JOURNAL of the TRANSPORTATION RESEARCH FORUM

Volume 51, Number 3 Fall 2012

TRF

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 51, Number 3 • Fall 2012

Table of Contents

On the cover: The U.S. has four million miles of roads and streets that have to be maintained at an adequate service level. In "Pavement Pre-and Post-Treatment Performance Models Using LTPP Data," Pan Lu and Denver Tolliver find that differences in pavement deterioration in severe weather regions are greater than in less severe weather regions, and that pavement deterioration increases with freeze-thaw cycles.

A Message from the JTRF Co-General Editors

A Message from JTRF Co-General Editors

The Fall 2012 issue of the *Journal of the Transportation Research Forum* contains the articles below:

- Assessment of Sustainable Infrastructure: The Case of Exurban Dallas
- Applying the Highway Safety Manual to Two-Lane Road Curves
- Predicting Block Time: An Application of Quantile Regression
- A Dynamic Programming Optimization Approach for Budget Allocation to Early Right-of-Way Acquisitions
- Pavement Pre- and Post-Treatment Performance Models Using LTPP Data
- Productivity Improvements in the U.S. Rail Freight Industry, 1980-2010
- Stopping Behavior of Drivers at Stop-Controlled Intersections: Compositional and Contextual Analysis.

Arvidson and co-authors write about sustainable infrastructure to answer two questions: "When is a particular infrastructure project sustainable and how are its impacts assessed?" In answering these questions they focus on two exurban communities and a control site outside of Dallas, Texas, and use a number of indicators to assess the impacts of sustainable infrastructure, where sustainable infrastructure includes street improvements such as sidewalks on both sides of the streets. The authors divide the impact area of each improvement site into three concentric rings whose distances apart vary by the site and are chosen to reflect typical walking distances. They then use economic, environmental, and social indicators drawn from the literature to assess the impacts of the sustainable infrastructure they study. The economic indicators include business density, employment density, property value, income and sales tax revenue, and vacancy rate. For environmental indicators, the authors used housing stock, land use mix, sidewalk density, and street density. The social indicators are average daily traffic, household density, population density, residential ethnicity, and walking and bicycling trips. The authors found consistency between their results and the expected and desired outcomes of dense development as well as diverse land use mix and compact circulation.

In the second paper, Findley et al. evaluate the Highway Safety Manual's (HSM) crash prediction model for horizontal curves with data on two-lane roads in North Carolina. The authors note that these models require more data than previous models and must be calibrated and validated for the state where they are to be applied. Further, they note that the model is credible, has been approved by a committee of experts, and its documentation provides instructions on how to apply it. By applying the model, the authors intend to inform decision-making by providing practical advice in how to use it to analyze horizontal curves to improve safety. They perform the application by preparing a field investigation form, which was distributed to North Carolina Department of Transportation (NCDOT) personnel to collect data on 21 variables for each of the 50 horizontal curves selected by NCDOT. The HSM model was then used to predict crash frequency, severity, and types of crashes. Among the findings are that there are not statistical differences between reported and predicted collisions; a large number of sites are needed for the calibration to meet the recommendations in the HSM; and annual average daily traffic (AADT), curve radius and curve length are the most important predictors of crashes in horizontal curves.

The third paper is by Tony Diana and it presents a quantile regression model to predict block time for airlines using data on the Seattle/Tacoma International Airport. As he argues, this method is most appropriate for skewed data; that is, it is more robust in handling outliers than do ordinary regression methods. Diana points out that predictability of block time can be affected by ground delays such as weather or congestion, delays propagated by a sequence of flights, and they can be stochastic, such as from crew and equipment problems. He argues that his approach allows predictability to be studied more accurately than previous studies have done. In particular, Diana's objectives are to assess the impacts of selected variables on block delays, derive predictable block times based on these variables, and test his model. Diana uses time series data for the years before and after September 11, 2001 for the Seattle/Tacoma International Airport to control for extreme variability in air travel. After estimating a number of equations, he concluded that quantile regressions can help airlines develop robust schedules.

The fourth paper is by Albitres et al. and its title is "A Dynamic Programming Optimization Approach for Budget Allocation to Early Right-of-Way Acquisitions." The objective of the paper is to identify candidate projects that have gone through preliminary environmental analysis and meet NEPA standards for early acquisition. Thus, "early acquisition of right-of-way is … the purchase of parcels before the approval of … environmental study." The rationale for doing this study is that property values may increase from such factors as change of use and speculation. Early acquisitions thus avoid the increases in costs that these changes could bring, as well as possible delays in project construction from protracted negotiations. The authors accomplish this objective using data for Texas and employing dynamic programming, which breaks down a problem into smaller ones, and are then solved recursively. Dynamic programming also allowed the authors to consider various right-of-way acquisition scenarios. Their study resulted in identification of parcels of land as candidates for early acquisition under different budget scenarios.

In the fifth paper, Lu and Tolliver write on "Pavement Pre- and Post-Treatment Performance Models using LTPP." They determine pavement roughness using the International Roughness Index (IRI) and associate it with exogenous variables, including pavement age, precipitation, freeze-thaw level and other maintenance strategies. To avoid endogeneity, they develop separate exponential models for pre- and post-treatment performance with the latter accounting for the effectiveness of various treatment strategies applied over time. A method of determining a post-treatment performance model using pre-treatment performance models and short-term performance effectiveness was presented. From the pre-treatment models, the authors conclude that minor preservation reduces a pavement's IRI deterioration rate; differences in deterioration in severe weather regions are greater than in less severe weather regions; deterioration increases with freeze-thaw cycles; and deterioration is higher in wet regions than in dry regions.

In "Productivity Improvements in the U.S. Rail Freight Industry, 1980-2010," Carl Martland documents the causes of the rail productivity gains during this period. Martland analyzes three sources of productivity improvement, which are (1) fewer service units per unit of output, (2) fewer resources per service unit, and (3) network rationalization. In connection with (1) he examines changes in traffic mix, length of haul, equipment, tons per load, and trip distances. He analyzes productivity gains from improvements in resource utilization (2) by examining changes in fuel efficiency, freight car utilization, labor productivity, improved track materials, and track maintenance technology. With regard to network rationalization, Martland discusses line abandonment, short line and regional railroads, mergers, and terminal consolidation and transformation. Martland concludes that productivity improvements were greatest for bulk traffic moving in unit trains, containers moving in double-stack trains, and high volume shipments moving long distances in specialized equipment.

The last paper, by Woldeamanuel, is on the stopping behaviors of drivers at stop-controlled intersections. The author's objective is to study how drivers conduct themselves at stop signs by considering their socio-demographic and physical attributes, which influence their stopping behaviors. It collects data for four intersections in St. Cloud, Minnesota, by observing drivers' stopping and other behaviors, including cell phone use as well as information on vehicle occupancy and the presence of a law enforcement officer. The observations were done at 12 different times for each intersection and resulted in 2,400 observations distributed equally among the intersections. Using these data, Woldeamanuel estimates a binary logit model with the dependent variable as the driver behavior of making a complete stop at a controlled intersection. The results from this model indicate that driver stopping behavior can be explained by contextual/ecological variables, including vehicle occupancy, the presence of law enforcement officers, and using headlights. Cell phone use was found to have a statistically insignificant effect on stopping behavior though its negative sign suggests to Woldeamanuel that it could prevent drivers from coming to complete stops at intersections.

Michael W. Babcock Co-General Editor	Kofi Obeng Co-General Editor

Assessment of Sustainable Infrastructure: The Case of Exurban Dallas

by Enid Arvidson, Stephen P. Mattingly, Asapol Sinprasertkool, and Siamak Ardekani

With increasing emphasis on sustainable infrastructure as a setting for sustainable development, two broad types of questions arise: when is a particular infrastructure project sustainable and how is its impact assessed? This paper presents a methodology, tested on two exurban town centers, for assessing the impact of sustainable infrastructure, using various assessment indicators found in existing literature and that fall within the triple bottom line approach (economic, environmental, social). Findings suggest that the method does yield useful information for gauging the impacts of sustainable infrastructure investment, and that the impacts are mostly consistent with the expected and desired outcomes of denser exurban development, increasingly diverse land-use mix, and compact circulation.

INTRODUCTION

An emerging consensus among planners, developers, civil engineers, architects, and other city builders emphasizes the importance of urban development that is "sustainable." Guided by the Brundtland Commission's widely cited definition of sustainable development as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs," this consensus seeks to tackle such "wicked problems" (to borrow Rittel and Webber's (1973) term describing social problems that are difficult to define and whose policy solutions are political) as low-density sprawl and dependency on private automobiles (United Nations General Assembly 1987). Policy responses to these problems include reshaping urban form and travel behavior, involving such efforts as compact, mixed-use, infill, and transit-oriented development[1] (Caves 2005, Wheeler 2004). Yet planning, at least in the United States where public ethos tends to frown upon direct governmental intervention in the market, has little direct control over private investment and development, and little direct power to enact these schemes. Given this "planning paradox," planners tend instead to focus on providing the infrastructure and profitable conditions under which private investment and development can occur (Gottdiener and Hutchison 2011, Dear and Scott 1981). The provision of infrastructure — defined as a non-excludable good and includes roads, railways, urban transport, water supply, sanitation and sewerage, solid waste collection and disposal — thus facilitates and provides the contextual environment in which private decisions and activities of households and businesses take place (The World Bank 1994). The provision of urban infrastructure is therefore a key component in public-sector efforts to provide the setting for private-sector development, and sustainable infrastructure is key in sustainable development efforts.

With the increasing emphasis on sustainable urban infrastructure as a setting for sustainable development, two broad types of questions arise: when is a particular infrastructure project sustainable and how is its impact assessed (Feiden and Hamin 2011)? These questions have led to a substantial and growing literature that propose and test a host of indicators and assessment methods. Yet there is no generally accepted single method or set of indicators, due both to the broadness of the UN General Assembly definition and the wide-ranging use of the term "sustainable development," as well as to the diversity of infrastructure types. Debates and literature on this topic are still evolving. This study contributes to these debates and literature by proposing an assessment method that has been developed and tested in two exurban cities in the Dallas, Texas, metropolitan area.[2] This paper first briefly outlines some of the definitions and assessment methods in the existing literature, next describes the background and approach for this study, and then presents the analysis and findings, concluding with a summary and remarks about strengths and weaknesses and directions for further research.

SELECTED LITERATURE

One increasingly common approach to defining and assessing impacts of sustainable infrastructure is known as the triple bottom line approach. This approach, which started in the private sector but has been quickly adopted in the public sector, focuses not only on economic efficiency but equally on the environmental and social aspects of a project in recognition that development must consider more than simply financial returns if it is to be sustainable in the long term (The World Bank 2002, Caldwell 2011, Institute for Sustainable Infrastructure 2011). There have been few attempts to operationalize the triple bottom line approach in defining and assessing sustainable infrastructure. Shen et al. (2011) operationalize the triple bottom line approach in assessing a large-scale transit infrastructure project by developing what they call Key Assessment Indicators (KAIs) based on surveys of experts and fuzzy set data analysis. They draw from Zadeh's (1965) original work on fuzzy sets as a method for dealing with classes of phenomenon that display no clear set boundaries (as opposed to an ordinary set with clearly defined criteria for set membership), such as experts' opinions of assessment indicators for infrastructure project sustainability. They identify eight key economic indicators, five key social indicators, and seven key environmental indicators, and test these on a case study. They conclude that while the case study produces helpful results, the reliance on expert opinion to identify the indicators may preclude the identification of other important indicators.

Koo et al. (2009) operationalize the triple bottom line approach in assessing underground infrastructure projects by developing what they call the Sustainability Assessment Model (SAM). However, they find that many indicators are difficult to quantify, especially environmental and social ones, and those that they do eventually measure present the problem of incommensurability with other indicators (i.e., the various measures cannot be reduced to a common measure that allows for assessment), suggesting that the model cannot be easily integrated into a final assessment of sustainability.

Other scholars have raised similar criticisms of the triple bottom line approach, arguing that its inherent incommensurability renders the TBL approach mere jargon at best (Robins 2006, Norman and McDonald 2004). The triple bottom line approach has also been criticized for implicitly portraying the three bottom lines as necessarily in competition with one another forcing "trade-offs" in practice, where "weak" sustainability tolerates trade-offs and "strong" sustainability discourages trade-offs (Coffman and Umemoto 2010, Hecht 2007). At the same time, the TBL approach is commended for broadening awareness of more than simply the economic bottom line, and some critics argue that, for "strong" sustainable development, not only must all three bottom lines be at play but also a "holistic mindset" as well as democratic public participation (to secure buy-in from community and stakeholders) (Robins 2006, Coffman and Umemoto 2010, Hecht 2007).

Regarding "trade-offs," skeptics often raise concerns about the potential negative economic impacts of sustainable infrastructure even when it is seen as socially or environmentally beneficial (Deakin 2011). Thus economic impacts, particularly impacts on property values, are frequently a key focal point of impact assessment, especially for cash-strapped local governments that need to consider opportunity costs of expenditures on infrastructure. Rising property values are seen as desirable not only for the local tax base but also as a catalyst for changing land uses and increased density (Alperovich 1983, Gatzlaff and Smith 1993). Many studies use hedonic pricing methods for assessing impacts of infrastructure on property values. For example, Hess and Almeida (2007), who also provide a useful summary of the extensive literature on this topic, assess the impact of transit development on nearby property values, and find that in a declining central city (Buffalo, NY), transit development has weak positive effects on property values, contrary to what studies have shown for growing cities, such as San Francisco. Nicholls and Crompton (2005) assess the impact of a different kind of sustainable infrastructure, namely urban greenbelt development, on nearby property values in a growing city (Austin, TX), and find that greenbelts have a clear positive effect on property values.

Yet, hedonic pricing methods have been criticized for possible multicollinearity as well as lack of depth and breadth of data considered (Nicholls and Crompton 2005). Rather than hedonic pricing methods, Clower et al. (2007) use a mixed measures approach to assess the impact of transit-oriented developments in central Dallas on nearby property values. They combine quantitative and qualitative data, including windshield surveys, interview data, and archival research, in an attempt to capture data that can be overlooked when using more conventional sources. They find that TOD has significant positive impact on property values, although they caution that some impacts may develop slowly given time lags in market responses to infrastructure development.

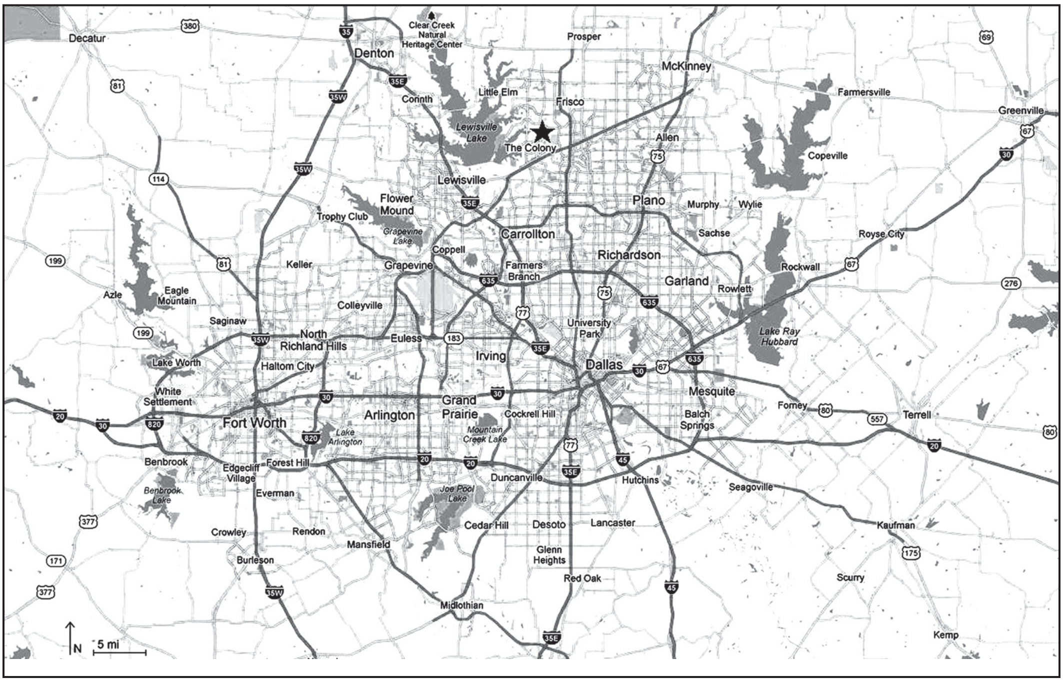
Many of these existing studies of sustainable infrastructure assessment focus on urban areas and central cities, whether declining or growing. This focal point may not be surprising given that compact development requires the higher densities found in urban areas. Yet, smaller edge cities and exurban neighborhoods that are not on transit lines, can and do utilize sustainable infrastructure, for example, street-design measures such as traffic calming, complete streets, context-sensitive design, etc., to foster compact development (Deakin 2011). This study describes and presents the results of a method for assessing the impact of sustainable infrastructure (in this case, street improvements for compact development) in two exurban town centers nearby each other within the Dallas metropolitan area, drawing on the triple bottom line approach. The results show strong positive impact on property values as well as on many of the other, albeit not all, indicators. As with the transit-oriented developments in central Dallas studied by Clower et al. (2007), impacts measured by some indicators in this study here may experience time lags due perhaps to the exurban location or perhaps to other constraints such as the Great Recession.

BACKGROUND AND STUDY APPROACH

With nearly 6.5 million people, the Dallas metropolitan region (also referred to as the Dallas-Fort Worth, or DFW, Metroplex) became, in the mid-2000s, the fourth largest metro region in the United States (behind New York, Los Angeles, and Chicago)—and its regional population is greater than the state-level populations of over 30 states (Mackun and Wilson 2011). Yet, it is far from the fourth densest region: sprawling over not quite 10,000 square miles, its population density is just one-quarter that of New York, Los Angeles, or Chicago, and more typical of recently developing, low-density urban areas in the U.S. sunbelt. Sustainable infrastructure planning takes on particular challenges in low-density sprawling regions. Thus to counteract low-density sprawl, the sustainable infrastructure programs of the North Central Texas Council of Governments (NCTCOG), the regional planning agency for the DFW metropolitan region, do not focus simply on rail and transit but rather "provide for a diverse range of mobility options, such as rail, automobiles, bicycling, transit, and walking" (North Central Texas Council of Governments n.d.).

Through its sustainable infrastructure programs, the NCTCOG provides seed and matching grant monies to cities and towns to foster sustainable development. For exurban town centers[3], the types of sustainable infrastructure that are funded through these grants include, among other things, investment in street construction and improvements to promote denser development, more diverse land use mix, and compact circulation within the center (Mandapaka 2010). This study assesses the impacts, using a number of indicators described below, of sustainable infrastructure investment in street construction and improvements, in two exurban town centers located toward the northern edge of the Dallas metro region (Figure 1). The two town centers, referred to in this study as AR and FS, are both new master-planned, mixed-use communities, including mixes of residential, commercial, office, and recreational uses. AR bills itself as "an up-scale, master planned community" while FR bills itself as "a multi-generational, master planned development… similar to a European village," and both communities describe themselves as, without citing Leinberger's oft-alluded-to description of edge cities from the 1980s, providing "the opportunity to live, work, shop, and play in the same geographic area" (The Colony Convention and Visitors Bureau n.d., Frisco Square Development n.d., Leinberger and Lockwood 1986). Both the AR and FS infrastructure projects involve street construction and improvements, including wider than normal sidewalks on both sides of the street to promote walkability, and wider than normal outside lanes to promote bicycling. Both projects were initiated in 2003 with the aim of promoting compact development within their respective town centers.

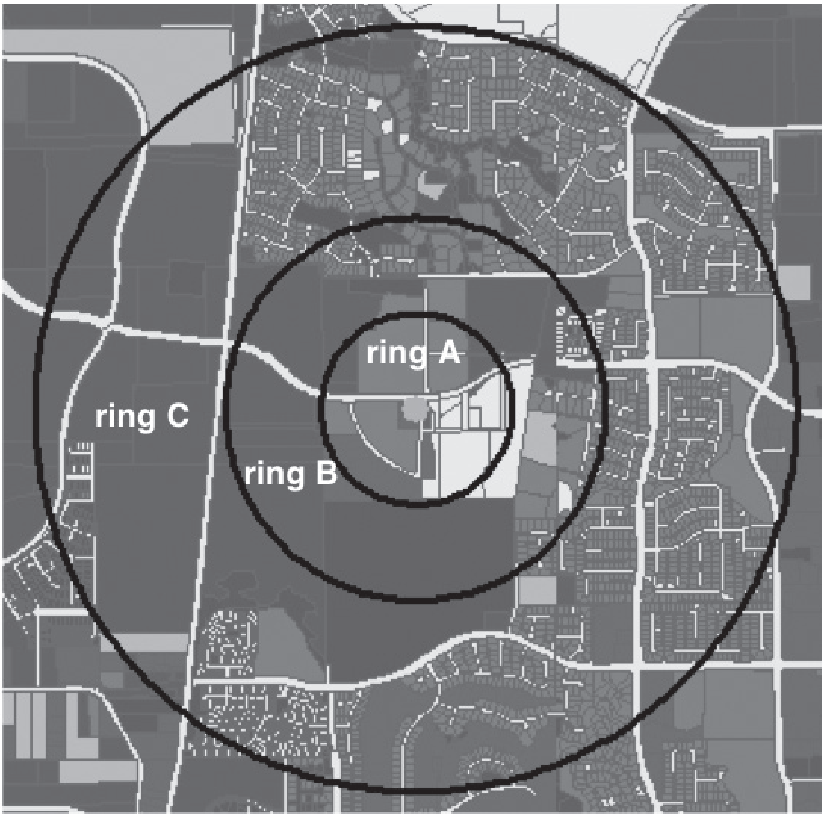
Figure 1: Location of the Study Area Within the Dallas Metropolitan Region



Both AR and FS are home to predominantly non-Hispanic white families with incomes above the region's median. A control site located nearby is also assessed as a reference point to help isolate the possible causes of observed changes at the study sites. The control site has similar historical, demographic, and growth characteristics as the study sites but does not have mixed land uses or publicly funded sustainable infrastructure and instead represents traditional exurban development.

To assess spatial variations in impacts as the distance from the sustainable infrastructure site changes, each of the three study areas (the two sustainable infrastructure sites and the control site) is divided into three concentric rings, A, B, and C, emanating from the center of each site (Figure 2). The cordons are set at one-quarter mile for ring A, one-half mile for ring B, and one mile for ring C, from the center of each study site. A quarter mile is considered the typical walking distance for bus transit, and half a mile is considered the typical walking distance for rail transit (Dittmar and Ohland 2003). The third cordon of one mile is placed to determine if the infrastructure's impact spreads beyond half a mile. Since data for this study are collected from several different geographic units, including census tracts, parcels, and traffic analysis zones (TAZ), and since the boundaries of these geographic units may not lie neatly within a concentric ring, data from these units are considered to belong within a ring if at least half their area falls within that ring (the actual boundary of each cordon is thus not so neat as portrayed in Figure 2). Depending on availability, data are collected from roughly one year before initiation of the sustainable infrastructure project to roughly five years after completion of the project.

Figure 2: Quarter Mile (Ring A), Half Mile (Ring B), and One Mile (Ring C) Cordons from the Center of the Study Site



Drawing on the triple bottom line approach, this study considers economic, environmental and social indicators in assessing the sustainable infrastructure impact. Jeon and Amekudzi (2005) review over a dozen studies on sustainable infrastructure assessment, providing a comprehensive list of measures and indicators used in the studies they review. The indicators selected for use in this study are typical of those used in other studies, albeit they are a small subset of Jeon and Amekudzi's comprehensive list. The specific indicators selected for this study are based on existing literature, along with the research team's knowledge and experience of their relevance for the DFW region, and are constrained by data availability at the geographically disaggregated level of the exurban town center, which is analyzed here. Table 1 lists the indicators used in this study, along with their measurement units normalized per acre (normalizing the data per acre helps address the critique of the TBL approach about incommensurability of data). Table 1 also lists the data source(s) for each indicator, as well as the expected direction of change for each indicator resulting from the impact of the sustainable infrastructure investment.

