abstract below the title of the manuscript.

    - First level headings should be all caps, left justified.
    - Second level headings should be initial caps, left justified.
    - Third level headings should be initial caps on the same line as the first paragraph which it heads.
    - Do not number the headings.

11. Submit your final copy in "MS WORD" only.

12. For proper preparation of all charts, figures, maps and photos call Beverly Trittin, North Dakota State University, at (701) 231-7137 or bev.trittin@ndsu.edu prior to submitting the final version of your paper. The journal is printed in black and white, therefore, all graphics must be submitted in black and white, or grayscale.

    TIPS:

    - Photographs to be submitted as 200 dpi or higher resolution (.jpg, .tif , .eps format).
    - Images taken from the web are 72 dpi and not print quality. Secure high resolution image if possible.
    - Do not embed graphics in Word – submit as individual files (.jpg, .tif, .eps format).
    - Improper graphics or images will be returned to author for correct formatting.

Book Review Information

Books for review should be mailed to:

Dr. Jack Ventura
Book Review Editor
Surface Transportation Board
1925 K Street NW, Suite 500
Washington, D.C. 20423
Phone: (202) 245-0319
Fax: (202) 245-0454
Email: jack.ventura@verizon.net

JTRF Information

The Journal of the Transportation Research Forum is distributed to members of the Transportation Research Forum.

ISSN 1046-1469

The Transportation Research Forum gratefully acknowledges the contributions of sponsoring and sustaining members.

Supporting Members

Michael Babcock, Kansas State University
Michael Bronzini, George Mason University
Dolores R. Buckenberger, TranSystems
Paul K. Gessner, Gessner Transportation Consulting LLC
Robert Harrison, University of Texas at Austin
Larry Jenkins, Embry-Riddle Aeronautical University
Jack S. Ventura, Surface Transportation Board
John (Jack) V. Wells, U.S. Department of Transportation

Sustaining Members

Gene Griffin, Upper Great Plains Transportation Institute
Carl Martland, Massachusetts Institute of Technology