Table 1: List of Indicators for Impact Analysis and Expected Direction of Change

Indicator	Units	Data Source(s)	Expected Value/Direction of Change for Sustainable Infrastructure Sites Ring A Compared with all Control Site Rings
Economic			
business density	sq. ft/ac	Council of Govts' parcel data (supplemented by google maps and windshield surveys)	higher/increase
employment density*	jobs/ac	Council of Govts' traffic analysis zones (TAZ)	higher/increase
income	median income/ac	U.S. census data	higher/increase
property value*	total value/ac	Council of Govts' appraisal data	higher/increase
sales tax revenue	revenue/ac	State Comptroller of Public Accounts	higher/increase
vacancy rate	sq. ft/ac	real estate companies' absorption studies	lower/decrease
Environment (Built And Physical)			
housing stock*	number of single-family houses/ac	Council of Govts' parcel data (denoted as a binary of single-family vs. other types)	lower/decrease
land use mix (dissimilarity index)*	sq. ft. residential/ ac, sq. ft. commercial/ ac, etc.	Council of Govts' parcel data and symbology function of GIS layer feature (see discussion for details)	higher/increase
sidewalk density*	sq. ft./ac	estimated from GIS maps (see discussion for details)	higher/increase
street density*	lane-mile/ac	roadway data from Council of Govts' GIS database	higher/increase
Social			
average daily traffic	number of trips/day/ac	city traffic studies, Council of Govts	lower/decrease
household density*	number of households/ac	TAZ data from Council of Govts	higher/increase
population density*	number of people/ac	TAZ data from Council of Govts	higher/increase
residential ethnicity	% white/ac, % black/ac, etc…	U.S. census data	same/no change
walking/bicycling trips	number of trips/day/ac	surveys, traffic studies, GPS aps (e.g., iTunes' Cycletracks)	higher/increase

*indicators for which data are readily available and that are considered in this study

Before presenting the results of the analysis, Table 1 is briefly discussed.

Economic Indicators

Business density. Business density is the total number of businesses per acre, and can be obtained from NCTCOG parcel data for each ring, supplemented by Google Maps and windshield surveys. Compared with all rings of the control site, business density is expected to increase the most in ring A of the sustainable infrastructure sites.

Employment density. Employment density is the number of jobs per acre and is calculated using data from the NCTCOG Traffic Analysis Zones (TAZ). Compared with all rings of the control site, employment density is expected to increase the most in ring A of the sustainable infrastructure sites.

Income. Income is average household income and can be obtained from census data. Higher incomes are expected in ring A of the sustainable infrastructure sites compared with rings B and C and with all rings of the control site because the infrastructure and subsequent development improves quality of life, attracting residents who can afford to live there.

Property value. Property value data are obtained from NCTCOG appraisal data. Average property value is expected to increase the most in ring A of the sustainable infrastructure sites compared with rings B and C and also compared with all rings of the control site. If the results of the analysis show that property values increase after the projects are completed, compared with the control site, this is an indication that sustainable infrastructure projects can increase the tax base as well as be a catalyst for compact development (see discussion above and also Alperovich 1983, Gatzlaff and Smith 1993).

Sales tax revenue. Sales tax revenue per acre can be calculated using gross sales data collected in the area, and obtained from The Texas State Comptroller of Public Accounts. Sales tax revenue is expected to increase the most in ring A of the sustainable infrastructure site due to expected increase in business density.

Vacancy rate. Vacancy rate is the percentage of all housing units that are unoccupied or all apartment units that are not rented, and can be obtained from absorption studies by local real estate companies. It is expected that vacancy rates are lower in ring A of the sustainable sites compared with all rings of the control site because sustainable infrastructure development improves quality of life, which attracts prospective residents.

Environmental Indicators (Built Environment and Physical Environment)

Housing stock. Housing stock for each cordon is obtained from NCTCOG parcel data, and is denoted as a binary variable of single-family vs. other types (such as mobile homes, condominiums, townhomes, multi-family, duplex, farm and ranch). It is expected that all rings of the control site would include more single-family homes than the corresponding rings of the sustainable development sites, and that growth in single-family homes would be greater at the control site.

Land Use Mix (Dissimilarity Index). To construct the dissimilarity index, NCTCOG parcel data are used in conjunction with the symbology function of the GIS layer feature, where different colors are assigned to represent each land use category. A grid, with 300 x 300 ft cells, is set on top of the land use layer and is positioned so that the center of the development coincides with the center of the grid. An index is created for each cell by considering the eight adjacent cells. If land use in an adjacent cell is akin to the land use of the cell under investigation, a 0 is assigned; otherwise, a 1 is assigned. These values are then summed and divided by the number of cells. A weighted index is then created by multiplying the index obtained in the previous step by the area of the cell captured within each ring and divided by the cell's overall area (900 ft^2). For each ring, an average index is obtained by summing all the indices for the cells contained within that ring. The interpretation of the index is simply a rule of thumb estimation: if a cell has ≥ 30% heterogeneous use, it is considered a mixed-use cell. Compared with all rings of the control site and rings B and C of the sustainable infrastructure sites, it is expected that land uses will be more mixed in ring A of the sustainable infrastructure sites.

Sidewalk density. Sidewalk density is estimated from GIS maps by dividing the square footage of sidewalk within each ring by each ring's acreage. Square footage of sidewalk is estimated by multiplying the total length of sidewalk by the width. In this study, it is assumed that all sidewalks are four feet wide as called for by city design codes, with the following assumptions as the basis for computing sidewalk length, and then multiplying by the segment length:

- "Primary Highway" (each direction) has one sidewalk along each segment
- "Major Arterial" (each direction) has one sidewalk along each segment
- "Minor Arterial" (both directions) has two sidewalks along each segment
- "Connecting Road" has two sidewalks along each segment
- "Service Road" has one sidewalk along each segment
- "Access Ramp" has one sidewalk along each segment
- "Other" roads have one sidewalk along each segment

Sidewalk density is expected to increase in ring A of the sustainable infrastructure sites as walkability and compact development increases, especially compared with rings B and C of the sustainable infrastructure sites and all rings of the control site.

Street density. Street density is measured by lane-miles, for example, a two-lane street that is one mile long has two lane-miles. Lane-mile data are obtained from roadway data in the NCTCOG's GIS database. Street density is calculated by dividing the total lane-miles within a ring by the ring area. It is expected that street densities will increase in ring A of the sustainable infrastructure sites as a result of the street construction and improvements, compared with rings B and C of the sustainable infrastructure sites and with all control site rings.

Social Indicators

Average daily traffic. Average daily traffic (ADT) can be measured using daily traffic counts in each ring. However, NCTCOG does not have historical data on ADT. Were these data available, we would expect a reduction in the number of trips in ring A of the sustainable infrastructure sites compared with rings B and C and also with all control site rings.

Household density. Household density is the number of households per acre, and is calculated using TAZ household data available from NCTCOG. Household densities are expected to increase in ring A of the sustainable infrastructure sites compared with rings B and C and with all rings of the control site.

Population density. Population density is calculated using TAZ data from NCTCOG. Population density is expected to increase in ring A of the sustainable infrastructure sites compared with rings B and C and with all rings of the control site.

Residential ethnicity. Residential ethnicity is the percentage of residents in different ethnic groups within each ring and can be obtained from census data. In exurban centers such as those analyzed in this study, residential ethnic diversity is expected to remain the same in all rings of the sustainable development sites. While the opposite trend (namely, sustainable development leading to increased ethnic diversity) would be desirable and in fact would be expected in central city sustainable infrastructure projects due to gentrification, it is not expected in homogenous exurbs (Gottdiener and Hutchison 2011).

Walking/bicycling trips. Walking/bicycle trip data can be obtained through surveys, manual or automated counts (taken by field data collectors or specialized equipment), or GPS applications (such as iTunes' CycleTracks). It is expected that non-motorized modes of transportation will increase in ring A of the sustainable infrastructure sites compared with rings B and C of the sustainable infrastructure sites and with all rings of the control site.

ANALYSIS AND FINDINGS

Using the measures and methods outlined above, this section discusses findings for those indicators for which data are readily available. Given the scope and time frame of this study, data are not readily available for all indicators, thus a subset is analyzed and compared among the sites and rings.

Economic Indicators

Employment density. Table 2 summarizes findings for employment density. Compared with the control site, employment density is expected to increase in ring A of the sustainable infrastructure sites. Yet contrary to expectations, the results show strong increase in employment density in the control site, as well as in ring B of one of the sustainable development sites (AR)—albeit these strong percentage changes reflect very minor absolute growth. On the other hand, the strong growth in employment density in AR's ring A appears to indicate, consistent with expectations, that a commercial center has been created subsequent to the sustainable infrastructure investment.

Table 2: Employment Density (Jobs/Acre)

Year	AR Rings			FS Rings			Control Site Rings		
	A	B	C	A	B	C	A	B	C
2010	4.8	2.4	5.9	0.9	1.2	0.8	1	1.5	1.2
2005	2.7	1.4	4.9	0.7	1.2	0.8	0.4	1	0.9
2000	0	0.1	3.5	0.6	1.1	0.7	0.1	0.6	0.6
annual % change 2000–2010	16%*	230%	7%	5%	1%	1%	90%	15%	10%

*annual % change 2005-2010

Property value. Table 3 shows findings for average property values. Consistent with expectations, there was a strong increase in average property values in ring A of the sustainable development sites, AR and FS, particularly compared with the control site. This increase could be a clear indication that sustainable development infrastructure can be a catalyst for development and an increased tax base.

Table 3: Average Property Values (000s Dollars/Acre)

Year	AR Rings			FS Rings			Control Site Rings		
	A	B	C	A	B	C	A	B	C
2007	277	199.1	389.6	678.4	320.7	167.5	454.8	475.1	357.4
2006	239.7	129.2	282.6	345.8	246.2	98.3	455.3	422.7	350.4
2005	209.9	105.6	273.6	341.4	260.6	111.5	384	356.8	318
2004	142.1	106.4	247.8	336.2	164.7	164.7	342.5	347	284.6
2002	38.9	120	265.9	88.8	147.6	67.3	235.1	164.8	212.3
annual % change 2002-2007	122%	13%	9%	133%	23%	30%	19%	38%	14%

Environmental Indicators

Housing stock. Table 4 presents findings for changes in single-family housing stock. For each ring, the single-family housing stock at the control site is greater than at the two sustainable development sites, except for 2002, which is before the sustainable infrastructure investment in 2003. This finding is consistent with expectations.

Table 4: Single-Family Housing Stock (Number Of Single-Family Houses/Acre)

Year	AR Rings			FS Rings			Control Site Rings		
	A	B	C	A	B	C	A	B	C
2007	0	0.21	1.14	0.06	0.18	0.27	0.49	0.7	1.41
2004	0	0.16	1.06	0.02	0.14	0.27	0.43	0.64	1.06
2002	0	0.16	1.05	0.02	0.14	0.27	0.41	0.61	0.72
annual % change 2002-2007	-	6%	2%	40%	6%	0%	4%	3%	19%

Land use mix (dissimilarity index). Table 5 summarizes the land-use mix of the two sustainable development sites compared with the control site. Consistent with expectations, prior to the infrastructure investment in 2003, ring A of both sustainable development sites had little (AR) to some (FS) mixed land uses, yet after the investment both rings have significant mixed land use. The increasing diversity of land use in ring A of the sustainable development sites is all the more pronounced compared with lack of changes in land-use mix in rings B and C and in all rings at the control site, suggesting that the sustainable infrastructure investment has achieved one of its key goals.

Table 5: Land Use Mix/Dissimilarity Index (≥ 30% ≡ Mixed Use)

Year	AR Rings			FS Rings			Control Site Rings		
	A	B	C	A	B	C	A	B	C
2007	0.316	0.01	0.086	0.7	0.481	0.298	0.161	0.064	0.104
2006	0.316	0.01	0.087	0.449	0.481	0.278	0.205	0.073	0.037
2005	0.316	0.01	0.085	0.449	0.482	0.294	0.161	0.073	0.038
2004	0	0	0.052	0.365	0.431	0.35	0.137	0.064	0.019
2002	-	0	0.067	0.415	0.413	0.33	0.156	0	0

Sidewalk density. Table 6 presents the findings for sidewalk density. Consistent with expectations, sidewalk density increased in ring A compared with rings B and C of the sustainable infrastructure sites and compared with all rings of the control site.

Table 6: Sidewalk Density (Sq. Ft./Acre)

Year	AR Rings			FS Rings			Control Site Rings		
	A	B	C	A	B	C	A	B	C
2007	1,989	394	1,066	1,369	2,664	1,766	910	1,029	1,333
2004	1,640	394	984	887	1,658	1,206	840	1,029	1,333
2002	0	372	984	528	1,546	1,176	840	869	1,261
annual % change 2002-2007	7%*	1%	2%	32%	14%	10%	2%	4%	1%

*annual % change 2004-2007

Street density. Table 7 presents findings for street density. Consistent with expectations, ring A of the sustainable infrastructure sites shows a significant increase in street density compared with rings B and C and with all rings of the control site. This increase in ring A suggests that the sustainable infrastructure investments may have contributed to denser development within the two town centers, which is one of the key goals of the sustainable infrastructure investment.

Table 7: Street Density (Lane Miles/Acre)

Year	AR Rings			FS Rings			Control Site Rings		
	A	B	C	A	B	C	A	B	C
2007	0.122	0.019	0.066	0.172	0.252	0.167	0.043	0.049	0.068
2004	0.078	0.019	0.062	0.084	0.157	0.114	0.04	0.049	0.068
2002	0	0.018	0.062	0.05	0.146	0.111	0.04	0.041	0.064
annual % change 2002-2007	19%*	6%	6%	71%	42%	33%	8%	16%	5%

*annual % change 2004-2007

Social Indicators

Household density. Table 8 shows findings for household density. The greatest increases in household density are in ring A of all sites (i.e., the control site and both of the sustainable infrastructure sites), as well as ring B of one sustainable infrastructure site (AR). Yet these notable percentage increases reflect relatively small changes in absolute numbers in all rings except one of the sustainable infrastructure sites (AR), particularly in ring A of this site. Thus, these data, in terms of percent change, appear to somewhat support the expected finding that household densities will increase in ring A of the sustainable infrastructure sites compared with rings B and C and with all rings of the control site. Yet when considering the absolute numbers, the most dramatic increases in household density occur in all rings of AR but not so much in the rings of FS or the control site, suggesting that perhaps something else is going on in AR, in addition to the sustainable infrastructure investment (such as continued build-out of the master planned community), to cause this increase in household density.

Table 8: Household Density (Number of Households/Acre)

Year	AR Rings			FS Rings			Control Site Rings		
	A	B	C	A	B	C	A	B	C
2010	10.8	5.3	4.2	0.3	1	1.1	1.9	1.4	1
2005	5	2.8	3.7	0.3	0.9	1	1.3	1.1	0.7
2000	0	0.4	2.2	0.005	0.2	0.5	0.2	0.5	0.4
annual % change 2000-2010	23%*	123%	9%	590%	40%	12%	85%	18%	15%

*annual % change 2005-2010

Population density. Table 9 presents findings for population density. Consistent with expectations, the sustainable infrastructure sites (particularly one of them — AR) have higher population densities, and greater percent increases in population densities, than the control site. However, contrary to expectations, the FS numbers are unexpected, and the higher densities, in terms of absolute numbers, are in the outer rings rather than ring A of FS. In addition, contrary to expectations, the control site had significant percentage increases in population density in ring A.

Table 9: Population Density (Number of People/Acre)

Year	AR Rings			FS Rings			Control Site Rings		
	A	B	C	A	B	C	A	B	C
2010	27.4	13.6	11.9	1	2.8	3.1	3.7	2.9	2.2
2005	12.7	7.3	10.8	1	2.6	2.8	2.5	2.1	1.6
2000	0	1.2	6.5	0.2	0.8	1.4	0.4	1.1	1
annual % change 2000-2010	23%*	103%	8%	40%	25%	12%	83%	16%	12%

*annual % change 2004-2007

Summary of Findings

Table 10 summarizes these findings, showing whether the findings for each indicator confirm, or compromise, the expected values and direction of change of the sustainable infrastructure sites' ring A compared with all control site rings. Several findings warrant highlighting. First, according to most of the *economic* and *environmental* indicators analyzed in this study (two of the three factors in the TBL approach), the sustainable infrastructure investments had the desired impacts. Specifically, the sustainable infrastructure sites show clear increases in property values, land use mixes, sidewalk and street densities, and relatively less growth in single-family housing stock, compared with the control site. These changes at the sustainable infrastructure sites compared with the control site suggest that the sustainable infrastructure investment (*viz*., investment in various types of street improvements) provided the hoped-for catalyst for denser development, increasingly diverse land-use mix and compact circulation within the town centers. Thus, sustainable infrastructure programs targeted at Sunbelt-style exurbs need not necessarily emphasize traditional strategies of rail and transit, more appropriate for older or central cities, to produce the desired outcome of fostering sustainable development.

Table 10: List of Indicators Comparing Expected With Actual Direction of Change

Indicator	Expected Value/Direction of Change for Sustainable Infrastructure Sites Ring A Compared With All Control Site Rings	Actual Direction of Change for Sustainable Infrastructure Sites Ring A Compared With All Control Site Rings
Economic		
employment density	higher/increase	ambiguous: higher for AR but not FS/increase greatest at Control Site and Ring B of AR
property value	higher/increase	strongly confirmed
Environment (Built And Physical)		
housing stock	lower/decrease	mostly confirmed
land use mix (dissimilarity index)	higher/increase	strongly confirmed
sidewalk density	higher/increase	confirmed
street density	higher/increase	strongly confirmed
Social		
household density	higher/increase	ambiguous: higher for AR but not FS/increases at all sites various rings
population density	higher/increase	ambiguous: higher for AR but not FS/increases at all sites various rings

Second, the positive effect on property values confirms findings of previous studies, reinforcing the notion that local governments need not worry about trade-offs between environmental vs. economic impacts. Rather, at least in this case, positive economic impacts have accompanied positive environmental impacts, leading to a virtuous cycle of development. Third, most of the hoped-for impacts occurred in rings A and B of the sustainable infrastructure sites, suggesting that the infrastructure's impact does not spread beyond half a mile from the center. Since half a mile is considered a comfortable walking distance to a destination, this finding reinforces the notion that the sustainable infrastructure is promoting walkability and compact development at the town center. Fourth, both of the *social* indicators analyzed in this study (the third leg of the TBL approach) show mixed results. Specifically, the data for household and population densities do not show the expected clear increase in density at the sustainable infrastructure sites, especially compared with the control site. Perhaps if additional social indicators were used, such as travel habits (i.e., average daily traffic counts, or walking/bicycle trips), then clearer impacts on social choices could be assessed. Fifth, the ambiguous social impacts might also be rendered more certain by directly involving citizens in the planning process. The TBL approach has been criticized for privileging the economic and environmental over the social, particularly democratic public participation. Increased citizen participation, through, for example, community meetings, focus groups, and participatory media, could increase public input and support for sustainable infrastructure and development, contributing to changed social choices and behavior that reinforce the economic and environmental impacts.

CONCLUSION AND RECOMMENDATIONS

This paper presents a methodology, tested on two exurban town centers, for assessing the impact of sustainable infrastructure, using various assessment indicators commonly found in existing literature and that fall within the triple bottom line categories (economic, environmental, social). To better assess spatial variation in impacts, the study areas are cordoned into concentric rings, a control site is selected as point of comparison, and the data are normalized per acre to address the problem raised in previous studies about incommensurability of economic, environmental, and social data. Findings suggest that the method used in this study does yield useful information for gauging the impacts of sustainable infrastructure investment, and that the impacts are mostly consistent with the expected and desired outcomes of denser development, increasingly diverse land-use mix and compact circulation, within the town centers. Limitations of this study include the limited number of indicators for which data are available at the geographically disaggregated level, and the limited observations (years) for this data. Future applications of the methodology presented here ideally would utilize a wider variety of indicators over more observations (spanning more years), particularly social indicators such as direct citizen input. Despite the limitations of the existing study, the methodology presented here can be a useful contribution to the literature on sustainable infrastructure assessment, in a field where there is no generally accepted single method or set of indicators to assess sustainable infrastructure. The test application of this methodology on two exurban town centers suggests that sustainable infrastructure programs in low-density exurbs, even if not focused on traditional strategies of rail and transit, can be effective in promoting mixed land use and compact development.

Acknowledgements

The authors wish to thank the North Central Texas Council of Governments for funding research related to this paper, and Michael Babcock as well as the anonymous reviewers for their helpful comments on earlier drafts of this paper.

Endnotes

1. These terms are often found in policies that address sprawl and dependency on private automobiles, and generally can be defined or described as follows. Compact development is typically described as contiguous high-density development that enables walking or public transportation (Caves 2005). Mixed-use development is contradistinguished from single-use,

separated, non-diverse land uses promoted by traditional zoning of the early-to-mid twentieth century (Wheeler 2004). Infill development can be described as the use or re-use of vacant or underused parcels or buildings within the central urban area (Caves 2005). Transit-oriented development is development in high-density, mixed-use areas where various forms of mass-transit (such as bus or rail) serve as focal point (Caves 2005). All these forms of development are seen as complementary ways of addressing unsustainable development.

2. Exurbia can be defined, following Caves (2005), as the low-density development beyond a metropolitan area's suburbs but within its commuting shed. Its development is facilitated by the shift to a service and information economy and the accompanying spatial decentralization of economic activity. This development is further facilitated by the extension of transportation systems on the urban fringe, pursuit of the "American Dream" lifestyle, and White flight (defined as spatial segregation resulting from the massive population shift to the suburbs and exurbs along racial and class lines [see Caves 2005, and Gottdiener and Hutchison 2011]).

3. According to Caves (2005), exurbia is predominantly residential but is serviced by small urban centers, often with locally owned small businesses, resulting in a diverse mix of land uses and a small-town feel.

References

Alperovich, G. "An Empirical Study of Population Density Gradients and Their Determinants." *Journal of Regional Science* 23 (4), (1983): 529–540.

Caldwell, K. "'Envisioning' the Future." *Civil Engineering* May, (2011): 10.

Caves, R. (ed). *Encyclopedia of the City*. Routledge, New York, 2005.

Clower, T.L., B. Weinstein, and M. Seman. *Assessment of the Potential Fiscal Impacts of Existing and Proposed Transit-Oriented Development in the Dallas Area Rapid Transit Service Area*. Center for Economic Development and Research, University of North Texas, Denton, TX, 2007.

Coffman, M. and K. Umemoto. "The Triple-Bottom-Line: Framing of Trade-Offs in Sustainability Planning Practice." *Environment, Development and Sustainability* 12 (5), (2010): 597-610.

Deakin, E. "Climate Change and Sustainable Transportation: The Case of California." *Journal of Transportation Engineering* 137 (6), (2011): 372-382.

Dear, M. and A. Scott. *Urbanization and Urban Planning in Capitalist Society*. Methuen, London, 1981.

Dittmar, H. and G. Ohland. *The New Transit Town: Best Practices in Transit-Oriented Development*. Island Press, Washington, D.C., 2003.

Feiden, W. and E. Hamin. *Assessing Sustainability: A Guide for Local Governments*. APA Press, Chicago, 2011.

Frisco Square Development. Frisco Square Overview. n.d. http://www.friscosquare.com/projectoverview

Gatzlaff, D. and M. Smith. "The Impact of the Miami Metrorail on the Value of Residences Near Station Locations." *Land Economics* 69 (1), (1983): 54-66.

Gottdiener, M. and R. Hutchison. *The New Urban Sociology, 4th ed*. Westview, Boulder, CO, 2011.

Hecht, J. *Can Indicators and Accounts Really Measure Sustainability? Considerations for the U.S. Environmental Protection Agency*. U.S. Environmental Protection Agency, Washington, D.C., 2007, http://www.epa.gov/sustainability/pdfs/hecht-epa-ord-paper.pdf

Hess, D.B. and T.M. Almeida. "Impact of Proximity to Light Rail Rapid Transit on Station-Area Property Values in Buffalo, New York." *Urban Studies* 44 (5/6), (2007): 1041-1068.

Institute for Sustainable Infrastructure (ISI). *New Organization to Develop Tool for Measuring Infrastructure Sustainability*. ISI Website: Washington, D.C., 2011, http://sustainableinfrastructure.org/news/pr020111.cfm

Jeon, C.M. and A. Amekudzi. "Addressing Sustainability in Transportation Systems: Definitions, Indicators, and Metrics." *Journal of Infrastructure Systems* 11 (1), (2005): 31-50.

Koo, D-H., S.T. Ariaratnam, and E. Kavazanjian Jr. "Development of a Sustainability Assessment Model for Underground Infrastructure Projects." *Canadian Journal of Civil Engineering* 36 (5), (2009): 765–776.

Leinberger, C. and C. Lockwood. "How Business Is Reshaping America." *Atlantic Monthly* October, (1986): 43-52.

Mackun, P. and S. Wilson. *Population Distribution and Change: 2000 to 2010 — 2010 Census Briefs* (pub. no. C2010BR-01). U.S. Census Bureau, Washington, D.C., 2011.

Mandapaka, P. Personal Communication, April 21, 2010, *Transportation Planner*, NCTCOG, Arlington, TX.

Nicholls, S. and J.L. Crompton. "The Impact of Greenways on Property Values: Evidence from Austin, Texas." *Journal of Leisure Research* 37 (3), (2005): 321-341.

Norman, W. and C. MacDonald. "Getting to the Bottom of 'Triple Bottom Line'." *Business Ethics Quarterly* 14 (2), (2004): 243-262.

North Central Texas Council of Governments (NCTCOG). *Sustainable Development: Land Use/Transportation Planning Website*, Arlington TX, n.d. http://www.nctcog.org/trans/sustdev/landuse/index.asp

Rittel, H. and M. Webber. "Dilemmas in a General Theory of Planning." *Policy Sciences* 4, (1973): 155-169.

Robins, F. "The Challenge of TBL: A Responsibility to Whom?" *Business and Society Review* 111 (1), (2006): 1–14.

Shen, L., Y. Wu, and X. Zhang. "Key Assessment Indicators for the Sustainability of Infrastructure Projects." *Journal of Construction Engineering and Management* 137 (6), (2011): 441-451.

The Colony Convention and Visitors Bureau. The Colony Attractions: Austin Ranch. n.d. http://visitthecolonytx.org/attractions_artsandculture.php

The World Bank. *World Development Report 1994: Infrastructure for Development*. World Bank and Oxford University Press, Oxford, 1994.

The World Bank. *World Development Report 2003: Sustainable Development in a Dynamic World: Transforming Institutions, Growth, and Quality of Life*. World Bank and Oxford University Press, Oxford, 2002.

United Nations General Assembly. *Report of the World Commission on Environment and Development*. December 11, 1987, http://www.un.org/documents/ga/res/42/ares42-187.htm

Wheeler, S.M. *Planning for Sustainability: Creating Livable, Equitable, and Ecological Communities*. Routledge, New York, 2004.

Zadeh, L.A. "Fuzzy Sets." *Information and Control* 8 (3), (1965): 338-353.

Enid Arvidson *is associate professor in the School of Urban and Public Affairs at the University of Texas at Arlington. She holds a doctorate in economics and has published primarily in the areas of critical urban theory, urban political economy, and the intersection of economic/political/social processes in the development and restructuring of capitalist cities. Her work has appeared among other places in* Economic Geography, Rethinking Marxism, *and* Publius*, as well as in book chapters of peer-reviewed volumes. She currently serves on the Editorial Board of the Review of* Radical Political Economics*, and in the past has served on the Governing Board of the Association of Collegiate Schools of Planning, the Editorial Board of* Rethinking Marxism*, and as guest co-editor of a special issue on cities and government policy of the* Journal of Urban Affairs.

Stephen Mattingly *is associate professor in civil and environmental engineering at the University of Texas at Arlington. He recently completed and is currently working on research projects that address a variety of topics including airport operations, managed lane traveler behavior, high-speed rail compatibility with existing freeway right-of-way, teaching/educational strategies, transportation emissions and sustainability, and freeway and frontage road operational modeling and performance measures. Previously, he worked at the Institute of Transportation Studies, University of California, Irvine, and University of Alaska, Fairbanks. He has published several articles in* Transportation Research Record*, as well as other journals and conferences on these and other related topics. He currently serves on the Transportation Research Board (TRB) Committee on Traffic Flow and Characteristics and is past member of the TRB Committee on Transportation Network Modeling. Stephen is also member of the Transportation Research Forum, Institute of Transportation Engineers, American Society of Civil Engineering and American Society for Engineering Education.*

Asapol Sinprasertkool *is research associate of transportation engineering at the University of Texas at Arlington. He received his doctoral degree in transportation engineering from the University of Texas at Arlington. During his Ph.D. program, he was employed as research assistant on several transportation research projects including the IH-30 managed lane toll revenue study funded by Texas Department of Transportation and the sustainable development impact study funded by North Central Texas Council of Governments. He was also a graduate teaching assistant for various transportation engineering classes, including analytical models in transportation, TransCAD, and surveying.*

Siamak Ardekani *is professor of civil engineering at the University of Texas at Arlington. He has served on a number of national committees, including the TRB Committees on Traffic Flow Theory and Characteristics and Alternative Transportation Fuels, the International Committee on Engineering Design and Automation, the ASCE Subcommittee on Human-Powered Transportation, and the ASCE Technical Council on Lifeline Earthquake Engineering. He is also a charter member*

of the ASCE Transportation and Development Institute and has served as a member of the Board of ITS Texas. He is also member of the Institute for Operations Research and Management Science and the American Society of Engineering Education, and has served as associate editor for the Transportation Science Journal *as well as the* Journal of Transportation Research-A. *He is a registered professional engineer in Texas and Virginia.*

Applying the Highway Safety Manual to Two-Lane Road Curves

by Daniel Findley, Charles Zegeer, Carl Sundstrom, Joseph Hummer, and William Rasdorf

This paper evaluates the Highway Safety Manual (HSM) crash prediction model using data on two-lane rural horizontal curves in North Carolina. An analysis of the local conditions calibration factor for the HSM predictive model in North Carolina found that a large number of sites (approximately 300) are required to meet HSM recommendations. The results showed that annual average daily traffic, curve radius, and curve length were the most important factors in determining crash prediction accuracy, but that average or default values may be used for other parameters with less risk to accuracy.

INTRODUCTION

Horizontal curves are relatively risky portions of the highway system in the United States and elsewhere. Collisions on two-lane curves have been found to be more than twice as likely to result in a fatality as all two-lane roadway segments (Hummer et al. 2010). Fortunately, curves are also places where highway agencies have many options and opportunities for making safety improvements. Agencies can add signs, markings, beacons, guardrails, and/or superelevation (cross-slope of the roadway), or can widen, straighten, and flatten sideslopes, just to name some common and proven examples of potential improvements. This paper focuses on the analysis of two-lane rural horizontal curves.

Typically, the analysis of a horizontal curve or a set of curves for safety purposes by a highway agency is based on field visits and the judgments of experienced personnel. Many agencies seem to rely on a drive-through by an engineer or a technician and a small set of countermeasures that seem to have proven themselves through the years. Analytical tools have existed for a number of years, such as the 1991 FHWA curve crash prediction model (Zegeer et al. 1991). That study developed a model to predict the number of curve crashes based on such geometric factors as degree of curve (amount of curvature of an arc), length of curve, roadway width, roadside hazard rating, superelevation on the curve, and presence or absence of spiral transitions to the curve (i.e., a smooth transition from the straight tangent to the curve) as well as traffic volume (ADT). While this model predicted curve crashes well for these mostly geometric variables, it did not incorporate the effects of traffic control devices such as signs, markings, flashing lights, rumble strips, lighting, and other variables. Also, the model that was developed by Zegeer et al. (1991) was based exclusively on data from a single state (Washington), even though it is recognized that there are state-to-state differences in crash reporting thresholds, climate, terrain, driver characteristics, and other factors that can affect crash frequencies and rates for a given set of roadway conditions.

Therefore, there has been a need for a model or tool to allow agencies to predict the crash potential of horizontal curves within a state or jurisdiction based on a wide variety of geometric, traffic, and other site-specific features, and for that model to be validated for the state where it is to be applied. There have also been barriers to widespread implementation of past curve crash models and tools due to the large number of competing highway safety objectives, real or perceived difficulties in collecting the necessary data, and possibly the need to calibrate the model for local conditions, among other reasons.

The publication of the Highway Safety Manual (HSM) offers a chance to overcome this impasse and get a crash model in use in the field (AASHTO 2010). The HSM contains a crash prediction model for horizontal curves and estimates of crash modification factors (CMFs) for popular curve countermeasures. A CMF is defined as the expected change in crashes that results from a given safety treatment. For example, if a countermeasure is expected to reduce crashes by 20%, the CMF is expressed as 1.00 - 0.20 = 0.80. The model and CMFs have been approved by a committee of leading safety researchers and practitioners, and this provides credibility to the tools it provides. The HSM also contains detailed instructions for applying the model and CMFs to the usual steps in a safety program, including evaluating installed countermeasures. Despite the promise of the HSM and the fact that draft versions circulated widely for several years before publication, it is yet to be widely applied in curve safety studies. Perceived or real difficulties in calibrating the HSM models and collecting needed data may be contributing to this slower-than-expected adoption process.

The objectives of this paper are to provide highway agencies with practical advice on how to use the new HSM to analyze horizontal curves to supplement the usual methods that identify curves with abnormally high crash experience. This paper answers the following questions: Can agencies use the new HSM to identify and analyze horizontal curves in need of safety improvements? If so, how should an agency calibrate the HSM curve crash prediction model to fit local conditions? If an HSM analysis is possible, how much effort should the agency expect to make? Is an HSM analysis possible without a field visit? If so, what accuracy can be expected? What steps should agencies take to make the HSM analysis more efficient so that they can utilize the results and apply them to improve curves in a more cost-effective manner? Satisfying these objectives should shorten the learning curve for agencies in using the HSM curve crash prediction procedure and reduce the risk agencies and professionals assume in using this new tool.

LITERATURE REVIEW

Due to the recent release of the HSM, only a few studies have been completed on calibrating its crash prediction models. This literature review consists of relevant studies that have evaluated the HSM application for two-lane and rural roads, the variance in crash modeling, and the calibration of HSM models. Sun et al. (2006) evaluated the applicability of the HSM safety prediction model to states from which crash data were not used in the original model development. The prediction model evaluated in this study was that for two-lane rural roads in the draft HSM. Data from state routes in Louisiana were used. Due to data limitations, the authors did not follow the recommended HSM procedure for calibrating the predictive model. However, the research team was able to create a database with important highway variables, including average daily traffic (ADT), segment length, lane width, shoulder width and type, and driveway density. Since the average predicted values were smaller than the observed values, a calibration parameter was calculated as a function of ADT. The results of this analysis were presented for two sets of road sections: the first consisted of 26 randomly selected sections, and the second, 16 sections ranked in the top 30 in the state for crash frequencies over three years. The analysis indicated that the HSM model successfully predicted crash frequencies, but the level of effort required to obtain the data necessary to calibrate the model was a challenge.

Martinelli et al. (2009) calibrated the HSM crash prediction model for the Italian Provence of Arezzo using 1,300 kilometers of rural two-lane highways. A comparison of observed crashes and results from four models with different calibration procedures showed they strongly overestimated crashes. Additionally, it was found that the models overestimated crashes at low crash locations and underestimated crashes at high crash locations. The authors concluded that calibration of the model is absolutely necessary to avoid over prediction in the base model. They also note that a primary issue with calibration exists because the high segmentation of the HSM procedures leads to low or zero crash segments, which are not predicted accurately by the HSM.

The accuracy of models using baseline data is also of interest for this paper. A recent study by Lord et al. (2010) compared crash prediction models for rural four-lane highways in Texas. Two full models with several covariates and the product of baseline models and accident modification factors (AMFs) were compared using predicted mean values and variances. The results of this analysis showed that the full models have much smaller variances than the product of baseline models and AMFs. This finding led the authors to conclude that when a study's objective includes variance as part of the decision-making process, a full model should be used.

Further details on which elements are critical to the outcome of a crash prediction model are also of interest in determining which elements will have the least effect if they remain as default settings. A study by Nowakowska (2010) developed logistic models for crash severity based on road characteristics of rural highways in Poland. This study found that shoulder presence and type, area type, sidewalk presence, and interactions had a statistically significant influence on crash severity. Easa et al. (2009) evaluated crash prediction models for three-dimensional alignments of rural two-lane highways in Washington State. They found that the most significant predictors of crashes were degree of curvature, roadway width, access density, grades, section length, and average annual daily traffic (AADT).

Xie et al. (2011) applied the HSM procedures to roadway segments in Oregon for the purpose of calibrating the model for local conditions. They included randomly selected roadway segments and found a two-lane roadway calibration factor of 0.74 across 75 sites with 394 reported collisions and 533 HSM predicted collisions. The authors presented a methodology for sites that did not meet the recommended 100 collisions per year among 30 to 50 locations. To overcome the under-represented collision locations, the authors applied sample size estimation procedures based on average Oregon crash history for that type of site to modify the expected total yearly collisions. Another study also examined the calibration of the HSM as well as the development of new models (Banihashemi 2011) and found that a calibrated HSM model performs as well as the newly developed models, and it is the preferred safety model. Banihashemi (2011) also predicted a total of 150 collisions per year for the sites employed in the calibration process.

METHODOLOGY

Data Collection

The collection of different data needed for calibrating one HSM model is described below for each of the selected curve sample sites in North Carolina. The selected samples of curves were all on two-lane rural roads. The researchers asked NCDOT to select 50 curve sites, with no more than five of the curves on any given roadway for the calibration effort. Field investigation forms developed by the researchers were distributed to NCDOT personnel who were assigned the task of collecting the necessary data on 21 variables for each curve. The procedure for measuring curve radius and the superelevation of the curve is described in Findley and Foyle (2009). In this validation effort, each selected curve must be isolated from other curves by tangent segments on both ends. Then relevant variables were collected for each curve along with similar data for the adjoining tangent sections on both ends of the curve. Table 1 lists the 21 variables on which data were collected for this study and provides a brief description of the data collection process for some variables.

HSM Predictive Method Calibration

The HSM predictive method is used to estimate crash frequency, severity, and types of crashes on a highway with known characteristics. To improve the accuracy of the model, the HSM predictive methods were developed such that they can be calibrated and adjusted based on local conditions.

Table 1: Field Data Collection Elements

Feature		Value
1.	Posted Speed Limit (mph):	
2.	Lane Width (feet): *(Measure from center of the lane-line of the roadway to center of edgeline, round to the nearest foot)*	
3.	Inside Shoulder Width (feet): *(Measure from center of edgeline to edge of shoulder, round to the nearest foot)*	
4.	Inside Shoulder Type: *(Paved, Gravel, Turf, or Composite)*	
5.	Outside Shoulder Width (feet): *(Measure from center of edgeline to edge of shoulder, round to the nearest foot)*	
6.	Outside Shoulder Type: *(Paved, Gravel, Turf, or Composite)*	
7.	Length of Section (feet): *(Measure from beginning of the curve to the end of the curve along the edgeline, in feet, measure tangents from end of curve to within 100' of the nearest intersection or next curve)*	
8.	Radius of Horizontal Curve (feet): *(Determine the radius using the attached Field Investigation Procedure and completed Field Investigation Form below)*	
9.	Roadside Hazard Rating (1-7): *(See the attached photos for examples)*	
10.	Inside Lane Superelevation (%): *(Determine the superelevation using the attached Field Investigation Procedure and completed Field Investigation Form below)*	
11.	Outside Lane Superelevation (%): *(Determine superelevation using the attached Field Investigation Procedure and completed Field Investigation Form below)*	
12.	Grade (%): *(Determine the grade using the digital level to find the steepest grade)*	
13.	Number of Driveways: *(Record the total number of driveways along the length of the roadway from beginning to end of segment on both sides)*	
14.	Presence of Raised Pavement Markers (Yes/No):	
15.	Presence of Passing Lanes* (Yes/No):	
16.	Presence of Roadway Lighting* (Yes/No):	
17.	Presence of Centerline Rumble Strips* (Yes/No):	
18.	Presence of Two-Way Left-Turn Lanes* (Yes/No):	
19.	Presence of Shoulder Rumble Strips (Yes/No):	
20.	Presence of Skid Treatments (overlay) (Yes/No):	
21.	Presence of Skid Treatments (groove pavement) (Yes/No):	

Examples of local conditions that may differ from the given predictive model include climate, geographic conditions, driver characteristics, and crash reporting thresholds.

The HSM predictive method for rural two-lane, two-way highways was applied in this evaluation to North Carolina highways. Other roadway types are available for analysis within the HSM through similar, but different methods which are specific to the characteristics that influence safety on those roadways. This application followed the steps provided in the HSM to estimate the expected average crash frequency of curve segments. The HSM predictive model contains 18 steps starting with defining the segment and period of study to evaluating the results. The focus of this paper is on Step Nine, which selects and applies safety performance functions (SPF); Step Ten, which applies CMFs to the segments; and Step 11, which involves applying a local calibration factor. These steps are applied after the roadway segments have been identified and the data collection, including crash history and geometric conditions, is complete. Each step must be completed separately for all identified segments to develop a SPF, CMF, and a calibration factor, which are then used to predict crashes for each segment.

Step Nine: Select and Apply SPF

This step develops the SPF for each selected roadway segment. The SPFs are used to determine predicted crash frequency with HSM base conditions. The SPF is adjusted to local conditions using the calibration factor in Step 11. For each segment, the SPF is found using the following equation in the HSM:

(1) $N_{spfrs} = AADT \times L \times 365 \times 10^{-6} \times e^{(-0.312)}$

Where:

N_{spfrs} = predicted total crash frequency for roadway segment base conditions (spfrs refers to the SPF for the roadway segment)

AADT = average annual daily traffic volume (vehicles per day)

L = length of roadway segment (miles) or length of curve

The HCM also provides default distributions for crash severity and collision type which are based on data for Washington State. These distributions may also be updated using local data for improved accuracy.

Step Ten: Apply the Appropriate CMFs to SPF to Account for the Difference in Base and Site-Specific Conditions

After an SPF is found for base conditions in each segment, it is multiplied by the appropriate CMFs to adjust the estimated crash frequency to site specific conditions. For example, if the road segment does not have a shoulder, the SPF estimate is adjusted by a CMF of 1.50 to show an increase in predicted crashes.[1] The HSM identifies 12 appropriate CMFs for horizontal curves. The most common CMFs used to adjust for local conditions on the curves are: lane width, shoulder width and type, length, radius, and presence or absence of spiral transition, superelevation, grade, and driveway density.

The Federal Highway Administration (FHWA) has established a CMF Clearinghouse for CMFs.[2] These CMFs are multiplicative factors used to estimate the change in the number of crashes after a given countermeasure is implemented under specific conditions. Included in this Clearinghouse are the horizontal CMFs that have been developed. Of the 2,546 CMFs from 150 studies that are included in the Clearinghouse (as of March 2011), 221 CMFs and 18 studies relate to horizontal curves. However, due to the base conditions in this HSM analysis, only the CMFs presented in Section 10.7 of the HSM can be used with the SPFs developed in the previous step. The development process for CMFs are presented in Part D of the HSM, but the focus of improvements is toward future editions of the HSM, not for inclusion of additional CMFs by the user.

Step Eleven: Apply a Calibration Factor to the Result of Step 10

Once the estimated crash frequency for each segment is found and adjusted for site-specific conditions, it is multiplied by an appropriate calibration factor developed for local conditions. The calibration factor is used to adjust the results of the HSM predictive model to local conditions and it is calculated as the ratio of total observed crash frequency to total expected average crash frequency during the same period. For example, for one group of curves in this analysis, the reported collisions per year is 8.8 and the predicted number of collisions from the HSM procedure is 6.6, resulting in a calibration factor of 8.8/6.6 or 1.33. In this analysis, several calibration factors were developed, including an overall factor for all segments, a non-random selection curve segment factor, and a random selection curve segment factor. The results of this analysis are presented in the next section.

ANALYSIS

Calibration Factor Analysis

The calibration factor is a critical component of the HSM procedure to adjust the standardized factors presented in the manual to account for local conditions (e.g., crash reporting thresholds, climate and geographic features, and driver factors for a given state). This paper focuses on calculating a calibration factor for two-lane rural road segments, including curved segments, tangent segments, and composite segments (including all curves and tangents). The HSM recommends that the calibration factors should be calculated every two or three years for those who wish to implement the procedures in the manual regularly. Additionally, the manual specifies a desirable minimum sample size of 30 to 50 sites that experience a total of at least 100 collisions per year. This analysis included 51 sites that experienced 85 collisions per year on average, over a five-year period (Table 2). However, these 51 sites include 26 curve segments that have abnormally high collision histories or have previously been identified as hazardous locations. The other 25 sites were selected randomly by NCDOT personnel by arbitrarily choosing a curve site while on other assignments.

Table 2: HSM Calibration Factors Calculated

Sample Type (Sample Size)	Roadway Type	Calibration Factor	Reported Collisions (Collisions per Year)	Predicted Collisions (Collisions per Year)
All Segments (51)	Curve	2.82*	35.4	12.5
	Tangent	1.12	49.4	44.0
	Composite	1.50*	84.8	56.5
Random Selection (25)	Curve	1.33	8.8	6.6
	Tangent	1.00	20.4	20.4
	Composite	1.08	29.2	27.0
Non-random Selection (26)	Curve	4.5*	26.6	5.9
	Tangent	1.23	29.0	23.6
	Composite	1.88*	55.6	29.5

*Denotes a statistical difference from a calibration factor of 1.00 at the 95% confidence level.

HSM calibration factors were calculated by first applying the HSM method to calculate the predicted number of crashes using site characteristic data like lane width, shoulder width, and roadside design. Once these predicted crashes were found, the calibration factor was calculated as the ratio of observed to predicted crashes. For example, in this analysis, the observed number of curve crashes for all 51 segments was 35.4 and the predicted 12.5, resulting in a calibration factor of 2.83. The calibration methodology implied by the HSM involves using extended roadway sections consisting of numerous tangent and curve sites. However, to examine the differences between tangents and curves, this analysis considered curve and tangent sections individually (and combined as “composite” sections). The HSM does not specify how calibration segments should be selected or if high crash location data should be used for this purpose. But Table 2 shows that the inclusion of high crash locations significantly impacts the calibration factor. When considering curved roadway segments, the calibration factor varies from 2.83 when including all 51 sites to 1.33 when counting only those sites that were randomly selected, and to 4.5 when incorporating only high crash sites. To meet HSM recommendations for collisions, additional sites would be needed in each sample type. For instance, if a user decided to develop a two-lane curve calibration factor based on randomly selected curves to meet the criterion of 100 total crashes, almost 300 sites would be needed in the analysis. Collecting the detailed data needed to calibrate the HSM for 300 curve sites would require an appropriate amount of additional labor.

A paired t-test was conducted to examine the importance or need for the calibration factors in Table 2. The test compared the reported and predicted collisions among each type of sample. The comparison found a difference in reported and predicted collisions in four of the nine samples and roadway types, indicating that only four of the calibration factors differed significantly from a calibration factor of one. Besides this finding, annual variations could exist when calculating calibration factors. Table 3 shows five years of calibration factors from the same data in Table 2. The calibration factor chosen in Table 3 for each year used only one year of data, so the samples of collisions were small. This table can provide users with an estimate of how much variation could exist when calculating annual calibration factors.

Table 3: Annual Calibration Factors (All Segments, Random Segments, and Non-Random Segments)

Sample Type (Sample Size)	Roadway Type	2004 Calibration Factor	2005 Calibration Factor	2006 Calibration Factor	2007 Calibration Factor	2008 Calibration Factor	Standard Deviation
All Segments (51)	Curve	2.63	2.07	3.19	3.75	2.47	0.65
	Tangent	1.04	1.14	1.11	1.32	1.00	0.12
	Composite	1.40	1.34	1.57	1.86	1.33	0.22
Random Selection (25)	Curve	1.36	1.51	1.97	1.06	0.76	0.46
	Tangent	0.88	0.98	0.78	1.13	1.22	0.18
	Composite	1.00	1.11	1.07	1.11	1.11	0.05
Non-random Selection (26)	Curve	4.05	2.70	4.56	6.75	4.39	1.46
	Tangent	1.19	1.27	1.40	1.48	0.80	0.26
	Composite	1.76	1.56	2.03	2.54	1.52	0.42

At each site, the field investigation to collect all necessary elements for HSM analysis took approximately 30 minutes to complete (not including driving time). Thus, the requirement of 300 sites to develop a calibration factor for curve sites could be expected to require at least one person-month of labor, plus drive time between sites. However, most of these elements do not change much or at all over time. So the data collected intensively for the first HSM calibration or application can likely be used for many years. The effort required to collect collision data varies in the way the data are stored and how efficiently they can be retrieved. And that effort may be substantial in some agencies. For example, field data collection efforts could vary considerably by agency, depending on the desired precision of crash prediction and available data sources within the agency. However, field data collection might not be necessary for some agencies or could require similar time commitments as noted in this study. Appendix A to Part C of the HSM defines data needs for each element as required or desirable, and provides suggested assumptions for defaults, average values, and actual data. Implementing the concepts for data collection presented in the HSM, along with utilizing available computer-based techniques (inventories, GIS data, and design plans), can significantly reduce or eliminate the need for field data collection, thereby reducing labor requirements substantially for the calibration effort.

Sensitivity Analysis

The sensitivity analysis focused on the effect of changing various HSM inputs on the number of predicted collisions. The objective of this analysis was to understand the most critical HSM inputs that might lend themselves more readily to default values, thus saving data collection effort. Several HSM inputs were not included in this sensitivity analysis because little or no variation existed among the curves in our sample. These included spiral transition, passing lanes, roadway lighting, centerline rumble strips, two-way left-turn lanes, and automated speed enforcement. Table 4 shows descriptive statistics about the data.

Table 4: Input Values for HSM (Minimum, Maximum, and Average)

HSM Input Factor	Minimum Value	Maximum Value	Mean Value
AADT	240	21,000	3,885
Lane Width (feet)	9	12	10.4
Inside Shoulder Width (feet)	3	12	7.4
Outside Shoulder Width (feet)	3	12	8.0
Length of Horizontal Curve (feet)	200	1,550	579
Radius of Horizontal Curve (feet)	202	6,011	1,360
Super elevation (feet/feet)	0.010	0.102	0.056
Grade (%)	0.0	5.1	1.3
Driveway Density (driveways/mile)	0.0	54.6	9.6
Roadside Hazard Rating (1-7)	3.0	6.0	3.8

Utilizing the field data resulted in a predicted collision rate of 12.5 collisions per year for the set of 51 curves. Table 5 shows the HSM outputs from the sensitivity analysis. The table emphasizes the importance of collecting and using individualized data for AADT, curve radius, and curve length of the segment. The AADT had a range of 62.6 predicted collisions per year between using the minimum value and using the maximum value. There was also a 0.9 collisions per year (or 7%) difference between the predicted collisions using the averages of the inputs and actual field values. Radius had a range of 18.9 predicted collisions per year between using the minimum and maximum values. There was also a 0.8 collisions per year (or 7%) difference between the predicted collisions using the average input and actual field values. Length had a range of 19.5 predicted collisions per year between using the minimum value and maximum values, and there was a 0.5 collisions per year (or 4%) difference between predicted collisions using the averages of the inputs and actual field values.

Table 5: Output Values from HSM (Predicted Collisions Per Year)

HSM Input	HSM Predicted Collisions per Year for Set of 51 Curves				
	Using Minimum Value from Table 1	Using Maximum Value from Table 1	Difference Between Using Maximum Value and Minimum Value	Using Mean Value from Table 1	Difference Between Using Mean Value and Actual Field Measured Values
AADT	0.8	63.4	62.6	13.4	0.9
Lane Width	14.7	11.5	3.2	13.4	0.9
Inside Shoulder Width	13.4	12.3	1.1	12.5	0.0
Outside Shoulder Width	13.4	12.3	1.1	12.2	0.3
Length of Curve	6.6	26.1	19.5	12.1	0.5
Radius of Curve	28.4	9.5	18.9	11.7	0.8
Superelevation	14.1	11.7	2.4	12.5	0.0
Grade	12.3	13.5	1.3	12.6	0.1
Driveway Density	11.5	21.5	10.0	12.4	0.1
Roadside Hazard Rating	11.9	14.6	2.6	12.6	0.0

The number of predicted crashes on curves for various traffic and geometric conditions using the HSM base model (not the calibrated North Carolina model) is in Table 6. Specifically, the variables that had the most effect on the number of crashes on curves (as measured by the difference between the minimum and maximum values) are AADT, curve radius, and length of curve. In Table 6, the predicted crashes on curves for five-year periods are based on the crash-prediction model for AADT's of 500, 1,000, 2,000, 5,000, 10,000, and 20,000. This table has a range of curve radius from 250 feet to 5,000 feet, and a range of curve lengths from 250 to 1,500 feet. For example, for a curve on a road with an AADT of 1,000, a 500-foot radius, and length of 750 feet, the expected number of curve crashes per five years would be approximately 0.54 (i.e., one curve crash every 10 years), as shown in Table 6. The calculations in Table 6 assume average NC conditions for the other variables included in the prediction model for crashes on curves. Specifically, it assumes a lane width of 10.4 feet, inside shoulder width of 7.4 feet, outside shoulder width of eight feet, superelevation of 0.056, grade of 1.3 %, 9.6 driveways per mile, average roadside hazard rating of four (on a seven-point scale), speed limit of 55 mph, no transition spiral, no passing lane, no roadway lighting, no centerline rumble strips, no two-way left-turn lane, and no automated speed enforcement.

Calibration Factor Validation

Calibration is a critical task to adjust a broad model for local analysis. In the case of the HSM, the calibration provides a multiplicative factor to adjust the predicted model to account for differences that are not determined by physical roadway elements, such as driver population, reporting threshold, and others. However, the calculation of a calibration factor in a research setting is incomplete without validating the model with the predetermined calibration factor. Therefore, an additional set of curve geometric and collision data was acquired to validate the previously calculated calibration factor.

Table 6: Predicted Collisions (over 5 years) for Two-Lane Road Horizontal Curves

Radius (Degree of Curvature)	Predicted Collisions for 500 vehicles/day						Predicted Collisions for 1,000 vehicles/day					
	Length (feet)						Length (feet)					
	250	500	750	1,000	1,250	1,500	250	500	750	1,000	1,250	1,500
250 feet (22.9°)	**	0.30	0.35	**	**	**	**	0.66	0.76	**	**	**
500 feet (11.5°)	0.15	0.20	0.25	**	**	**	0.33	0.43	0.54	**	**	**
1000 feet (5.7°)	0.10	0.15	0.20	0.24	0.29	0.34	0.22	0.32	0.42	0.53	0.63	0.73
2000 feet (2.9°)	0.07	0.12	0.17	0.22	0.26	0.31	0.16	0.26	0.37	0.47	0.57	0.68
3000 feet (1.9°)	0.07	0.11	0.16	0.21	0.26	0.30	0.14	0.24	0.35	0.45	0.55	0.66
4000 feet (1.4°)	0.06	0.11	0.16	0.20	0.25	0.30	0.13	0.23	0.34	0.44	0.54	0.65
5000 feet (1.1°)	**	0.11	0.15	0.20	0.25	0.30	**	0.23	0.33	0.44	0.54	0.64

Radius (Degree of Curvature)	Predicted Collisions for 2,000 vehicles/day						Predicted Collisions for 5,000 vehicles/day					
	Length (feet)						Length (feet)					
	250	500	750	1,000	1,250	1,500	250	500	750	1,000	1,250	1,500
250 feet (22.9°)	**	1.52	1.76	**	**	**	**	3.80	4.40	**	**	**
500 feet (11.5°)	0.76	1.00	1.24	**	**	**	1.90	2.50	3.10	**	**	**
1000 feet (5.7°)	0.50	0.74	0.98	1.22	1.46	1.69	1.25	1.85	2.44	3.04	3.64	4.24
2000 feet (2.9°)	0.37	0.61	0.85	1.09	1.32	1.56	0.92	1.52	2.12	2.71	3.31	3.91
3000 feet (1.9°)	0.33	0.56	0.80	1.04	1.28	1.52	0.81	1.41	2.01	2.61	3.20	3.80
4000 feet (1.4°)	0.30	0.54	0.78	1.02	1.26	1.50	0.76	1.36	1.95	2.55	3.15	3.75
5000 feet (1.1°)	**	0.53	0.77	1.01	1.25	1.49	**	1.32	1.92	2.52	3.12	3.71

Radius (Degree of Curvature)	Predicted Collisions for 10,000 vehicles/day						Predicted Collisions for 20,000 vehicles/day					
	Length (feet)						Length (feet)					
	250	500	750	1,000	1,250	1,500	250	500	750	1,000	1,250	1,500
250 feet (22.9°)	**	7.61	8.80	**	**	**	**	15.22	17.61	**	**	**
500 feet (11.5°)	3.80	5.00	6.19	**	**	**	7.61	10.00	12.39	**	**	**
1000 feet (5.7°)	2.50	3.69	4.89	6.08	7.28	8.47	5.00	7.39	9.78	12.16	14.55	16.94
2000 feet (2.9°)	1.85	3.04	4.24	5.43	6.62	7.82	3.69	6.08	8.47	10.86	13.25	15.64
3000 feet (1.9°)	1.63	2.82	4.02	5.21	6.41	7.60	3.26	5.65	8.04	10.42	12.81	15.20
4000 feet (1.4°)	1.52	2.71	3.91	5.10	6.30	7.49	3.04	5.43	7.82	10.21	12.59	14.98
5000 feet (1.1°)	**	2.65	3.84	5.04	6.23	7.43	**	5.30	7.69	10.08	12.46	14.85

Notes: Assumed values are mean values for lane width (10.5 feet), inside shoulder width (composite - six feet), outside shoulder width (composite - eight feet), superelevation (0.056 ft/ft), grade (1.3%), driveway density (9.6 driveways/mile), roadside hazard rating (4),speed limit (55 mph), no spiral transition, no passing lanes, no roadway lighting, no centerline rumb le strips, no two -way left-turn lanes, and no automated speed enforcement.

** = Data does not support generation of collisions for this combination of radius and length

This validation effort included two two-lane roads—NC42 and NC96—which are predominantly rural and run through central and eastern North Carolina. Horizontal curve data collection for this effort used GIS techniques (Rasdorf et al. 2012). The analysis included all the curved sections of each route except where a higher order route (i.e., a US route) ran concurrently with the NC route. The entire route of NC42 is 223 miles long; the analysis sections included 168 miles of this route and 246 curves. The entire route of NC96 is 107 miles long; the analysis sections included 95 miles of it and 174 curves. None of the calibration sites was included in the validation data set.

The HSM analysis of the curves predicted 114 collisions per year, while the curves experience 174 reported collisions per year, giving a calibration factor of 1.5267. Applying the suggested calibration factor of 1.33 from Table 2 to the HSM prediction gives 152 collisions per year. This is approximately 10% less than the reported collisions. Comparatively, the original prediction is approximately 35% less than the reported collisions. A paired t-test of the collision rates for each curve and for each of the three data sets (reported collisions, HSM prediction, and calibrated HSM prediction) shows that pairing reported collisions and calibrated HSM predicted collisions did not result in a statistically significant difference between them (95% confidence at p = 0.05). Comparatively, the other two pairings were statistically different. Differences in the data collection methods and randomness of collisions could contribute to the difference between reported collisions and the calibrated HSM prediction. Similar differences between predicted and actual crashes on horizontal curves could also be due to randomness of collisions. Furthermore, since the difference between the actual and HSM predicted curve crashes was not statistically significant, it might be reasonable to assume a calibration factor of one.

CONCLUSIONS

The publication of the Highway Safety Manual (HSM) offers agencies an analytical tool to evaluate the safety of a horizontal curve or set of curves efficiently and proactively. The HSM provides a crash prediction model for horizontal curves that can be applied to identify the highest priority locations for safety treatments as well as common and effective countermeasures. The calibration of the HSM predictive method was evaluated and tested on horizontal two-lane rural roads in North Carolina in this paper. Based on the analysis, it is found that approximately 300 curve sites are needed to meet HSM recommendations for the number of collisions in the calibration data set. This large number of sites is partly due to the finding that the selection of random segments provides a more accurate outcome (in terms of matching the HSM prediction model) than the crash results from a high-crash location group as identified by a transportation agency.

One challenge with requiring a large number of sites to develop an accurate model based on local conditions is the manpower needed for data collection. For each of these sites, field investigations took approximately 30 minutes to complete (not including driving time) for the collection of necessary elements for HSM analysis. However, most of the elements on which data were collected do not change much or at all over time. So, some data collection may not be needed during each calibration. Also, considerably less manpower may be required by some agencies that have curve inventories and/or in-house data sources for some of the needed curve features. To further lessen the data collection burden, an analysis of differences in predicted collisions based on field data collection and average or default values was performed. It was found that for AADT, curve radius, and curve length of the segment, individualized data are necessary for accuracy, but that the other data inputs may be assumed with less penalty for the accuracy of overall predicted crash value. It is possible for each of these three elements, which are the most sensitive to individual curve data, to be collected from existing agency GIS data or roadway inventory databases if available. These findings can allow agencies to more efficiently apply and utilize HSM procedures for the cost-effective improvement of curves, in some cases without a field visit.

To properly calibrate the predictive models to HSM standards, the research team found that at least for North Carolina conditions, approximately 300 segments are required to meet the HSM recommendations for collisions. This number of sites may vary for other local and state situations. Additionally, while it will require many sites, randomly selected road segments are recommended for this process because they have a low calibration factor. AADT, curve radius, and curve length are the most important data in terms of the prediction of curve crashes in the HSM model. Fortunately, as demonstrated elsewhere (Rasdorf et al. 2012), many of these important variables can be collected without a field visit, thus saving time and resources. A calibration factor or 1.33 was found to be appropriate to apply the HSM prediction method for it to match North Carolina crash values.

However, a calibration factor of one could be justified since the differences between the actual and HSM predicted number of curve crashes are not statistically significant.

Although some of the analyses performed for specific two-lane rural roads in this paper are not applicable to all road types or jurisdictions, several key findings can be applied by traffic engineers and researchers conducting similar analyses. First, engineers should consider the impact of site-specific or average data on their analysis. If some variables are not highly sensitive to small changes in value or if the variable of interest is fairly consistent among the type of roadways under analysis, an average value might be sufficient to minimize data collection costs. Secondly, annual or ongoing costs should be considered. Although the initial resource needs for the type of analyses presented in this paper is considerable, the costs to update the data in future years will be considerably less. Thirdly, site selection is a critical component of this process and randomly selected locations are preferred over high crash locations. Finally, with available data through internet-based sources and centralized databases, along with average values for non-sensitive elements, it is possible to achieve reasonable estimates of collisions on a roadway without a field visit.

Acknowledgement

This paper was based on a study funded by the North Carolina Department of Transportation entitled "Procedure for Curve Warning Signing, Delineation, and Advisory Speeds for Horizontal Curves." Brian Mayhew was the NCDOT Chair of the Steering and Implementation Committee. We thank the NCDOT for their project and data collection support.

Endnotes

1. See Table 10-9 of the HSM.

2. See www.cmfclearinghouse.org.

References

AASHTO. *Highway Safety Manual*. American Association of State Highway Transportation Officials, Washington, D.C., 2010.

Banihashemi, M. "Highway Safety Manual, New Model Parameters vs. Calibration of Crash Prediction Models." *Proceedings of the Transportation Research Board 90th Annual Meeting. Transportation Research Board*. National Research Council, Washington, D.C., 2011.

Easa, S. M. and Q. C. You. "Collision Prediction Models for Three-Dimensional Two-Lane Highways: Horizontal Curves." *Transportation Research Record: Journal of the Transportation Research Board* 2092, (2009): 48–56.

Findley, D. J. and R.S. Foyle. "Procedure for Identification and Investigation of Horizontal Curves with Insufficient Superelevation Rates." *North Carolina Department of Transportation*. NCDOT Report No. FHWA/NC/2009-08. Raleigh, NC, 2009.

Hummer, J. E., W. Rasdorf, D.J. Findley, C.V. Zegeer, and C.A. Sundstrom. "Curve Crashes: Road and Collision Characteristics and Countermeasures." *Journal of Transportation Safety and Security* 2 (3), (2010): 203-220.

Lord, D., P. Kuo, and S. R. Geedipally. "Comparison of Application of Product of Baseline Models and Accident-Modification Factors and Models with Covariates: Predicted Mean Values and Variance." *Transportation Research Record: Journal of the Transportation Research Board* 2147, (2010): 113-122.

Martinelli, F., F. La Torre, and P. Vadi. "Calibration of the Highway Safety Manual Accident Prediction Model for Italian Secondary Road Network." *Proceedings of the Transportation Research Board 88th Annual Meeting*. Transportation Research Board. National Research Council, Washington, D.C., 2009.

Nowakowska, M. "Logistic Models in Crash Severity Classification Based on Road Characteristics." *Transportation Research Record: Journal of Transportation Research Board* 2148, (2010): 16–26.

Rasdorf, W., D.J. Findley, C.V. Zegeer, C.A. Sundstrom, and J.E. Hummer. "Evaluation of GIS Applications for Horizontal Curve Data Collection." *Journal of Computing in Civil Engineering* 25 (2), (2012): 191-203.

Sun, X., Y. Li, D. Magri, and H.H. Shirazi. "Application of the Highway Safety Manual Draft Chapter: Louisiana Experience." *Transportation Research Record: Journal of the Transportation Research Board* 1950, (2006): 55-64.

Xie, F., K. Gladhill, K.K. Dixon, and C.M. Monsere. "Calibrating the Highway Safety Manual Predictive Models for Oregon State Highways." *Proceedings of the Transportation Research Board 90th Annual Meeting.* Transportation Research Board. National Research Council, Washington, D.C., 2011.

Zegeer, C., D. Reinfurt, T. Neuman, R. Stewart, and F. Council. "Safety Improvements on Horizontal Curves for Two-Lane Rural Roads – Informational Guide." *Federal Highway Administration.* Report FHWA-RD-90-074. Washington, D.C., 1991.

***Daniel J. Findley**, Ph.D., P.E., is a senior research associate in the Highway Systems Group at the Institute for Transportation Research and Education at North Carolina State University in Raleigh, NC. He specializes in horizontal curve safety, economic impact analysis, asset management, unique transportation engineering studies, and logistics. He holds B.S., M.S., and Ph.D. degrees in civil engineering from North Carolina State University. Findley also possesses a professional engineer's (PE) license in the state of North Carolina.*

***Charles V. Zegeer**, M.S., P.E., is associate director for engineering and planning for the Highway Safety Research Center (HSRC) at the University of North Carolina in Chapel Hill, NC. He specializes in safety research related to roadway geometric design, traffic control devices, crash analysis systems, and pedestrian and bicyclist safety. Zegeer holds a B.S. degree in civil engineering from Virginia Tech and an M.S. eegree in civil engineering (transportation specialty) from the University of Kentucky.*

***Carl A. Sundstrom**, P.E., is a senior research associate at the University of North Carolina Highway Safety Research Center. He specializes in general safety studies, crash modification factors, and bicycle and pedestrian research and education. Sundstrom holds B.S. and M.S. degrees in civil engineering from the Georgia Institute of Technology.*

***Joseph E. Hummer**, Ph.D., P.E., is a professor of civil engineering at North Carolina State University, where he has taught and researched traffic safety, traffic operations, highway design, and transportation policy since 1992. He has developed many collision modification factors for safety countermeasures and collision prediction models. Hummer has a Ph.D. from Purdue University and bachelor's and masters degree's from Michigan State University.*

***William Rasdorf**, Ph.D., P.E., is a professor at North Carolina State University in Raleigh, NC. He specializes in modeling of engineering business processes, information technology, database design and development, and asset management for the solution of societal civil engineering problems. Rasdorf holds a B.A.E. and M.S. from Penn State University and M.S. and Ph.D. degrees in civil engineering from Carnegie Mellon University.*

Predicting Block Time: An Application of Quantile Regression

by Tony Diana

Airlines face three types of delay that make it difficult to build robust schedules and to support block time predictability. Block time is the time elapsed from gate departure to gate arrival and refers to the time when blocks are off the wheels at the departure airport to the time they are back on at the destination airport. These delays can be induced (i.e., ground delays), propagated, or stochastic. With capacity constrained at major airports and regulators facing greater public pressure to alleviate congestion and tarmac delays, aviation practitioners have renewed their interest in the predictability of block time. This study presents a methodology based on the case study of the Seattle/Tacoma International (SEA) and Oakland International airport (OAK) city pair to determine the predictability of block time. The methodology based on quantile regression models is appropriate for a skewed distribution where analysts are interested in the impact of selected operational variables on the conditional mean of block times at given percentiles. Quantile regression provides a measure of on-time performance based on the percentile results that show the most significance and best fit.

INTRODUCTION

Block time refers to the time that an aircraft spends from gate departure to gate arrival. Pilots are usually paid on the basis of "lock time or better," meaning the greater of scheduled or actual gate-to-gate time. Actual block time depends on external factors such as available airport capacity, ground surface congestion, en route delays, weather events, air traffic control delays, and airline operational issues, among others. To minimize the impact of these unanticipated conditions, airlines have some incentive to pad their schedules so as to make on-time performance look better. The padding is all the more important as U.S. airlines schedule for good weather condition (visual meteorological conditions) compared with European airlines that take into account reduced capacity (instrument meteorological conditions) and whose traffic at large hubs is slot-constrained. Airlines often use on-time performance as an important marketing argument to attract passengers. Both on-time performance and the causes of delay are published monthly by the Bureau of Transportation Statistics[1] (BTS) in the Airline Service Performance Quality (ASQP) report. Therefore, scheduled block time is often construed as a measure of passenger experience.

This article proposes a methodology to determine the predictability of block time based on the case study of the Seattle-Oakland city pair. The proposed methodology relies on quantile regression to determine how some selected operational variables are likely to affect actual block times at different percentiles. This is of importance to aviation practitioners and, especially, airline schedulers who have often resorted to schedule padding to make up for ground and en route delays.

Predictability is all the more difficult to achieve as airlines often face three types of delay. First, delays can be induced: The Federal Aviation Administration (FAA) can initiate a ground delay program in case of adverse weather conditions or heavy traffic volume on the ground or en route. These delays are reported by air carriers as National Airspace System (NAS) delay when non-extreme weather conditions or airspace/airport conditions prevent on-time operations. Second, delays can be propagated; in a sequence of legs operated by the same tail-numbered aircraft, a flight may accumulate delays that cannot be recovered by the end of an itinerary. These delays are usually reported as late arriving aircraft delays. Finally, delays can be stochastic because they are the results of random events such as equipment breakdown or crew problems (air carrier-related delays), security, or an extreme weather event.

Predictability represents an important key performance indicator in the aviation industry for several reasons.

- For the International Civil Aviation Organization (ICAO), predictability refers to the “ability of the airspace users and ATM [Air Traffic Management] service providers to provide consistent and dependable levels of performance.” [2] Air Traffic Management (ATM) services can be public agencies such as the FAA or a private non-share capital corporation, such as NavCanada. The fundamental mission of ATM is to ensure flight safety by enforcing separation between aircraft.
- One of the goals of the U.S. Next Generation of Air Transportation System (NextGen) is to foster the transition from an air traffic control to more of an air traffic managed system where pilots have more flexibility to select their routes, utilize performance-based navigation (PBN), and make decisions based on automated information sharing. Performance-based navigation refers to either required navigation performance (RNP) when navigation entails on-board performance monitoring and alerting or area navigation (RNAV) when there are no requirements for monitoring and alerting. PBN procedures enable aircraft to fly more efficient arrival and departure trajectories not previously available due to the constraints of ground-based navigation aids such VHF omnidirectional radio range.[3] PBN makes it possible for aircraft to operate at airports that are difficult to access because of surrounding terrain or airspace congestion. Presently, it is very difficult to assess the impact of NextGen-related technologies on flight performance because surveillance data do not account for the difference between the use of required navigational performance and instrument landing systems (ILS) when flight tracks for both types of procedure overlay, for instance. Surveillance data are generated by radar such as the Traffic Flow Management System data.
- According to Rapajic (2009), “cutting five minutes off average of 50% of schedules thanks to higher predictability would be worth some €1,000 million per annum, through savings or better use of airlines and airport resources.” Unpredictability imposes considerable costs on airlines in the forms of lost revenues, customer dissatisfaction, and potential loss of market share.

Recently, much discussion has revolved around the validity of using airlines’ schedules as a measure of on-time performance and the variance of block delay as an indicator of predictability. Both airlines’ limited control over the three types of delay and airport congestion make it difficult to build robust schedules. In this discussion, the predictable block time is located at the percentile where the sign and magnitude of the pseudo coefficient of determination (a measure of goodness of fit) is the highest with all the explanatory variables significant at a given confidence level. Ordinary Least Square (OLS) regression models enable analysts to evaluate the percentage of variation in actual block time explained by changes in selected operational variables. However, quantile regression is more robust to outliers than the traditional OLS regression because the latter does not focus on the conditional mean. The attributes of quantile regression will be addressed later in the discussion.

This article presents a different perspective on the study of predictability with the intent of helping aviation practitioners achieve the following objectives:

- To assess the impact of selected independent variables at different locations of the distribution of block delays in order to anticipate block time based on selected operational variables.
- To derive more predictable block times based on the impact of operational independent variables at various percentiles.
- To test a model without any assumption about the distribution of errors and homoscedasticity.

After a brief background, the discussion will proceed with the methodology, an explanation of the outcomes, and, eventually some final comments.

BACKGROUND

There has been much discussion recently about the impact of schedule buffers and their reliability as a tool to measure airline or even the National Airspace System (NAS) performance. Constructing robust schedules is important for an airline because they support profitability. Lohatepanont and Barnhart (2004) focused on fleet assignment to determine where and when flights should be offered and what type of equipment should be used to maximize profits.

Wu (2005) highlighted the difference between the real operating delays, the inherent delays (from simulation) and the zero-delay scenario. The reliability is also affected by delay propagation when an aircraft accumulates delays over a series of legs that cannot be recovered at the end of the total trip. Wu (2005) recommended that airlines integrate buffers in their schedule in a way that strengthens reliability. This article provides a methodology that evaluates the effect of selected operational variables on block time in order to support more reliable schedules.

Lan et al. (2006) proposed two methodologies to minimize passenger disruptions and achieve robust scheduling based on aircraft routing and retiming flight departure times. The purpose of their research was to identify ways to minimize passenger disruption and minimize delay propagation through mixed integer programming. The disadvantage of such methodology is that it does not take into consideration the impact of key operational variables on block time that includes ground movement operations and flight time.

Robust airline scheduling is the outcome of four sequential tasks, including schedule generation, fleet assignment, aircraft routing, and crew pairing/rostering (Wu 2010, Abdelghany and Abdelghany 2009). Fleet assignment models (FAM) are often used to determine how demand for air travel is met by available fleet (see Abara 1989 and Hane et al. 1995). Moreover, the fleet assignment models present two challenges: complexity and size of the problem that the FAM can handle.

Rapajic (2009) identified network structure and fleet composition as sources of flight irregularities. Wu (2010) provided an excellent exposition of issues related to delay management, operating process optimization, and schedule disruption management. Wu (2010) explained that "airline schedule planning is deeply rooted in economic principles and market forces, some of which are imposed and constrained by the operating environment of the [airline] industry." He presented a schedule optimization model to improve the robustness of airline scheduling. However, such a model does not consider how selective operational variables are likely to impact scheduling.

Morrisset and Odoni (2011) compared runway system capacity, air traffic delay, scheduling practices, and flight schedule reliability at 34 major airports in Europe and the United States from 2007 to 2008. The authors explained that European airports limit air traffic delay through slot controls. The other difference is that declared capacity (therefore, the number of available slots) is based mainly on operations under instrument meteorological conditions. In Europe, slot refers to a time window when a flight is scheduled to depart. By not placing restrictions on the number of operations, schedule reliability in the United States is all the more dependent on weather conditions as European airports.

METHODOLOGY

The Sample and the Assumptions

The sample includes daily data for the months of June to August in 2000, 2004, 2010, and 2011 for the Seattle/Tacoma International (SEA)-Oakland International (OAK) city pair. The summer season is usually characterized by low ceiling and visibility that determine instrument meteorological conditions and weather events such as thunderstorms—all likely to skew the distribution of block times.

Figures 1 to 4 show the shape of the distribution for each summer and provides some key statistics. The skewness coefficients are 0.11, -0.44, 0.37, and 0.19 respectively for summer 2011, 2010, 2004, and 2000. A negative skew indicates that the left tail is longer. While the standard deviation is appropriate to measure the spread of a symmetric distribution, interquartile ranges are more indicative of spread changes in skewed distributions. The Jarque-Bera statistic indicates whether the data are from a normal distribution. A normal distribution has an expected skewness and kurtosis of zero. A small probability value implies the rejection of the null hypothesis (H0: the distribution is normal).

Figure 1: Histograms of Block Time
(Counts of Flights by Average Minutes of Block Time, June-August 2011)

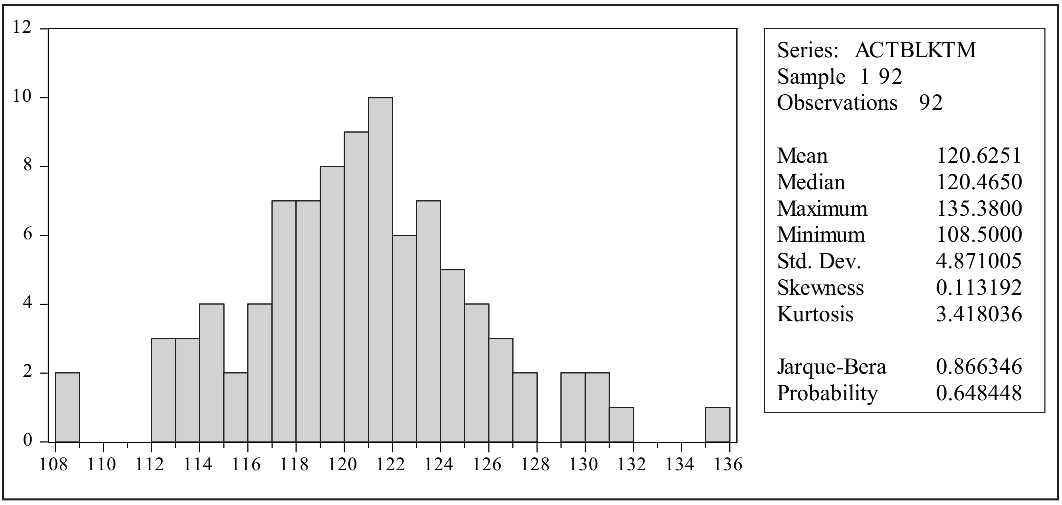


Figure 2: Histograms of Block Time
(Counts of Flights by Average Minutes of Block Time, June-August 2010)

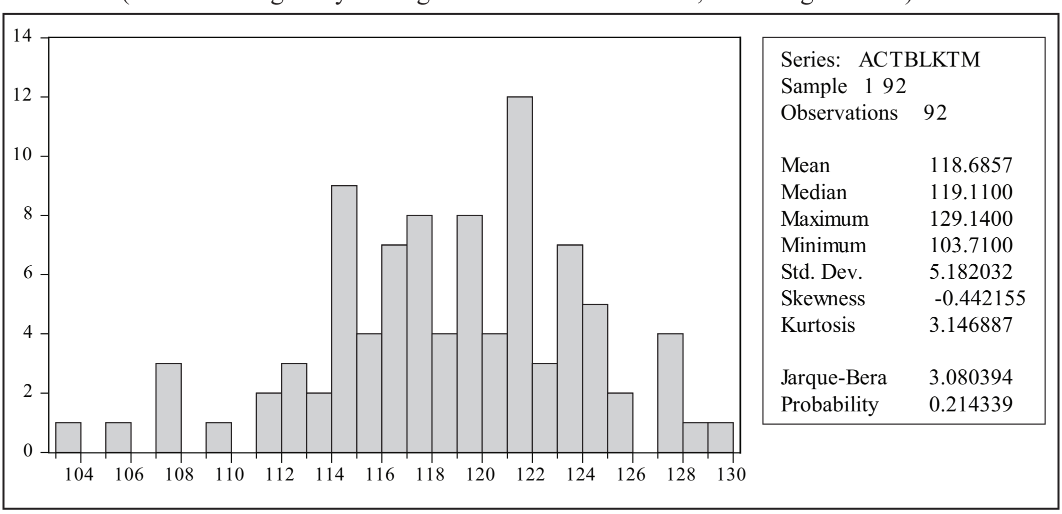


Figure 3: Histograms of Block Time
(Counts of Flights by Average Minutes of Block Time, June-August 2004)

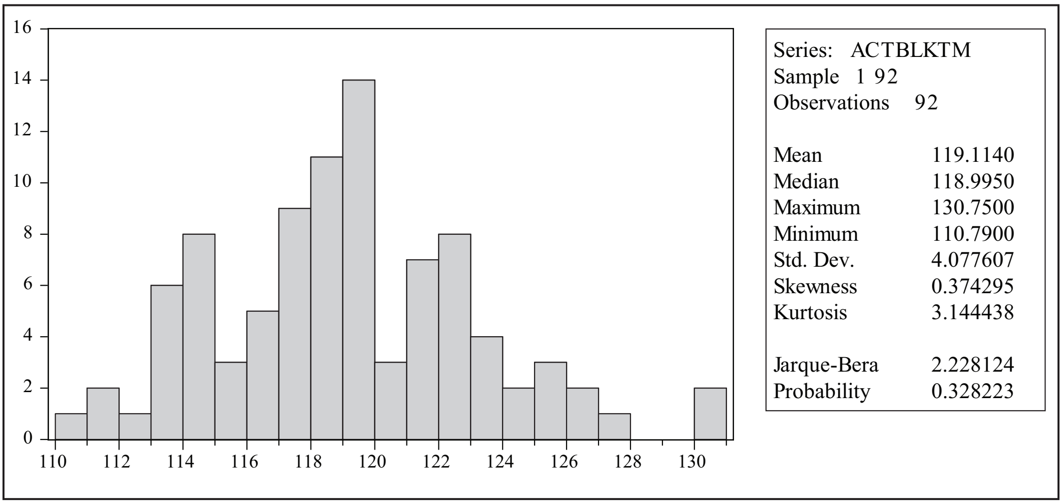


Figure 4: Histograms of Block Time
(Counts of Flights by Average Minutes of Block Time, June-August 2000)

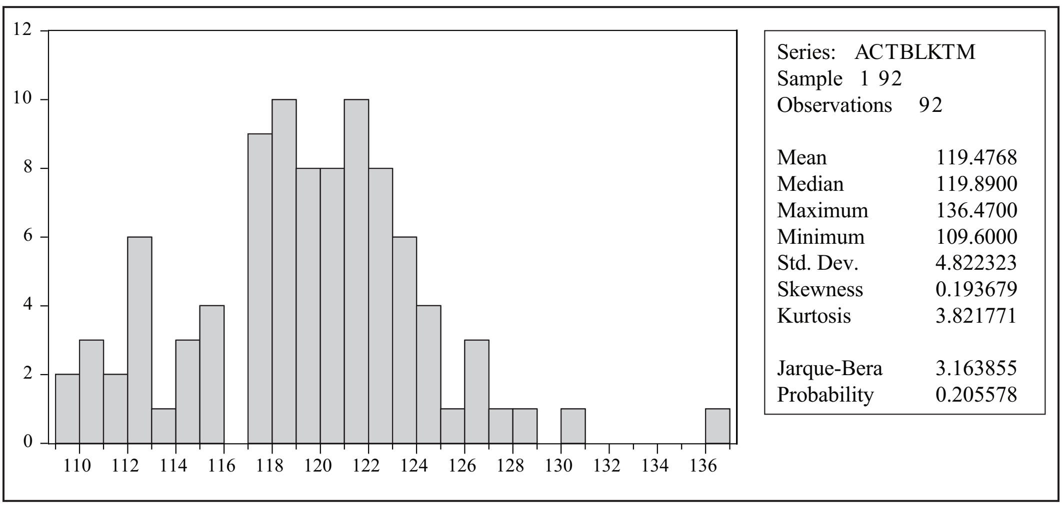
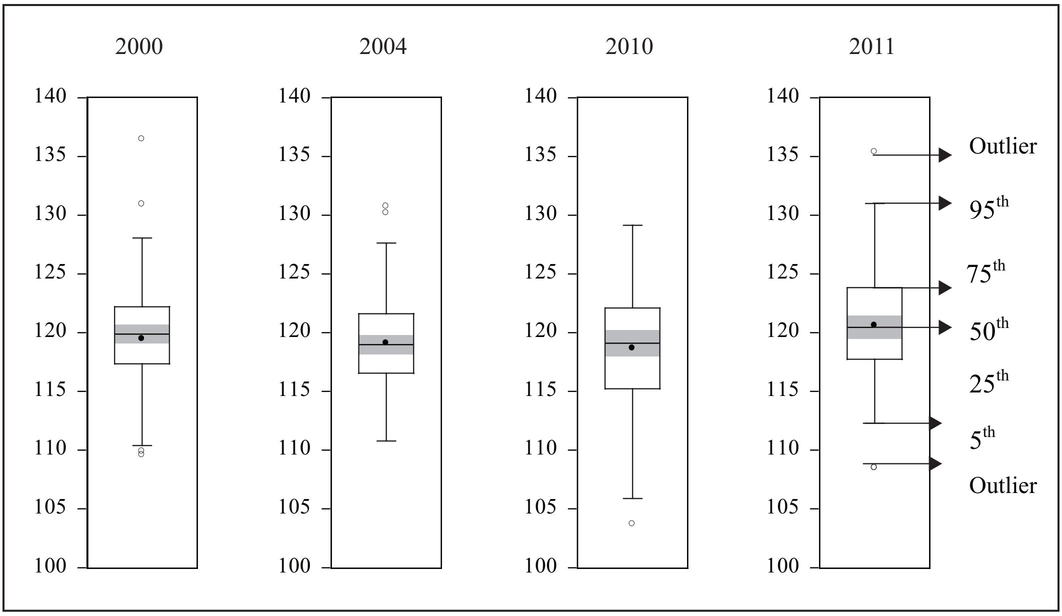


Figure 5 compares the boxplots of actual block times in minutes for the four summers under investigation. The boxplot shows the spread of the distribution, the selected quantile values, the position of the mean and median block times, and the presence of outliers that make it important to consider a regression model at different quantiles. The boxplots reveal an increase in the actual block times between summer 2004 and 2011. Summer 2010 features the largest range as well as the lowest block times at the 5th percentile among the four samples. It is also characterized by the highest proportion of operations in instrument meteorological conditions compared with the other three samples (Table 1).

Secondly, summer is part of the high travel season when demand is usually at its peak. This, in turn, is likely to increase airport and en route congestion and subsequently impact block time. Finally, the years were selected to account for the following conditions: (1) pre- and post-September 11, 2001, traffic, (2) lower traffic demand resulting from the 2008-2009 economic recession, and

Figure 5: Boxplots of Actual Block Time (June - August, in Minutes)



Note: The black dot represents the mean

(3) the introduction of the Green Skies over Seattle after 2010. It is an airline industry-wide initiative (i.e. airline, airport, aircraft manufacturer and FAA) designed to maximize performance-based navigation through the use of satellite navigation in order to ensure more direct and optimized landing approaches.

In Table 1, although the number of flights decreased between 2000 and 2011 and the average minutes of expected departure clearance times (EDCT) were higher in 2011 than in 2000, the percentage of on-time gate departures and arrivals and other key delay indicators such as taxi-out delay (a measure of ground congestion) improved in 2011. It is interesting to point out that the percentage of flights in IMC did not change significantly at OAK among the four selected summers. IMC operations were, however, much higher in 2010 and 2011 than in 2000 at SEA, which may explain the existence of average minutes of EDCT in 2010 and 2011.

The sample does not include a variable that measures performance-based navigation. The available surveillance data such as Traffic Flow Management System (TFMS) do not capture whether a pilot had requested a required navigation performance procedure, whether air traffic control had granted the request, and whether the procedure had actually been implemented. Surveillance data refer to information generated by radars. Moreover, it is presently difficult to differentiate flown performance-based navigation procedures from instrument landing system (ILS) approaches in the case of flight track overlay.

Secondly, the availability of Q-routes makes it possible for RNAV/RNP capable aircraft to reduce mileage, to minimize conflicts between routes (especially in a congested airspace such as the San-Francisco/Oakland area), and to maximize high-altitude airspace. The Q-routes are en route high altitude RNAV airways identified by a Q number. For instance, the great circle route between OAK and SEA is 584 nautical miles. The Q5 between the two airports is 523 nautical miles. Q-routes are designed to reduce flight distance and travel time between a city pair. They are available for use by RNAV/RNP capable aircraft between 18,000 feet MSL (mean sea level) and FL 450 inclusive (flight level 45,000 feet). Q-routes help minimize mileage and reduce conflicts between routes.

Table 1: Performance Metrics for the SEA-OAK City Pair

SEA-OAK	Flight Count	% On-Time Gate Departures	% On-Time Airport Departures	% On-Time Gate Arrivals	Arrivals With EDCT*
2000	1,411	76.26	67.54	77.18	0
2004	1,473	73.52	68.30	79.16	0
2010	1,147	91.63	85.27	92.85	1
2011	1,196	93.73	90.64	94.90	1

SEA-OAK	Average EDCT Where EDCT>0	Gate Departure Delay (min)	Taxi Out Delay (min)	Average Taxi Out Time (min)	Airport Departure Delay (min)
2000	0	9.07	4.48	14.29	12.83
2004	0	11.03	3.49	12.80	13.84
2010	94	4.48	3.01	13.34	6.97
2011	104	3.83	2.79	13.03	5.92

SEA-OAK	Airborne Delay (min)	Taxi In Delay (min)	Block Delay (min)	Gate Arrival Delay (min)	Percent IMC** SEA	Percent IMC** OAK
2000	4.33	0.95	2.61	9.23	10.58	29.49
2004	6.93	1.22	1.44	9.26	8.33	29.71
2010	5.05	0.50	1.67	3.76	30.29	29.86
2011	3.92	0.70	1.90	3.24	22.83	29.78

* In the event of a ground delay, airlines are issued an expected departure clearance time (EDCT).
Flights held by FAA at the departure airport due to problems at the arrival airport.
EDCT hold delay is computed by comparing EDCT wheels-off time to the flight plan's wheels-off time.
** Instrument Meteorological Conditions
Performance is compared with the last flight plan filed before take-off.

Source: FAA, Aviation System Performance Metrics

Thirdly, block time as a measure of gate-to-gate performance is sensitive to delays on the ground and en route. To account for this, airborne delay represents a surrogate for en route congestion, while increases in taxi times imply surface movement congestion.

Sources and Definition of the Variables

The sources for the variables are ARINC's[4] Out-Off-On-In times and the FAA's Traffic Flow Management System (TFMS). The directional city pair data originated from the "En Route" and "Individual Flights" sections of FAA's Aviation System Performance Metrics data warehouse.[5]

The choice of variables reflects operational and statistical considerations. On the one hand, some model variables represent core factors in airport congestion (taxi times) and en route performance (airborne delays). On the other hand, the model with the highest values for the Akaike Information Criterion (AIC)[6] and Bayesian Information Criterion (BIC)[7] was selected in order to prevent overfitting and to reduce the number of explanatory variables.

The dependent (response variable) and independent variables (also called covariates in the literature on quantile regression) are defined as follows:

- ***Actual Block Time*** (ACTBLKTM) is the dependent variable. It refers to the time from *actual* gate departure to *actual* gate arrival.
- ***Block Buffer*** (BLKBUFFER) represents the difference between planned and optimal block time.[8] The latter is the sum of unimpeded taxi-out times and filed estimated time en route. Block buffer is the additional minutes included in planned block time in order to take into account potential induced, propagated, and stochastic delays. According to Cook (2007), the block buffer is "the additional time built into the schedule specifically to absorb delay whilst the aircraft is on the ground and to allow recovery between the rotations of aircraft." Donohue et al. (2001) explained that "to obtain their desired on-time performance, airlines will add padding into a schedule to reflect an amount above average block times to allow for delay and seasonally experienced variations in block times."
- ***Departure Delay*** (DEPDEL) corresponds to difference between the actual and planned gate departure time at the departure airport in a city pair.
- ***Arrival Delay*** (ARRDEL) refers to the difference between the actual and planned gate arrival time at the arrival airport in a city pair.
- ***Airborne Delay*** (AIRBNDEL) accounts for the total minutes of airborne delay. It is the difference between the actual airborne times (landing minus take-off times) minus the filed estimated time en route.
- ***Taxi-Out Time*** (TXOUTTM) refers to the duration in minutes from gate departure to wheels-off times.

The dependent variables except block buffer represent some key flight operations likely to impact block time adversely as they increase. As taxi time increases, take-off is delayed; the airport may experience congestion, and block times may increase.

Quantile Regression

Quantile regression is a type of regression that makes it possible to study the relationship between an independent and dependent variables at different percentiles of the dependent variable distribution. This is all the more important as the distribution is skewed. Quantile regression features several advantages compared with the traditional ordinary-least-square (OLS) regression in assessing the influence of selected operational factors on the variations of block time at various locations of its distribution:

- Quantile regression specifies the conditional quantile function. It permits the analysis of the full conditional distributional properties of block delays as opposed to OLS regression models that focus on the mean.
- It defines functional relations between variables for all portions of a probability distribution. Quantile regression can improve the predictive relationship between block times and selected variables by focusing on quantiles instead of the mean. As Hao and Naiman (2007) pointed out, "While the linear regression model specifies the changes in the conditional mean of the dependent variable associated with a change in the covariates, the quantile regression model specifies changes in the conditional quantile."
- It determines the effect of explanatory variables on the central or non-central location, scale, and shape of the distribution of block times.
- It is distribution-free, which allows the study of extreme quantiles. Outliers influence the length of the right tail and make average block time irrelevant as a standard for identifying the best-possible block time. A single rate of change characterized by the slope of the OLS regression line cannot be representative of the relationship between an independent variable and the entire distribution of block time. In the quantile regression, the estimates represent the rates of change

conditional on adjusting for the effects of the other model variables at a specified percentile. Therefore, the skewed distribution of block times calls for a more robust regression method that takes into account outliers or the lack of sufficient data at a particular percentile (especially at the extremes of the distribution) and generates different slopes for different quantiles.

OUTCOMES AND IMPLICATIONS

The estimates as well as the key regression outputs at the 5th, 25th, median, 75th, and 95th percentile are summarized in Table 2. The 50th quantile estimates can be used to track changes in the location of the median from the lowest to highest observed values of block times. According to Hao and Naiman (2007), the 5th and 95th percentiles "can be used to assess how a covariate predicts the conditional off-central locations as well as shape shifts of the response." The shape shift refers to a movement of the mean (location on the X-axis) due to the presence of outliers. Based on the graphs in Appendix 1, the coefficient estimates show a positive relationship between the quantile value and the estimated coefficients at higher percentiles for scheduled block times, taxi out times, and airborne delay.

If we take the example of the 50th percentile in summer 2011, the quantile regression model for at $\tau = 0.50$ (50th percentile) is as follows:

$$\text{(1)}\quad \text{Block Time}_{\tau = 0.50} = -0.9105*X_{BLKBUFFER} + 0.8888*X_{SCHEDBLKTM} - 0.3090*X_{DEPDEL}$$

$$+ 0.2702*X_{ARRDEL} + 1.1015*X_{AIRBNDEL} + 1.1372*X_{TXOUTTM} + \varepsilon$$

In equation (1), 1.1372 represents the change in the median of block time between SEA and OAK corresponding to a one minute change in taxi-out time at SEA. Since the p value is zero, we reject the null hypothesis, at a 95% confidence level, that taxi-out times at SEA have no effect on the median block time between SEA and OAK in summer 2011. The pseudo coefficient of determination is a goodness-of-fit measure.[9]

Overall, summer 2011 is the only period when all the independent variables have a significant effect on block times at all the considered percentiles. Remarkably, block buffer, scheduled block time, departure, arrival and airborne delays, as well as taxi-out times are significant at the 95th percentile, at a 95% confidence level, in summer 2011, 2010, 2004, and 2000. This suggests that the difference between actual and planned departure and arrival times are more likely to have an incidence on the conditional mean of block times at the highest percentile as a result of taxi-out delays and surface area movement congestion. Moreover, the magnitude of block buffer and departure delays have a negative impact on the conditional mean of block time for all samples at all selected percentiles. This calls for airline schedulers to understand the reasons for the gap between planned and actual block time and for airport analysts to evaluate the times and conditions when departure operations are delayed.

The results imply that arrival and departure delays have a significant impact on block times at the 95th percentile for all years and at all the percentiles in 2011. Arrival and departure delay imply that an aircraft departed or arrived later than the time filed by the pilot in the last flight plan before take-off. This may be due to ground stops when traffic volume or weather requires departures to be delayed. While the average minutes of EDCT were zero in 2000 and 2004, they increased from 94 in 2010 to 104 in 2011 (Table 1). However, the magnitude of the sparsity value is important to evaluate the relevance of the impact of the independent variables on block time. Sparsity refers to the density of data at a given percentile level. A low value indicates that there are many observations near the quantile. For instance, there were few observations around the 5th percentile (12.64) in summer 2011 than at the other percentiles: The sparsity values were 2.40, 2.27, 2.69, and 4.88 respectively for the 25th, 50th, 75th, and 95th percentiles.

Table 2 shows that 95% of the distribution of block times between SEA and OAK was below 129.14 minutes in the June-to-August time period based across the samples. While the standard distribution is appropriate to measure the spread of a symmetric distribution, interquartile ranges are more indicative of spread changes in skewed distributions. One benefit of quantile regression is that it facilitates the evaluation of scale and magnitude changes across samples and percentiles.

In a comparison of summer 2000 with summer 2011, there has been an increase of 2.21 minutes in block times at the 95th percentile. The SEA-OAK city pair has been mainly operated by Southwest Airlines (SWA) and Alaska Airlines (ASA) with a fleet of Boeing 737s. It was not possible to separate the types of aircraft in the ASPM block en route city pair data that were used for this study. However, based on scheduled data in the Official Airline Guide (OAG) and Innovata, the total number of ASA operations declined to 356 in summer 2011 from 693 in summer 2000, while 91 flights were operated by Horizon's Bombardier Q400 (capable of RNAV/RNP-capable) on behalf of ASA. Nevertheless, ASA operated larger capacity aircraft such as the dash 400, 800, and 900 series, while SWA utilized a combination of dash 300, 500, and 700 aircraft.

The reason for the increase in block time may be attributed to airlines' corporate policy to slow down aircraft speed in order to save on fuel costs and not to large differences in aircraft type (turboprop versus jet aircraft). Based on schedule data, Horizon's Q400's represented a small proportion of the overall traffic between SEA and OAK. Weather conditions characterized by the percentage of operations in instrument meteorological conditions (IMC) did not vary substantially at OAK: It was, respectively, 29.78, 29.86, 29.71, and 29.49% in summer 2011, 2010, 2004, and 2000. At SEA, the percentages were, respectively, 22.83, 30.29, 8.33, and 10.58% during the same time periods (Table 1). Finally, although the use of RNAV/RNP may lengthen the path to the runway, it plays an instrumental role in de-conflicting approaches and departures at neighboring airports, thus minimizing their mutual impact on block times.

CONCLUSION

Predictability is a key performance area identified by the International Civil Aviation Organization. Moreover, it is a cornerstone of the Next Generation of Air Transport System (NextGen) initiatives in the U.S. to ensure the transition from an air traffic controlled to a more air traffic managed environment. As air transportation regulators are under public pressure to crack down on tarmac and other types of delays, it has become imperative for airline schedulers to evaluate models that reflect the influence of key operational variables on actual performance. The complexity of the air traffic system, the inability of airline schedulers to fully anticipate both airport and en route congestion, and the imbalance between travel demand and capacity that results in delay all make it more significant for aviation practitioners to assess the impact of operational variables at different locations of the distribution of block times.

Based on the analysis of the SEA-OAK city pair case study, this article showed how quantile regression can help aviation practitioners develop more robust schedules. First, it enables aviation analysts to consider the impact of explanatory variables at different locations of the distribution of block times. Secondly, the significance of the selected variables and the strength of the impact of selected independent variables on block times make it possible to assess the probability that gate-to-gate operations are likely to reach a specific duration. This is made possible by looking at the conditional mean in the case of quantile regression as opposed to the mean of the distribution of block times in the case of OLS models. Thirdly, quantile regression makes it easier to evaluate the scale and magnitude of changes across specific percentiles over a sample. Finally, quantile regression can help analysts study the impact of explanatory variables from different perspectives. In the present case, the quantile regression models focused on constraining factors such as airborne, departure, and arrival delays on the conditional means of block times, which explains the identification of the 95th

Table 2: The Quantile Regression Outputs

Alpha = .95	2011		2010		2004		2000	
	Coefficient	Probability	Coefficient	Probability	Coefficient	Probability	Coefficient	Probability
5th Percentile								
BLKBUFFER	-0.7594	0.0000	-1.0083	0.0000	-0.9693	0.0000	-0.9904	0.0000
SCHEDBLKTM	0.9263	0.0000	0.9006	0.0000	0.9316	0.0000	0.9051	0.0000
DEPDEL	-0.7318	0.0000	-0.0404	0.3929	-0.0448	0.5115	-0.0616	0.5885
ARRDEL	0.7623	0.0001	0.0788	0.1598	0.0621	0.4139	-0.0131	0.9229
AIRBNDEL	0.7751	0.0000	1.0951	0.0000	0.8766	0.0000	1.2568	0.0000
TXOUTTM	0.5935	0.0127	1.0119	0.0000	0.9374	0.0000	1.0042	0.0000
Pseudo R-squared	**0.7014**		**0.8893**		**0.8878**		**0.8758**	
Adjusted R-squared	0.6841		0.8829		0.8813		0.8686	
S.E. of regression	2.4713		1.0580		1.0013		1.3161	
Quantile dependent var	112.8600		107.5600		113.0000		110.9400	
Sparsity	12.6438		3.4399		2.8121		4.0616	
25th Percentile								
BLKBUFFER	-0.8525	0.0000	-1.0084	0.0000	-0.9134	0.0000	-0.8522	0.0000
SCHEDBLKTM	0.8916	0.0000	0.9066	0.0000	0.9253	0.0000	0.9230	0.0000
DEPDEL	-0.4393	0.0000	-0.0439	0.4714	-0.1207	0.1369	-0.1860	0.0400
ARRDEL	0.3976	0.0000	0.0411	0.6096	0.1134	0.1995	0.1727	0.0665
AIRBNDEL	1.1222	0.0000	1.0805	0.0000	0.9252	0.0000	0.9989	0.0000
TXOUTTM	1.0369	0.0000	1.0048	0.0000	0.9607	0.0000	0.8326	0.0000
Pseudo R-squared	**0.7694**		**0.8952**		**0.8753**		**0.8843**	
Adjusted R-squared	0.7560		0.8891		0.8680		0.8776	
S.E. of regression	1.2083		0.7264		0.7354		0.7577	
Quantile dependent var	117.7100		115.0000		116.5300		117.2100	
Sparsity	2.4094		1.8315		1.5958		1.8171	
50th Percentile								
BLKBUFFER	-0.9105	0.0000	-0.9907	0.0000	-0.9152	0.0000	-0.7903	0.0000
SCHEDBLKTM	0.8888	0.0000	0.9074	0.0000	0.9383	0.0000	0.9285	0.0000
DEPDEL	-0.3090	0.0275	-0.0187	0.8068	-0.1183	0.1299	-0.2783	0.0260
ARRDEL	0.2702	0.0398	-0.0170	0.8388	0.1205	0.1602	0.2709	0.0433
AIRBNDEL	1.1015	0.0000	0.9973	0.0000	0.7753	0.0000	0.9507	0.0000
TXOUTTM	1.1372	0.0000	1.0547	0.0000	0.9383	0.0000	0.7760	0.0000
Pseudo R-squared	**0.7994**		**0.8922**		**0.8631**		**0.8689**	
Adjusted R-squared	0.7877		0.8859		0.8551		0.8613	
S.E. of regression	1.1793		0.6220		0.6015		0.6263	
Quantile dependent var	120.4300		119.0000		118.9300		119.8600	
Sparsity	2.2739		1.4886		1.5010		1.7504	
75th Percentile								
BLKBUFFER	-0.8864	0.0000	-0.9900	0.0000	-0.9185	0.0000	-0.7780	0.0000
SCHEDBLKTM	0.9255	0.0000	0.9140	0.0000	0.9344	0.0000	0.9385	0.0000
DEPDEL	-0.3590	0.0340	-0.0885	0.3238	-0.0900	0.2008	-0.2509	0.0089
ARRDEL	0.3720	0.0402	0.0561	0.5520	0.1094	0.1546	0.2650	0.0144
AIRBNDEL	0.9501	0.0000	0.9861	0.0000	0.8565	0.0000	0.8778	0.0000
TXOUTTM	0.8492	0.0000	1.0249	0.0000	0.9496	0.0000	0.7251	0.0000
Pseudo R-squared	**0.8138**		**0.8954**		**0.8548**		**0.8695**	
Adjusted R-squared	0.8030		0.8893		0.8464		0.8619	
S.E. of regression	1.4336		0.7203		0.6748		0.7868	
Quantile dependent var	123.8000		121.9200		121.6000		122.1500	
Sparsity	2.6969		1.6380		1.9727		2.0322	
95th Percentile								
BLKBUFFER	-0.8970	0.0000	-0.9377	0.0000	-0.7299	0.0000	-0.6846	0.0000
SCHEDBLKTM	0.9308	0.0000	0.9097	0.0000	0.9750	0.0000	0.9456	0.0000
DEPDEL	-0.3469	0.0108	-0.2347	0.0023	-0.3725	0.0002	-0.3753	0.0000
ARRDEL	0.4219	0.0047	0.2056	0.0127	0.3844	0.0001	0.4111	0.0000
AIRBNDEL	0.6948	0.0000	1.0246	0.0000	0.6003	0.0000	0.7797	0.0000
TXOUTTM	0.9168	0.0000	1.0483	0.0000	0.5678	0.0000	0.6431	0.0000
Pseudo R-squared	**0.8430**		**0.8986**		**0.8761**		**0.9103**	
Adjusted R-squared	0.8339		0.8927		0.8689		0.9051	
S.E. of regression	1.9295		1.0992		1.2589		1.1371	
Quantile dependent var	129.1400		127.4300		126.0600		126.9300	
Sparsity	4.8819		3.2418		3.6042		3.4603	

Not significant at $\alpha = .95$

percentile optimal values as the expected block time given the impact of the explanatory variables and the strength of the pseudo R-square.

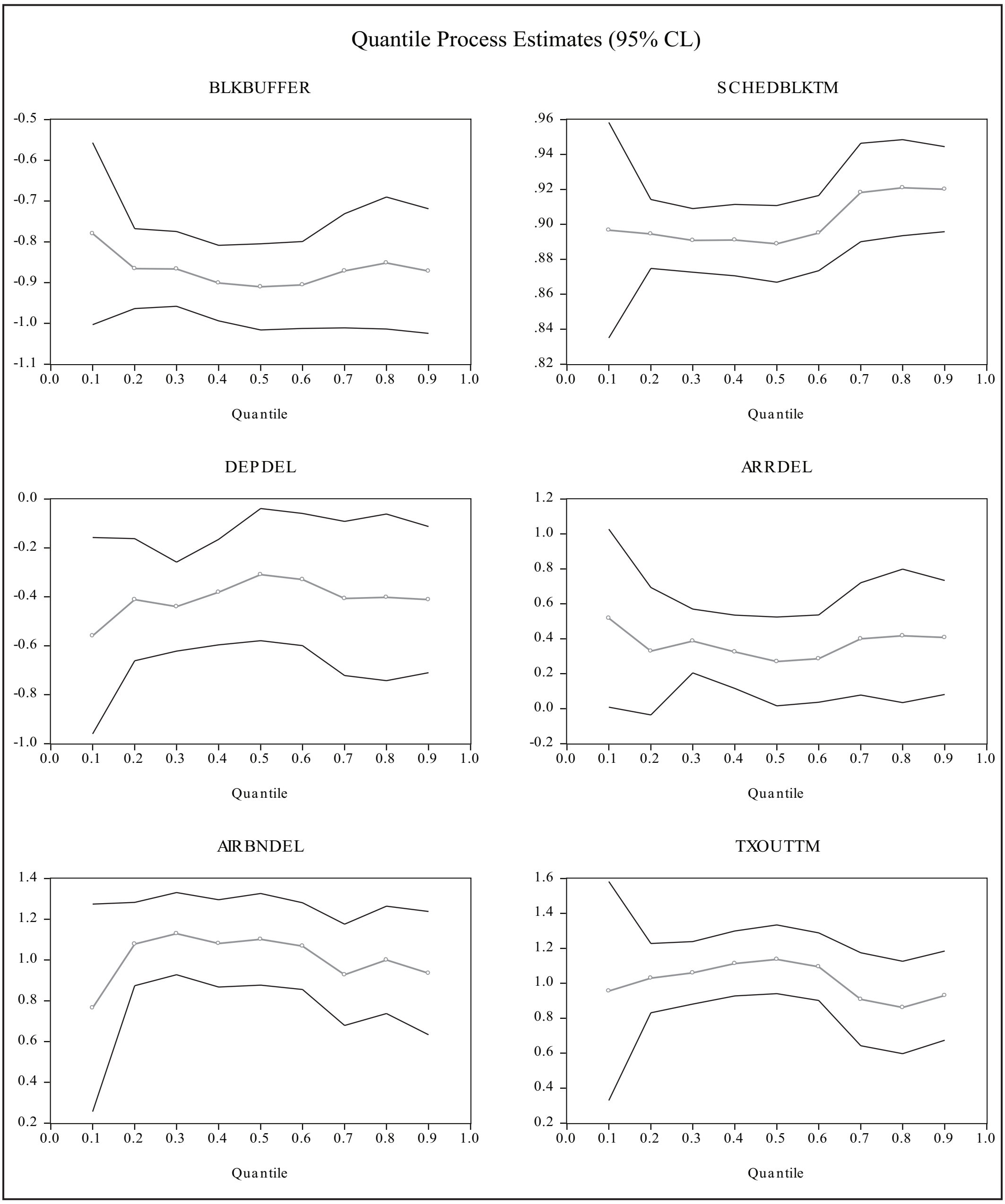
The results suggest that the impact of the dependent variables on block time is holding more consistently at the higher percentile. As a measure of gate-to-gate performance, block time depends on the variability of taxi times. Traffic volume and weather events that can trigger ground delay programs make it more difficult for airlines to predict block times. In fact, the average Expected Departure Clearance Time (EDCT) for the arrivals having an EDCT delay increased from 94 to 104 minutes between summer 2010 and 2011. As the percentage of operations in instrument meteorological conditions did not vary at OAK, it was more variable at SEA. Moreover, while the use of precision approach to OAK has made it possible for RNAV/RNP capable aircraft to avoid airspace congestion around San Francisco, it is likely to increase flight time as aircraft must follow a specific path into OAK.

Although the quantile regression models could provide some indication as to the scale and magnitude of change, they would have benefited by measuring the impact of technical changes introduced by NextGen programs and initiatives between summer 2000 and summer 2011. However, the assessment of such changes requires that available surveillance data keep track of the use of procedures such as RNAV/RNP, optimal profile descents, among others.

Abbreviations

ARINC	Aeronautical Radio, Inc.
ASA	Alaska Airlines
AIC	Akaike Information Criterion
ASPM	Aviation System Performance Metrics
ASQP	Airline Service Quality Performance
ATM	Air Traffic Management
BIC	Bayesian Information Criterion
BTS	Bureau of Transportation Statistics
ICAO	International Civil Aviation Organization
ILS	Instrument Landing System
IMC	Instrument Meteorological Conditions
EDCT	Expected Departure Clearance Time
FAA	Federal Aviation Administration
FAM	Fleet Assignment Models
FL	Flight Level
MSL	Mean Sea Level
NAS	National Airspace System
NextGen	U.S. Generation Air Transportation System
OAG	Official Airline Guide
OLS	Ordinary Least Squares
PBN	Performance-Based Navigation
RNAV	Area Navigation
RNP	Required Navigation Performance
SWA	Southwest Airlines
TFMS	Traffic Flow Management System
VMC	Visual Meteorological Conditions

APPENDIX 1: QUANTILE PROCESS ESTIMATE GRAPHS (95% CONFIDENCE LEVEL)



The graphs show the 95% confidence intervals around the regression coefficients listed in Table 2 on the Y-axis and the different quantiles on the X-axis in the case of the June to August 2000 sample.

Endnotes

1. On-time performance and the causes of delay are reported by major carriers in the Airline Service Quality Performance report. There are five categories of delay: air carrier, extreme weather, National Airspace System, late arriving aircraft, and security. The information is available at http://www.bts.gov.

2. Hof, J. "Development of a Performance Framework in Support of the Operational Concept," *ICAO Mid Region Global ATM Operational Concept Training Seminar*, Cairo, Egypt, November 28–December 1, (2005): 36.

3. VHF Omni directional radio range (VOR) is a short-range radio navigation system that allows an aircraft to determine its position.

4. AIRINC stands for Aeronautical Radio, Inc. (http://www.arinc.com).

5. The TFMS (formerly ETMS) and ARINC data, as well as the ASPM delay metrics, are available at http://aspm.faa.gov.

6. The Akaike Information Criterion is defined as 2k – 2 ln(L) where k is the number of parameters and L the maximized value of the likelihood function for the estimated model.

7. The Bayesian Information Criterion is -2 ln(L) + k*ln(n) where n is the number of observations.

8. Block buffer in this paper is determined by the difference between block time based on the last flight plan before takeoff and optimal block time. Therefore, the relationship between block buffer and arrival delay is not as strong as if schedules were used as a benchmark.

9. See Koenker and Machado (1999) for further explanations.

References

Abara, J. "Applying Integer Linear Programming to the Fleet Assignment Problem." *Interfaces* 19 (4), (1989): 20-28.

Abdelghany, F. and K. Abdelghany. *Modeling Applications in the Airline Industry*. Ashgate Publishing Company, Burlington, Vermont, 2009.

Cook, A. *European Air Traffic Management: Principles, Practices, and Research*. Ashgate Publishing Company, Burlington, Vermont, 2007.

Donohue, G.L., A. Zellweger, H. Rediess, and C. Pusch. *Air Transportation System Engineering: Progress in Astronautics and Aeronautics*. American Institute of Aeronautics and Astronautics, Danvers, Massachusetts, 2001.

Federal Aviation Administration, *Aviation System Performance Metrics*, http://aspm.faa.gov.

Federal Aviation Administration, *Traffic Flow Management System*, https://aspm.faa.gov/tfms/sys/main.asp.

Hane, C.A., C. Barnhart, E.L. Johnson, R.E. Marsten, G.L. Nemhauser, and G. Sigismondi. "The Fleet Assignment Problem: Solving a Large Scale Integer Programming." *Mathematical Programming* 70 (2), (1995): 211-232.

Hao, L. and D. Q. Naiman. *Quantile Regression*. Sage Publications, Thousand Oaks, California, 2007.

Koenker, R. and J. Machado. "Goodness of Fit and Related Inference Processes for Quantile Regression." *Journal of the American Statistical Association* 94 (448), (1999): 1296-1310.

Lan, S., J.P. Clarke, and C. Barnhart. "Planning for Robust Airline Operations: Optimizing Aircraft Routing and Flight Departure Times to Minimize Passenger Disruptions." *Transportation Science* 40 (1), (2006): 15-28.

Lohatepanont, M. and C. Barnhart. "Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment." *Transportation Science* 38 (1), (2004): 19–32.

Morissett, T. and A. Odoni. "Capacity, Delay, and Schedule Reliability at Major Airports in Europe and the United States." *Transportation Research Record: Journal of the Transportation Research Board* 2214, (2011): 85-93.

Rapajic, J. *Beyond Airline Disruptions*. Ashgate Publishing Company, Burlington, Vermont, 2009.

Wu, C.L. *Airline Operations and Delay Management: Insights from Airline Economics, Networks and Strategic Schedule Planning*. Ashgate Publishing Company, Burlington, Vermont, 2010.

Wu, C.L. "Inherent Delays and Operational Reliability of Airline Schedules." *Journal of Air Transport Management* 11 (4), (2005): 273-282.

***Tony Diana** is the division manager, NextGen Performance at the U.S. Federal Aviation Administration. He is involved in the measurement of operational outcomes resulting from the implementation of NextGen initiatives at U.S. airports, metroplexes, and airspaces. Prior to that position, Diana was deputy division manager, Forecasting and Performance Analysis, in the Office of Aviation Policy and Plans of the FAA, where he managed the Aviation System Performance Metrics data warehouse. Diana's main interest is performance evaluation, airport performance benchmarking, and the study of delay.*

A Dynamic Programming Optimization Approach for Budget Allocation to Early Right-of-Way Acquisitions

by Carlos M. Chang Albitres, Paul E. Krugler, Iraki Ibarra, and Edith Montes

Maximizing potential savings when purchasing right-of-way within a limited budget is a challenge currently faced by state departments of transportation (DOTs) across the nation. Early right-of-way acquisitions promote smoother negotiations and are aimed to save money, time, and human resources. This paper describes an optimization approach based on dynamic programming developed for the Texas Department of Transportation (TxDOT) to identify projects with candidate parcels for early right-of-way acquisition in order to achieve the highest potential savings. Each candidate parcel must be subjected to a preliminary environmental analysis to ensure that each comply with the National Environmental Policy Act (NEPA) standards.

INTRODUCTION

Early right-of-way acquisitions promote smoother negotiations and are aimed at saving money, time, and human resources. Property improvements, speculation, and damages to remainders of properties are some of the major factors that increase the cost of right-of-way acquisitions. The Texas Department of Transportation (TxDOT) found that property costs double, triple, or more after property improvements, speculation, and damages to remainders of properties occurs. Some of the methods used to increase the cost of the properties include the subdivision of the property to sell more units at higher per unit cost and the development of the property to add more value to it (TxDOT 2009).

Early right-of-way acquisitions are intended to improve project efficiency and promote smooth negotiations; although sometimes condemnation cannot be avoided to acquire the land, causing a delay in the project and rougher negotiations with the owners. However, condemnation rates can be decreased by improving the valuation and negotiation processes (Caldas et al. 2011). Early right-of-way is an option to consider due to anticipated property cost increases; however, its use is constrained by laws and policies. Not all parcels are candidates for early right-of-way acquisition and each candidate parcel must be subjected to a preliminary environmental analysis to ensure compliance with the National Environmental Policy Act (NEPA) standards.

Purchase of right-of-way transportation projects must consider the environmental, social, economic, and political aspects of each project. This paper provides an insight into the parcel selection process to start early negotiations in order to maximize DOTs savings. A dynamic programming optimization method was applied to optimize the selection of the best candidate parcels with the highest rate of returns from early acquisition. A software called Early Right-of-Way (EROW) is the decision tool used to support this concept. EROW analyzes the optimal projects needed to be purchased in early right-of-way acquisition in a given scenario. The model used by EROW has the ability to input the project's costs, expected savings, and available budget for early right-of-way acquisitions. Savings are defined as the difference in budget between the early right-of-way scenario and the traditional right-of-way acquisition process. Each parcel included in the candidate list for early right-of-way acquisition must comply with NEPA's standards; this should be analyzed by performing a preliminary environmental analysis prior to the selection of candidate projects. The model identifies the parcels that may lead to higher savings; the final selection of projects should always be performed by the decision maker based on the individual characteristics of each project.

BACKGROUND

The process of acquiring right-of-way for a construction project can considerably impact the project's completion time and overall cost. The degree of impact to these important considerations fluctuates due to the number of variables associated with the project and individual parcel characteristics. One of the ways to reduce potential negative impact on time and cost from the right-of-way acquisition portion of the project is to acquire portions of the right-of-way early. U.S. DOTs have limited authority to acquire right-of-way prior to completion and approval of the environmental studies, which is the conventional time for the Right of Way Division (ROW Division) to issue a right-of-way release to the district office to begin purchase negotiations (TxDOT 2009). In addition to the challenges associated with limited ability to purchasing parcels early in the project planning process, no systematic analytical approach is available to decide when to allocate funds for early right-of-way acquisition. However, experience has shown that waiting until normally required steps in project planning have been completed, before acquiring certain right-of-way parcels, can result in substantially increased costs. TxDOT research project 0-5534 (2009) documented some examples of increase in land costs:

- Example 1: A parcel cost estimate was $0.5 million as pasture land in 2001. This property was purchased for $3.3 million in 2004 as residential property. The district estimated that the cost could have gone up to $6 million if purchasing had been delayed further.
- Example 2: A parcel cost increased from $0.17/sq-foot up to $0.23/sq-foot in three years.
- Example 3: A parcel cost for pasture land went from $7,000 per acre to $22,000 per acre in three years.
- Example 4: Total parcel costs for a group of parcels went from an estimated $5.4 million to $10 million when not acquired early.
- Example 5: Total parcel costs for a group of parcels went from an estimated $5 million to $15 million (Krugler et al. 2010).

It was concluded that property improvements, speculation, and damages to remainders of properties are some of the major factors that contribute to increases in land costs (TxDOT 2009). For this reason, the valuations of properties and the negotiations with property owners are two important aspects in the right-of-way acquisition process for transportation projects. If the valuation and negotiation processes are improved, the overall project delivery efficiency can increase due to a decrease in condemnation rates (Caldas et al. 2011).

LITERATURE REVIEW

Efficient allocation of resources is a critical component of successful transportation asset management practice. Since optimization is a mathematical approach that minimizes cost or maximizes benefit while satisfying pre-given constraints, it is adopted for many transportation problems including, the capital budgeting allocation problem. There are some publications on funding allocation, but none regarding the application of dynamic programming to right-of-way acquisitions. Armstrong and Cook (1979) developed a model for a single-year planning period. In this model, the objective was to maximize the total benefit from the highway subjected to fixed budget constraints. Later it was expanded to consider multiple planning years by using a financial planning model and a goal programming approach (Cook 1984). In contrast to maximizing benefit, another approach is to seek a solution minimizing total project costs. Davis and Van Dine (1988) developed a computer model to minimize user costs subject to budget and production capacity constraints for optimizing maintenance and reconstruction activities. They used linear programming formulation as an optimization technique. More recently, advanced computing power allows optimization techniques to solve more realistic and sophisticated pavement management problems, which is a part of a larger decision-support system. Ferreira et al. (2002) formulated a mixed integer optimization model for network-level Pavement Management Systems (PMS). They used genetic-algorithm heuristics to solve the optimization problem, minimizing the expected total discounted costs of pavement maintenance and rehabilitation actions over a planning period. Heuristic methods are used in optimization problems in order to provide an approximate answer to the problem when the optimal solution is very difficult to find. A Genetic Algorithm (GA) is an adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics. The GA exploits historical information to direct the search into the region of better performance within the search space. The basic techniques of the GAs are designed to simulate processes in natural systems necessary for evolution, especially those that follow the principles first laid down by Charles Darwin of "survival of the fittest," i.e. optimal answers. Wang et al. (2003) also used genetic-algorithm heuristics to solve the zero-one integer programming formulation, which is a special optimization case where the problem's solution is required to be 0 or 1.

As described before, resource allocation problems are among the classical applications of optimization techniques. However, the complexity of real-world problems associated with resource allocation in right-of-way acquisitions limits the applicability of classical methods. Few research efforts were found for right-of-way acquisition. A research report published by the Minnesota Department of Transportation addresses the question of whether there are financial benefits to acquiring transportation right-of-way far in advance of when the improvement will be done (MNDOT 2005). The main findings of this report suggest that early acquisition is recommended under the following conditions:

- Developed land that is in danger of being redeveloped to a higher value
- Developed land that is being offered for sale voluntarily
- Undeveloped land that is near developing areas or in desirable recreation areas and may appreciate rapidly enough to justify early acquisition
- Undeveloped land that is expected to be developed
- Land along major transportation corridors that may appreciate more rapidly than land in general in the vicinity

Hedonic price models to estimate the cost of right-of-way acquisition were proposed in Texas taking into consideration data from several corridors and commercial sales transactions in the state's largest regions. The findings showed that the value of improvements were more important to costs than the size of the parcels, and that utility costs were highly variable (Heiner and Kockelman 2005). Furthermore, an electronic appraisal methodology for right-of-way acquisition in highway projects has been suggested in order to improve the appraisal process and reduce the likelihood of inconsistent appraisal values (Caldas et al. 2012). Before any appraisal valuation takes place, DOTs must select the parcels that will be considered in the right-of-way acquisition process. Due to limited funding for ROW acquisition of transportation projects, decision makers must ensure that the projects selected for early right-of-way acquisition are the most cost efficient projects, represented by higher rate of return.

No optimization techniques for early right-of way allocation are mentioned in these reports. A dynamic programming optimization approach to optimize right-of-way acquisitions to maximize savings is presented in this paper. Dynamic programming is a method to solve complex problems by breaking them down into simpler interrelated sub-problems that affect the overall decision making process. The approach is based on a comparison of anticipated rates of returns from acquiring right-of-way at a certain time to a rate of return expected from other potential uses of the evaluated budgetary amount.

BUDGET ALLOCATION FOR EARLY RIGHT-OF-WAY ACQUISITION

Early acquisition of right-of-way is defined as the purchase of parcels before the approval of the environmental study. Early right-of-way acquisitions foster smoother negotiations, and save money, time, and human resources (TxDOT 2009). Projects considered in the early acquisition must not represent a hard case in environmental approval; this can be analyzed by performing a preliminary environmental analysis of the candidate projects under the National Environmental Policy Act (NEPA) regulations. It ensures that federal agencies consider the effect of their actions on the quality of the human environment and sets standards that must be met by these agencies in order to receive federal funding (NEPA 2012). In the process of early acquisition, because no environmental approval has been made, federal funding cannot be used to buy these properties. Therefore, the state DOT must provide its own money to buy the parcels selected in early acquisition and must wait for the environmental approval in order to receive federal reimbursement. All projects bought during the early or normal right-of-way acquisition must comply with NEPA's standards prior to receiving federal funding or reimbursement. Early right-of-way acquisitions are performed only under the following cases: to alleviate a particular hardship to the property owner, to prevent imminent parcel development, and as donations. Early acquisition does not avoid the environmental review of a project, influence a decision regarding the need to construct the project, or the selection of a specific location (TxDOT 2011).

A significant amount of cost savings are expected when parcels are purchased early in the right-of-way acquisition process. If a good property valuation and a good negotiation with the property owner take place during the early right-of-way acquisition, the likelihood of taking the parcel by condemnation or of delaying the project will be reduced due to the successful purchase (Hakimi and Kockelman 2005). Experiences of TxDOT show that cost savings could easily double the funds invested in purchasing selected parcels earlier when compared with the normal procurement process (Krugler et al. 2010).

Expected costs and savings resulting from early acquisition of a specified number of parcels included in a given right-of-way project is usually estimated by experienced personnel in the state DOTs. Knowledge of the characteristics associated with each individual parcel is essential for a good estimate. Among these characteristics are: location of the parcel, type of right-of-way acquisition process (either through negotiation or condemnation), time at which the parcel is acquired, ownership of the parcel, and likelihood of future improvements made on the parcel before acquisition. The cost estimation process is quite difficult due to the complex interaction among the factors that affect acquisition of the parcels, especially when parcel cost increases could be driven up by land speculation, a fairly common situation where transportation improvements are planned. These factors were considered in the study and included when calculating costs and potential savings.

When funds available for projects are constrained, senior managers in state DOTs face a difficult situation, making hard decisions on where and how to invest the limited budget. The levels of expertise of managers as well as tools available to conduct funding allocation are crucial to stretch the budget. In this situation, optimization methods can assist managers in selecting the set of projects with the highest rate of return. The rate of return (ROR) of an investment is the interest rate at which the net present value (NPV) of costs, i.e., negative cash flows, equals the NPV of the benefits, i.e., positive cash flows in the investment. In this case study, ROR was calculated as the ratio between the expected savings to the expected costs of a given project.

Numerous projects and parcel options are taken into consideration from a statewide perspective, and to obtain a broader based optimal solution, all possible scenarios must be included in a system-level optimization model, to ensure that at most only one of the scenarios corresponding to each critical project will be selected. The problem becomes more complex when the number of projects is large and budget constraints are taken into account. In fact, the resulting combinatorial optimization problem is known as NP-hard, meaning that the problem is computationally intractable in theory. Despite this discouraging theoretical property, the problem can be solved in practice by using pseudo-polynomial dynamic programming (DP) algorithms to output cost savings for all possible budget scenarios. Dynamic programming is a method to solve complex problems by breaking them down into simpler sub problems. It solves different parts of the problems, called sub problems, and then combines the solutions of the sub problems to get an overall solution. Furthermore, the use of pseudo-polynomial algorithms simplifies the problem even more by only storing and calculating the values that meet all the criteria, reducing running time.

DYNAMIC PROGRAMMING OPTIMIZATION FOR EARLY RIGHT-OF-WAY ACQUISITION

Dynamic programming (DP) is a methodology used to solve a large-scale problem by breaking it down into a number of smaller problems, which can then be solved in a recursive fashion. The nature of the early right-of-way acquisition process fits very well with the use of DP algorithms, which share the following common characteristics:

- The early right-of-way acquisition problem can be subdivided into separate stages.
- There are a number of states and possible early right-of-way scenarios associated with each stage. The division of sequence of an optimization problem into various subparts is called stages. A state is a measurable condition of the system; in this problem, a state represents the budget available for project funding.
- At each stage of the right-of-way acquisition process, a decision is made to move from the current to the following stage.
- An optimal decision at each state does not depend on the previous decisions or states since the right-of-way acquisition process is unique and depends on each individual parcel characteristics as discussed before (principle of optimality).
- There is a recursive relationship representing the optimal decision for each state of the right-of-way acquisition process at stage i in terms of previously computed optimal decisions for states at subsequent stages $i+1$, $i+2$, … (for backward recursion). Backward recursion is used to determine the list of optimal actions needed to solve the problem using a backwards sequence. It considers the last time a decision is made and computes its possible outcomes; then, it analyzes the second-to-last decision in the same manner. This process continues until every possible scenario is analyzed.

The resource allocation problem being dealt with in right-of-way funding allocation can be viewed as a modified version of the knapsack problem, in which a given set of N items, in our case projects, can be placed in a knapsack of a certain budget capacity B. Each project i is characterized by its cost c_i and its savings s_i. Savings is defined as the monetary value that the agency might save if a given parcel is bought during early right-of-way as compared with a scenario where no early right-of-way parcels are bought. Savings is an input that must be given to the EROW software. It can be calculated based on expert judgment, simulations, or any other method.

The dynamic programming optimization model was a result of a four-year study sponsored by TxDOT and validated through a one-year implementation project with reasonable results, demonstrating its applicability for early right-of-way acquisition decisions. The objective of the dynamic programming problem is to find the parcels from each project to maximize savings while staying within the budget constraint, while total acquisition costs of selected project scenarios must be less than or equal to the total available budget. The problem is formalized as follows:

(1) Maximize $\sum_{i=1}^{N} \sum_{j \in S_i} s_{ij} x_{ij}$

so that $\sum_{i=1}^{N} \sum_{j \in S_i} c_{ij} x_{ij} \leq B$

$$x_{ij} \in \{0,1\}.$$

where:

N= the total number of projects considered;

i = index representing the project number, $i=1,\dots, N$;

S_i = the set number of considered early acquisition scenarios for project i;

j = index used for the scenarios, $j \in S_i$;

$c_{ij} =$ the cost of project i under scenario j;

$s_{ij} =$ the savings for project i under scenario j;

b = a lower bound on the budget for early right-of-way acquisition (the default value is 0); and

B = the total available budget.

If the binary variable is denoted by x_{ij}, such that $x_{ij}=1$ if the j-th scenario of project i is selected and $x_{ij}=0$ if the j-th scenario of project i is not selected. The DP stages correspond to projects, and the state b_i at stage i will represent the budget available for projects $1\dots i$ using forward recursion. Forward recursion is the method to solve a problem that is decomposed into a series of n stages and analyzes the problem starting with the first stage in the sequence, working forward until the last stage is analyzed. The recursive function $v_i(b_i)$ at stage i expresses the maximum savings resulting from allocating the budget given by b_i among the options available for projects $1, \dots, i$.

Incremental rate of return (IRR) analysis is used to compare investment alternatives to the specified minimum attractive rate of return (MARR). The MARR is the minimum rate of return on a project a manager or company is willing to accept before selecting a project, taking into account the risk and the opportunity cost of not accepting other projects. In this case study, MARR was determined by calculating the average ratio of annual benefits to initial investment of various projects. In order to illustrate this concept, consider an example with two alternatives: Alternative 1 and Alternative 2. Alternative 1 has a base list of given ROW projects and Alternative 2 shows the same list of projects plus some extra projects that might be bought if additional funds were allocated. If the rate of return from just the additional amount of investment being considered is greater than the MARR, then the additional investment is worthwhile, and the second alternative, including the additional investment, is preferred. Otherwise, if the rate of return is less than MARR, the first alternative is more attractive. The incremental rate of return analysis allows comparison of alternative early right-of-way acquisition funding scenarios to determine the optimal budget. Optimal budget is given by the total amount of money needed to buy the ROW projects that will maximize the incremental rate of return.

The decision at stage i is which of the options in S_i , if any, should be chosen so that the optimal savings of $v_i(b_i)$ is achieved at stage i. Next, the recursive relation is defined to compute the values of $v_i(b_i)$. To simplify the formulation, an artificial stage 0 is introduced, which corresponds to having no projects. It is also assumed that the first early acquisition scenario in each S_i corresponds to not picking project i, i.e., $c_{i1} = s_{i1} = 0$. Another important assumption is that only a discrete set of possible budget levels b_i is considered; therefore, discretization is used to transfer the model from continuous to discrete equations to make the model easier to analyze. Notation $b_i = b,\dots, B$ will mean that b_i takes on all possible funding levels (according to the corresponding discretization) between the minimum considered budget b and the maximum considered budget B. This provides the following recursive relation for each $b_i = 0,\dots, B$:

$$(2) \qquad v_i(b_i) = \begin{cases} 0, & if\ i = 0 \\ max\{s_{ij} + v_{i-1}(b_i - c_{ij}) \mid j \in S_i\}, & if\ 0 < i \leq n \end{cases}$$

By solving this recursion for all projects starting from the first project to the last, the computational results for $v_n(b_n), b_n = b, \ldots, B$ will represent the optimal solution for all the considered budget increments. This dynamic optimization model was used to develop a software tool for Early-Right-of-Way Acquisition (EROW). EROW was developed at Texas A&M University to assist TxDOT in selecting projects under different right-of-way acquisition scenarios.

USING EROW TO IDENTIFY EARLY RIGHT-OF-WAY ACQUISITION PROJECTS

EROW input parameters include the maximum and minimum possible early right-of-way acquisition budgets and the budget increment size (interval between the maximum and minimum budgets). The costs and savings for each project scenario are also entered as well as the minimum attractive rate of return. A project scenario is defined by certain number of parcels acquired early.

A one-year implementation project was conducted with the participation of right-of-way personnel in each of the four TxDOT regions. The project objective was to apply EROW to analyze three construction projects with candidate parcels for early right-of-way acquisition. Project 1 is located in a metropolitan county in Houston and contains 28 parcels. Project 2 is located in an urban county in Austin and contains 20 parcels. Project 3 is located in an urban county in Dallas and contains 10 parcels.

EROW was used to compare the three construction projects under 10 possible parcel acquisition scenarios. In this example, only the 10 parcels with the highest rate of return per project were considered. Table 1 shows the costs and savings under each scenario for all three projects. Scenario 1 is when one parcel was acquired early. Scenarios 2 through 10 correspond to two to 10 parcels acquired early—one additional parcel acquired per scenario. Table 1 shows the costs of each scenario, which are merely the sum of the parcels that were considered. The savings shown is the difference in the budget produced by early right-of-way as compared to the case where no early right-of-way is used. Savings are calculated using historical data of previous acquisitions and based on expert feedback. Table 1 shows very high potential savings in the projects of Houston and Dallas because these parcels were mainly commercial and had a high likelihood of improvement and speculation.

Table 1: Data Inputs for Project Costs and Savings Under Different Right-of-Way Acquisition Scenarios

	Number of Parcels Acquired	Project 1-Houston		Project 2-Austin		Project 3-Dallas	
		Cost ($)	Savings ($)	Cost ($)	Savings ($)	Cost ($)	Savings ($)
Scenario 1	1	183,821	171,683	1,539	1,051	16,781	44,653
Scenario 2	2	450,444	418,850	5,292	2,604	108,594	286,675
Scenario 3	3	8,576,594	7,600,182	11,585	4,629	115,371	296,380
Scenario 4	4	9,123,858	8,043,000	16,327	6,029	119,868	302,611
Scenario 5	5	9,573,150	8,405,204	598,716	154,593	132,756	320,446
Scenario 6	6	9,737,632	8,532,623	622,080	161,031	144,755	336,302
Scenario 7	7	9,952,620	8,679,954	635,515	165,187	196,087	395,963
Scenario 8	8	10,120,755	8,792,406	674,702	175,171	197,086	398,280
Scenario 9	9	10,250,046	8,878,657	675,761	175,593	263,432	440,842
Scenario 10	10	11,100,703	9,413,069	1,099,439	270,938	327,902	568,036

Therefore, if acquired early, these parcels would produce more savings as compared with the case where no parcels are bought during early right-of-way acquisition.

Optimal Solutions

Further analysis was conducted to find optimal solutions corresponding to different early right-of-way acquisition budget alternatives. The total budget available was $1,000,000 and the lower bound on the budget was $100,000. The MARR was set by TxDOT at 20% based on the results from a study that concludes that the rate of return for transportation projects is around this range based on the new direct and indirect jobs created and the aggregate annual income increase due to the highway construction and indirect jobs created (Governor's Business Council 2006). EROW performs an IRR analysis to compare each resulting value to the specified MARR. When an IRR for a considered early acquisition budget does not exceed the MARR value, that budget amount is removed from solution consideration. In this analysis, the project scenarios providing the best rate of returns for each possible budget solution are identified. Using the input costs and savings from Table 1, the EROW software was run under the given budget and MARR constraints in order to maximize the IRR. Table 2 shows the rate of return, budget required, and total savings values for budget solutions corresponding to each optimal solution given by the EROW software.

Table 2: Summary of Optimal Solutions at Different Funding Levels

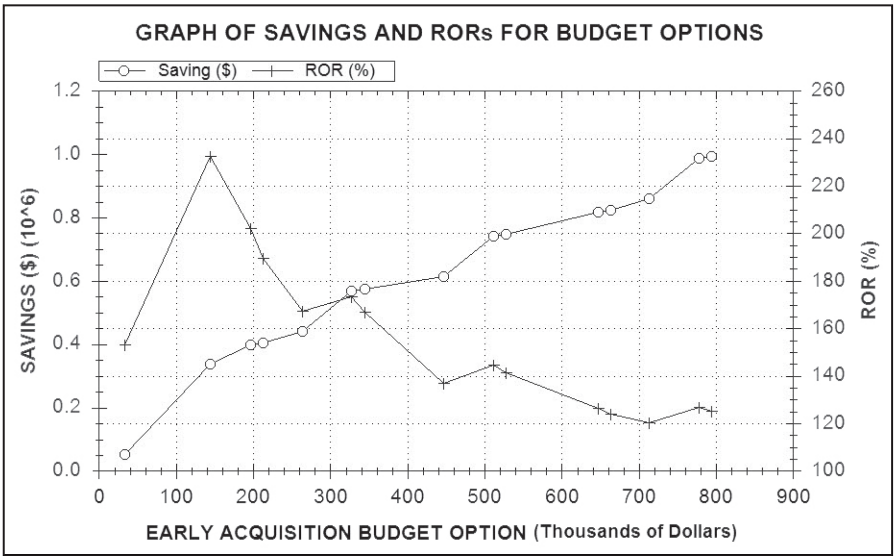
Funding Level ($)	Total Resulting Expenditure ($)	Total Resulting Savings ($)	Rate of Return (%)	Project 1 Houston	Project 2 Austin	Project 3 Dallas
100,000	33,108	50,682	153.08	-	Scenario 4	Scenario 1
150,000	144,755	336,302	232.32	-	-	Scenario 6
200,000	197,086	398,280	202.08	-	-	Scenario 8
250,000	213,413	404,309	189.45	-	Scenario 4	Scenario 8
300,000	263,432	440,842	167.35	-	-	Scenario 9
350,000	327,902	568,036	173.23	-	-	Scenario 10
400,000	344,229	574,065	166.77	-	Scenario 4	Scenario 10
500,000	447,253	612,525	136.95	Scenario 1	-	Scenario 9
550,000	511,723	739,719	144.55	Scenario 1	-	Scenario 10
650,000	528,050	745,748	141.23	Scenario 1	Scenario 4	Scenario 10
700,000	647,530	817,130	126.19	Scenario 2	-	Scenario 8
750,000	663,857	823,159	124.00	Scenario 2	Scenario 4	Scenario 8
800,000	713,876	859,692	120.43	Scenario 2	-	Scenario 9
850,000	778,346	986,886	126.79	Scenario 2	-	Scenario 10
900,000	794,673	992,915	124.95	Scenario 2	Scenario 4	Scenario 10

Table 2 shows the budget required to buy the parcels indicated in each project, i.e., the budget required is the sum of all the scenarios needed in all three projects to obtain the given rate of return under the given funding scenario. Total resulting savings indicates the money that will be saved due to the early acquisition of the parcels indicated in each scenario. Table 2 also shows the rate of return of the investment, which is merely the ratio of total savings to total expenditures of each case. For example, if the funding level is $550,000 then the total expenditure needed to maximize the rate of return is $511,723 and the resulting savings of this purchase is $739,719, which produces a rate of return of 144.55%. In order to have this rate of return, it is needed to buy the parcels on Scenario 1 from Houston and the parcels on Scenario 10 from Dallas.

From Table 2, it is observed that the best rate of return is given by Project 3, Scenario 6 with 232.32%. This rate of return occurs when the funding scenario is $150,000. This means that funding should be allocated primarily to six parcels in the Dallas project. The rates of returns shown are very high because the parcels chosen under each budget scenario were primarily commercial with a high likelihood of improvement and speculation. The budget estimates and funding recommendations based on EROWs were consistent with expert opinions from TxDOT's ROW personnel, including appraisers and planning staff.

Figure 1 shows several budget intervals with peaks that the user may want to analyze in greater detail by decreasing the uncertainty interval for the budget and decreasing the budget increment value.

Figure 1: EROW Output Plot Example



Peaks indicate favored budget solutions from a rate of return standpoint. Wherever the peaks occur, the represented funding scenario offers rate of return advantages over the budgets represented by valleys in the chart. It is observed that the highest rate of return may occur at a funding scenario rather low in the allowable budget range for the analysis. When this occurs, the user may prefer a lower rate of return at a considerably higher early right-of-way acquisition budget. Parcels with rates of return lower than MARR for a parcel should not be considered for early right-of-way acquisition. MARR is determined by the agency based on its own experience, and for TxDOT, the expected MARR is between 20%-25%. The outcome of the analysis is not sensitive to MARR since the IRR for all funding levels is several multiples higher than 20%.

CONCLUDING REMARKS

Efficient allocation of resources is critical for state DOTs, even more when budget limitations are experienced. Early right-of-way acquisition is considered a valid option to obtain substantial cost savings when buying parcels. However, the complexity and variability surrounding the acquisition of right-of-way parcels made this type of analysis extremely difficult and not all projects are qualified to use the early right-of-way acquisition process. In order to qualify for this process, a preliminary environmental analysis of the candidate projects under the NEPA regulations must be performed. Candidate parcels considered for early right-of way acquisition must show compliance with NEPA's standards. Furthermore, early right-of-way parcels must only be bought to alleviate a particular hardship to the property owner, to prevent imminent parcel development, and as donations. Early right-of-way acquisitions are intended to improve project efficiency and negotiations and avoid the condemnation process if possible. The use of a dynamic programming optimization approach allows comparing different right-of-way acquisition scenarios, to optimally allocate funds among candidate projects for a given funding scenario. The dynamic programming optimization model was a result of a four-year study sponsored by TxDOT and validated through a one-year implementation project with reasonable results, demonstrating its applicability for early right-of-way acquisition decisions.

Given the planning time horizon and the right-of-way sites to be acquired, the optimal solutions generated by the model for different budget levels can assist state DOTs make more efficient right-of-way acquisitions decisions through:

- Determination of which right-of-way projects within a district offer the greater potential for substantial savings.
- Determination on a statewide basis of candidate projects and their optimal number of parcels for early right-of-way acquisition.
- A sensitivity analysis varying the input parameters to analyze multiple right-of-way acquisition funding scenarios for different sets of projects providing managers with a broader perspective, and making explicit the previously hidden factors affecting cost savings.
- When selecting which parcels were going to be acquired early, the size of the parcel, the owner (private, public, military, government), project location (metropolitan, urban, rural county), and likelihood of improvements become relevant factors.

The model identifies the candidate parcels that may lead to potential higher savings; such parcels must meet the criteria of early right-of-way acquisitions to be considered in the analysis. A preliminary environmental study must be performed on the candidate parcels before considering them for early right-of-way acquisition. The final selection of projects should always be performed by the decision maker based on the individual characteristics of each project.

References

Armstrong, R. D. and W. D. Cook. "Capital Budgeting Problems in Pavement Maintenance." *International Journal of Urban System* 4, (1979): 175–181.

Caldas, C.H., Z. Zhang, K.M. Kockelman, K.R. Persad, R.A. Halabi, and E. Kincaid. "Development of Best Practices for Right-of-Way Valuation and Negotiations in Transportation Projects." *Journal of the Transportation Research Forum* 50 (3), (2011): 23-41.

Caldas, C.H., Z. Zhang, R.A. Halabi, and E. Kincaid. "Electronic Appraisal Methodology for Right-of-Way Acquisition in Highway Projects." *Journal of the Transportation Research Forum* 51 (1), (2012): 19-34.

Cook, W.D. "Goal Programming and Financial Planning Models for Highway Rehabilitation." *Journal of Operations Research Society* 35 (3), (1984): 217–223.

Davis, C. F. and P. Van Dine. "Linear Programming Model for Pavement Management." *Transportation Research Record* 1200, (1988): 71–75.

Ferreira, A., A. Antunes, and L. Picado-Santos. "Probabilistic Segment-Linked Pavement Management Optimization Model." *Journal of Transportation Engineering* 128 (6), (2002): 568–577.

Governor's Business Council. *Shaping the Competitive Advantage of Texas Metropolitan Regions*. 2006. Accessed on June 22, 2012. http://www.texasgbc.org/Trans%20Report% 20Docs/Shaping%20 the%20Competitive%20Advantage.pdf

Hakimi, S. and K. M. Kockelman. "Right-of-Way Acquisition and Property Condemnation: A Comparison of U.S. State Laws." *Journal of the Transportation Research Forum* 44 (3), (2005): 45-58.

Heiner, J. D. and K. M. Kockelman. "The Costs of Right of Way Acquisition: Methods and Models for Estimation." *Journal of Transportation Engineering* 131 (3), (2005): 193-204.

Krugler, P., C. Chang-Albitres, R. Feldman, S. Butenko, D. Kang, and R. Seyedshohadaie. *Report 0-5534-2: Development of Decision-Making Support Tools for Early Right of Way Acquisitions*. College Station, TX, 2010.

MNDOT. *The Financial Benefits of Early Acquisition of Transportation Right of Way*. Minnesota Department of Transportation, 2005.

NEPA. National Environmental Policy Act, 2012. Accessed on June 9, 2012. http://ceq.hss.doe.gov/

TxDOT. *Project Summary 0-5534:Asset Management–Texas Style,* 2009. Accessed on June 22, 2012. ftp://ftp.dot.state.tx.us/pub/txdot-info/rti/psr/5534.pdf

TxDOT. *Right of Way Manual Vol.2-Right of Way Acquisition,* 2011. Accessed on June 9, 2012. http://onlinemanuals.txdot.gov/txdotmanuals/acq/index.htm

Wang, K.C.P., C. Nunn, C. Mackey, W. Gong, D. Wilson, M. Evans, and J. Daleiden. "Network Level Crack Survey with the Automated Real-Time Distress Analyzer" (paper presented at the annual meeting for the Transportation Research Board, Washington, D.C., January 12-16, 2003).

Acknowledgements

The authors want to acknowledge the strong support of Ron Hagquist, who was TxDOT's project coordinator for the four-year research project 0-5534 "Asset Management–Texas Style" and followed for a one-year project optimization, and give special thanks to TxDOT for sponsoring this project. We also acknowledge the contribution of Texas A&M for developing the EROW model.

Carlos M. Chang Albitres, *Ph.D., P.E., is an assistant professor in the Department of Civil Engineering at the University of Texas at El Paso (UTEP). His major areas of interest are pavement evaluation and design, pavement management, construction management, knowledge management, infrastructure management, and asset management. Chang conducts research and teaches courses in his areas of expertise. Among his research sponsors are the National Highway Cooperative Research Program (NCHRP), Metropolitan Transportation Commission in California (MTC), the Texas Department of Transportation (TxDOT), and the City of El Paso. Chang is the academic advisor of the Chi Epsilon Chapter at UTEP and member of the American Society of Civil Engineering (ASCE). He is recognized as an international consultant and has participated as an expert in major highway projects. He is registered as an international consultant at the Inter-*

American Development Bank. Dr. Chang is the former Chair of the International Road Federation (IRF)Asset Management Task Force.

Paul E. Krugler, *P.E., is a former research engineer and manager of Research Implementation at the Texas Transportation Institute (TTI). Before joining TTI, he worked for the Texas Department of Transportation, and his experience includes: 14 years in the bituminous section of the Materials & Tests Division, five years as the director of Technical Operations in the Materials Section of the Construction Division, and five years as the TxDOT research office director, during which time he gained a broad perspective of statewide department technical and knowledge needs. In addition, Krugler has also led or been involved in numerous task forces, committees, and work groups pursuing paving materials, specifications, and forensic study objectives.*

Iraki Ibarra *earned his bachelor of science in civil engineering from The University of Texas at El Paso (UTEP) in December 2008. He started working in 2007 as an undergraduate research assistant in a project funded by City of El Paso that involved the inspection of the city's drainage system. He continued his education by entering the master's program also in civil engineering at UTEP focusing in transportation, infrastructure, asset management, and construction management. As a graduate research assistant, he has worked on several projects funded by City of El Paso, UTEP, and the Texas Department of Transportation.*

Edith Montes *received a bachelor of science in civil engineering from the University of Texas at El Paso (UTEP) in May 2011. She is currently working on her master of science in civil engineering at UTEP. Montes is a graduate research assistant in the Department of Civil Engineering of UTEP. Her thesis topic is "Fair Division Methods for Funding Allocation."*

Pavement Pre- and Post-Treatment Performance Models Using LTPP Data

by Pan Lu and Denver Tolliver

This paper determines that pavement performance in International Roughness Index (IRI) is affected by exogenous interventions such as pavement age, precipitation level, freeze-thaw level, and lower level preservation maintenance strategies. An exponential function of pavement age was used to represent pavement IRI performance curves. Moreover, this paper demonstrates a method which calculates short-term post-pavement performance models from maintenance effect models and pre-treatment performance models to avoid endogeneity bias but is still integrated with effects of maintenance activities on pavement performance.

INTRODUCTION

Predicting future pavement performance and condition is often the first and most important concern for pavement management plans. In modern pavement management systems (PMSs), pavement forecasting models are the most essential elements affecting many critical management decisions. Indeed, reliable performance prediction models are more important than ever before.

Pavement performance forecasting models are generally used to forecast changes in pavement condition over a future time period. Empirical regression models are the most popular models found in the literature and have greater practical value because of the infinite complexity of the underlying phenomena (AASHTO 1993). The models not only provide the pavement performance future changes in response to the known influential factors that cause the changes in those responses, but also provide the relationship between the performance indicators and influential factors. The model's ability to provide this relationship is important to the study in order to understand how the influential factors affect the performance indicators.

All the previous researchers' studies provide great insight on pavement performance, the influential factors that will affect pavement performance, and the potential problems in model formulation. Drawing from the previous studies, this paper will incorporate the idea of developing pre-treatment performance models and post-treatment performance models separately to avoid endogeneity bias. Additionally, the post-treatment performance models will take into account the effectiveness of different treatments applied at different times and pre-treatment performance models.

Long Term Pavement Performance Program (LTPP) is a comprehensive field-data pavement program. It monitors more than 2,400 pavement test sections across the United States and Canada (Elkins and Schmalzer 2009). LTPP was part of the Strategic Highway Research Program (SHRP) from 1987 to 1992. Now it is managed by the Federal Highway Administration (FHWA) and functions as a partnership with the states and provinces. Pavement performance data were available as International Roughness Index (IRI) in LTPP data. Pavement IRI survey values over years from the LTPP database are used in this study to establish pavement performance models. All qualified LTPP test sections, which include all precipitation, freeze-thaw, and maintenance information, are used in this analysis. So it covers the test sections across United States and some provinces in Canada combining all qualified specific pavement studies (SPS) and general pavement studies (GPS) sections. Only flexible pavement test sections were used in the anlysis. The standard data release (SDR) released in 2010 obtained through LTPP customer support service is the data source used in this research.

LITERATURE REVIEW

Hein and Watt (2005) developed a pavement performance prediction model with age and traffic as explanatory factors. Ozbay and Laub (2001) developed a basic IRI prediction model using initial IRI value (IRI value immediately after a pavement is built and open), pavement age, analysis age, structural number, and cumulative equivalent single axle load (ESAL) during analysis age. Gibby and Kitamura (1992) found that the most influential factors affecting the condition of local pavements are previous pavement condition, time elapsed since last major work, soil classification of roadway drainage, surface thickness, functional classification, and individual jurisdiction. Paterson and Attoh-Okine (1992) summarized that roughness progression in flexible pavement is developed by using traffic loading, strength, age, and environmental factors. Perera and Kohn (2001) summarized that environmental factors significantly affect roughness progression in AC pavement. The aforementioned authors commonly agree on the significance of climatic and age factors' effect on pavement deterioration. Maintenance and rehabilitations are also viewed as important factors that will affect pavement performance. However, some researchers suggest that maintenance and rehabilitations should always be viewed as endogenous variables (Prozzi and Mandanat 2004; Ramaswamy and Ben-akiva 1990). If such variables were incorporated into the model as one of the explanatory variables, then endogeneity bias will occur (Madanat, Bulusu, and Mahmoud 1995). Endogeneity bias is the main reason that some researchers have counterintuitive signs for the parameter estimates of important explanatory variables (Ramaswamy and Ben-akiva 1990).

Researchers such as Lytton (1987) have long recognized the need to develop models that respond to exogenous interventions but are also integrated with effects of maintenance activities. Many researchers tried to perform such tasks by accounting for maintenance and rehabilitation effects and also avoiding endogeneity bias. Maintenance and rehabilitations affect pavement condition directly and often are triggered by the condition of a pavement. In this situation, it is not recommended to include maintenance and rehabilitation directly as exogenous explanatory variables (Ramaswamy and Ben-akiva 1990).

Ramaswamy and Ben-Akiva (1990) developed a model that can simultaneously reflect pavement deterioration processes caused by exogenous influential factors and maintenance activities as a response to deterioration. Pavement condition and maintenance depend both on exogenous factors, as well as on each other. The results show a great improvement on having all the expected signs for all the significant parameters. Therefore, the model appears to be a more realistic one for predicting the deterioration of pavement with effects of maintenance activities. The study was the first to shed light on the difficulties associated with combining deterioration and maintenance. The drawbacks are that 1) the model's fit becomes less precise (R^2 value is 0.28), even though the simultaneous equation estimator gets rid of the endogeneity bias; and 2) the model assumes pavement condition and maintenance simultaneously depend on each other, which makes it difficult to forecast conditions under various maintenance and rehabilitation (M&R) policies. Therefore, the models are less useful to support M&R decision making.

Prozzi and Madanat (2004) developed a pavement performance model by combining experimental and field data. They first developed a ride quality model based on American Association of State Highway and Transportation Officials (AASHTO) road test experimental data and then re-estimated the parameters by applying joint estimation with the incorporation of the field data set. With well-designed experimental data, the endogeneity bias will be avoided. The model shows great benefits of using joint estimation such as improving the forecasts, lowering the estimate variance, and avoiding bias in the parameters. The main drawback of the model is that it requires both field data and well-designed experimental data for the regions with homogeneous weather conditions and level of maintenance activities if such data are not available.

Haider and Dwaikat (2010) introduced the idea of separate analysis pre-treatment performance curves, treatment performance changes, and post-treatment performance curves in simple format. Pre-treatment performance curves are connected with post-treatment performance curves by treatment performance changes. Haider and Dwaikat (2010) analyzed different treatment application timing effects and compared the pavement condition changes to the post-treatment deterioration rates for different timing policies. This method separates pavement performance models from maintenance effect models and finds a way to combine the effect with performance models. The drawback of Haider and Dwaikat's (2010) model is that it requires pre-treatment historical data to formulate pre-treatment performance curves and various post-treatment datasets for different treatment application timings to formulate different post-treatment performance curves. To research the different timing effect for each treatment requires too many post-treatment performance datasets. Such datasets are not always available, and may be too expensive to obtain.

PRE-TREATMENT PAVEMENT PERFORMANCE MODELS

In this paper, two types of pre-treatment performance models are developed. First, a pre-treatment performance model with an absolutely "do nothing" strategy will be developed. Second, the study will develop a pre-treatment performance model with minor preservation level maintenance activities (which only contains routine maintenances such as full depth patching, patching, skin patching, shoulder treatment, and seal coat). The rationale for developing two pre-treatment performance models stems from the fact that minor preservation level maintenance activities may not directly affect pavement surface condition indicators but will lower the deterioration rate with factors such as reducing the amount of moisture infiltrating the pavement structure and protecting the pavement system. To capture the effect of minor preservation level maintenance activities, one can compare performance curves with and without those minor preservation level maintenance activities. Minor preservation-level treatments are often categorized as routine maintenance. The study will assume the same level of minor preservation maintenance as those in LTPP program data.

Inspired by previous researchers' findings and to avoid endogeneity bias, simple regression models are developed using IRI as the dependent variable and pavement age as the independent variable for different precipitation and freeze thaw cycle regions. Traffic factors are not included in the research because of limited range of variation. Ninety percent of the observations used for this study have low traffic volumes. Because traffic factors are not included in the paper, the paper is more for demonstrating the methodology.

Three levels of freeze-thaw regions are defined according to the number of freeze-thaw days within a year. The regions are categorized as no freeze-thaw, medium freeze-thaw, and severe freeze-thaw. A freeze-thaw day is defined by a day's air temperature; if the air temperature changes from less than 0 degrees Celsius to greater than 0 degrees Celsius (or from less than 32 degrees Fahrenheit to greater than 32 degrees Fahrenheit), then that day is counted as one freeze-thaw day (USDOT and FHWA 2010)[1]. Regions are also classified based on levels of precipitation and are defined as a dry region or wet region based on the number of wet days per year. A wet day is defined by a day's amount of precipitation; an amount greater than 0.25 mm (or 0.01 inches), results in the day being counted as one wet day (USDOT and FHWA 2010)[2]. The detailed category information is shown in Table 1.

In Table 1, the values in Size 1 represent the number of observations of segments with do-nothing strategy in the corresponding region; Size 2 values represent the number of observations of segments with performing regular minor preservation activities.

Table 2 summarizes the number of observations for all six analysis regions: no freeze-thaw, dry region; no freeze-thaw, wet region; medium freeze-thaw, dry region; medium freeze-thaw, wet region; severe freeze-thaw, dry region; and severe freeze–thaw, wet region.

Table 1: Definition of Analysis Region for Pre-treatment Performance Models

Freeze-Thaw Region	**Definition**	**Size 1**	**Size 2**
No Freeze-Thaw	0 ≤ freeze thaw days per year < 70	48	697
Medium Freeze-Thaw	70 ≤ freeze thaw days per year < 140	52	955
Severe Freeze-Thaw	140 ≤freeze thaw days per year < 230	21	97
Precipitation Region			
Dry	0 ≤ wet days per year < 100	49	373
Wet	100 ≤ wet days per year < 270	72	1376

Table 2: Analysis Data for Pre-treatment Performance Models

Analysis Region	**Size 1**	**Size 2**
No Freeze-Thaw, Dry	15	181
Medium Freeze-Thaw, Dry	22	155
Severe Freeze-Thaw, Dry	12	37
No Freeze-Thaw, Wet	33	516
Medium Freeze-Thaw, Wet	30	800
Severe Freeze-Thaw, Wet	9	60

IRI data in LTPP have shown that IRI over time follows the shape of exponential functional form (Haider and Baladi 2010; Haider and Dwaikat 2010). In this study, the pre-treatment performance curve represented by exponential models as shown in equation (1) was tested, and promising statistical results shown later in the paper support the conclusion that IRI in LTPP can be represented by exponential models.

(1) $IRI_{pre}(t) = \alpha_1 * e^{\beta_1 * t}$

Where

α_1 = model parameters representing the initial value of IRI for pre-treatment performance curve

β_1 = model parameters representing the deterioration rate in IRI for pre-treatment performance curve

t = pavement age in months

Several factors have a role in pavement deterioration. Pavement age in months (t) represents the number of months for a pavement from the initial construction month or most recent reconstruction month. This variable is important because the rate of pavement deterioration is expected to change while pavement is aging. Because of the pavement age's relationship with deterioration, the variable is expected to have a positive sign.

Additionally, moisture is recognized as another important factor in pavement deterioration. The more moisture that penetrates a pavement under the surface layers, the faster the pavement will deteriorate. The freeze-thaw cycle is another important factor in affecting pavement deterioration rate. More frequent freeze-thaw cycles result in faster pavement deterioration.

The initial value of IRI for pre-treatment performance curve, ranged from 0.5 to 1.3 meter per kilometer. In this paper, 0.5 meter per kilometer (m/km) is selected as suggested by Haider and

Baladi (2010). The initial IRI benchmark value of 0.5 m/km was suggested to represent a brand new road condition (Haider and Baladi 2010; Haider and Dwaikat 2010). A typical initial serviceability value (which is considered to be the pavement service index value immediately after the pavement is built and open) of 4.5 pavement service index (PSI) was also suggested by AASHTO (1993). The PSI value of 4.5 roughly equals an IRI value of 0.5 m/km when correlation reported by Al-Omari and Darter (1994) was used: $PSI=5*e^{-0.26*IRI}$. All the above-mentioned literature suggest that when a pavement is brand new the IRI index value can be 0.5 m/km.

PRE-TREATMENT PAVEMENT PERFORMANCE MODEL RESULTS

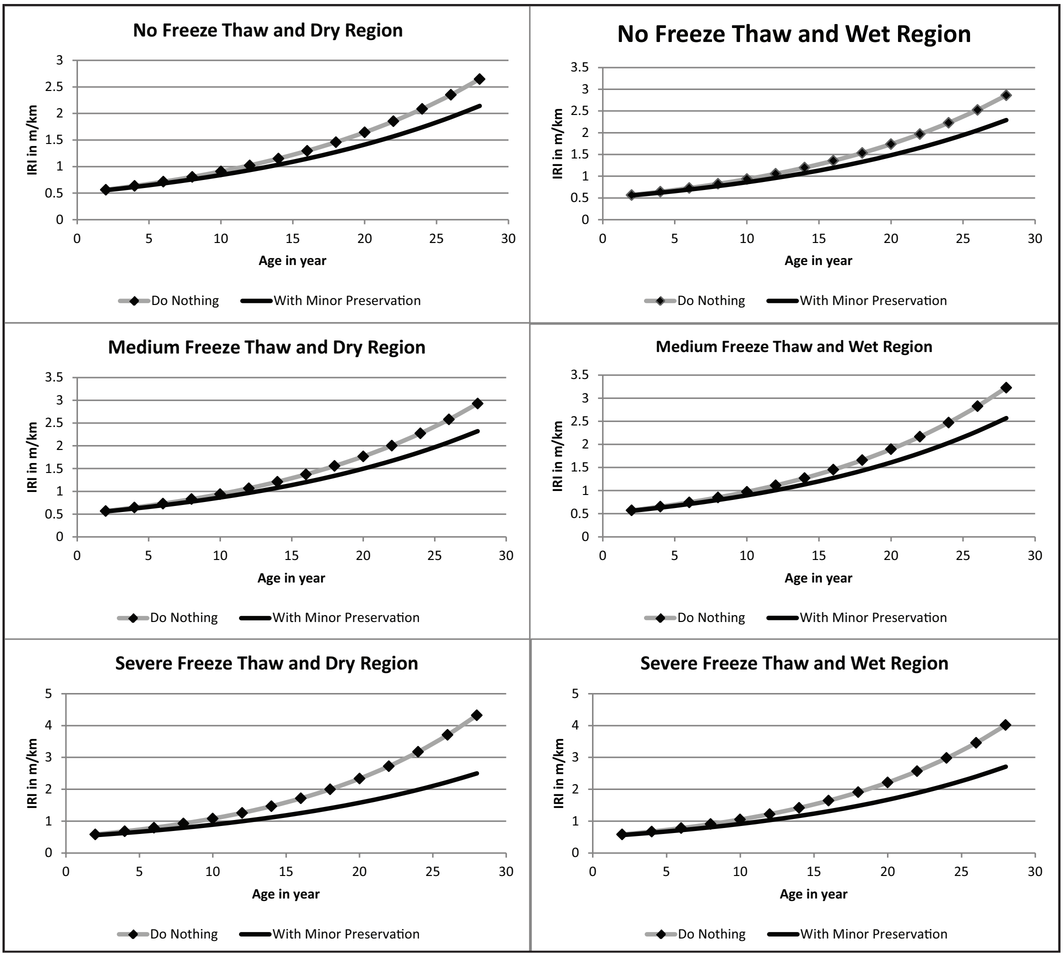
The model's key properties and parameter estimates are shown in Table 3. All 12 models have R-square values higher than 0.69, which suggests that all 12 models explain greater than 69% of the variation in pavement IRI values. The smallest number of observations for the models with minor level maintenance activities is 37, which is sufficient to realize large sample properties. It is rare to truly apply nothing to a pavement, especially to those segments located in a severe weather condition region. So it is not surprising to see that the available data with do-nothing strategy are limited. In this case the least number of observations are found in the severe freeze-thaw, wet region, which only contains nine observations, and the next severe weather condition region is severe freeze-thaw, dry region, which only contains 12 observations. The lowest deterioration rate is found in the no freeze-thaw, dry region with minor preservation activities. It matches expectations since the region is in the least freeze-thaw and the least precipitation affected region. The detailed analyses for each single influential factor will be discussed next.

Table 3: Key Model Properties and Parameter Estimates

Analysis Region	Estimated β_1	Prob. > | t | for β_1	Prob. > F for model	r-square	Adjusted r-square	Observations
With Minor Preservation No Freeze-Thaw, Dry	0.00433	< 0.0001	< 0.0001	0.8502	0.8494	181
With Do Nothing No Freeze-Thaw, Dry	0.00496	< 0.0001	< 0.0001	0.779	0.7632	15
With Minor Preservation Medium Freeze-Thaw, Dry	0.00457	< 0.0001	< 0.0001	0.777	0.7755	155
With Do Nothing Medium Freeze-Thaw, Dry	0.00526	< 0.0001	< 0.0001	0.8103	0.8103	22
With Minor Preservation Severe Freeze-Thaw, Dry	0.00479	< 0.0001	< 0.0001	0.8899	0.8033	37
With Do Nothing Severe Freeze-Thaw, Dry	0.00642	< 0.0001	< 0.0001	0.9143	0.9065	12
With Minor Preservation No Freeze-Thaw, Wet	0.00453	< 0.0001	< 0.0001	0.6913	0.6907	516
With Do Nothing No Freeze-Thaw, Wet	0.00519	< 0.0001	< 0.0001	0.796	0.7315	33
With Minor Preservation Medium Freeze-Thaw, Wet	0.00487	< 0.0001	< 0.0001	0.9055	0.895	800
With Do Nothing Medium Freeze-Thaw, Wet	0.00555	< 0.0001	< 0.0001	0.849	0.8438	30
With Minor Preservation Severe Freeze-Thaw, Wet	0.00503	< 0.0001	< 0.0001	0.8649	0.8626	60
With Do Nothing Severe Freeze-Thaw, Wet	0.0062	< 0.0001	< 0.0001	0.8835	0.8689	9

To compare minor preservation and do-nothing strategies as shown in Figure 1, it is found that the deterioration rate is higher with a do-nothing strategy than with a minor preservation strategy. The result is expected, because it is widely accepted that routine maintenance or minor preservation activities can reduce pavement deterioration rates.

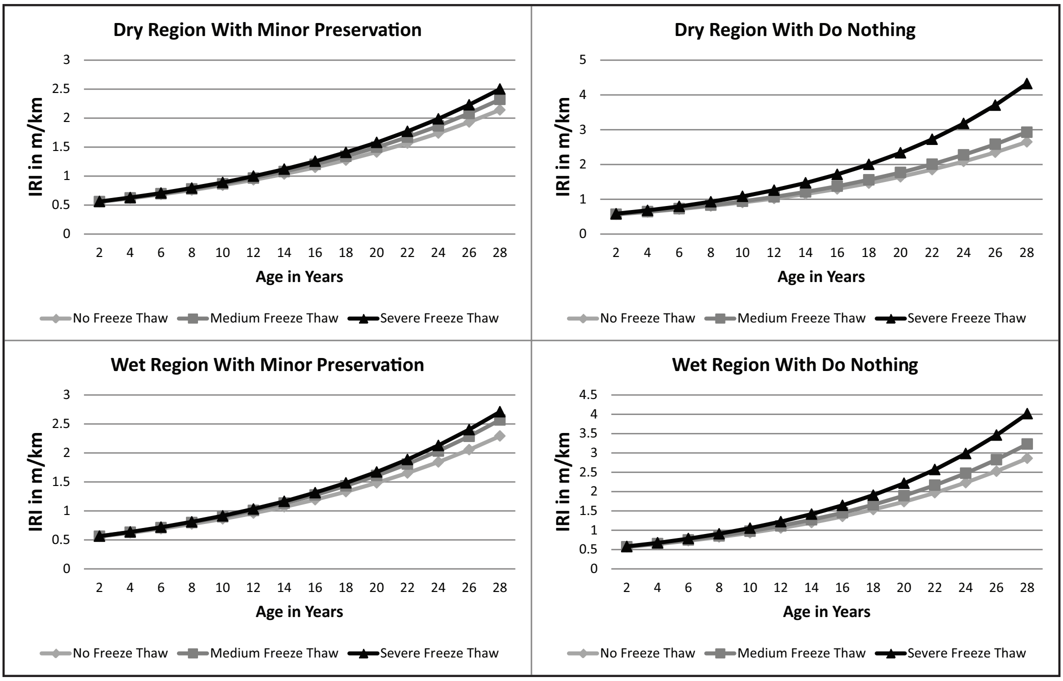
Figure 1: Pre-treatment Performances for Different Minor Preservation Strategies



The deterioration rates for the do-nothing strategy are higher than for the minor preservation strategy, but the magnitudes of the differences for each analysis region are different. Differences in severe weather condition regions tend to be greater than in other regions with less severe weather conditions.

To compare the differences of the freeze-thaw cycle effect shown in Figure 2, it is found that the deterioration rate increases as the freeze-thaw cycle level increases. The result is expected because freeze-thaw activity is commonly thought to accelerate pavement deterioration. The more freeze-thaw cycles within a year, the more quickly the pavement condition will deteriorate.

Figure 2: Pre-treatment Performances for Different Freeze Thaw Regions

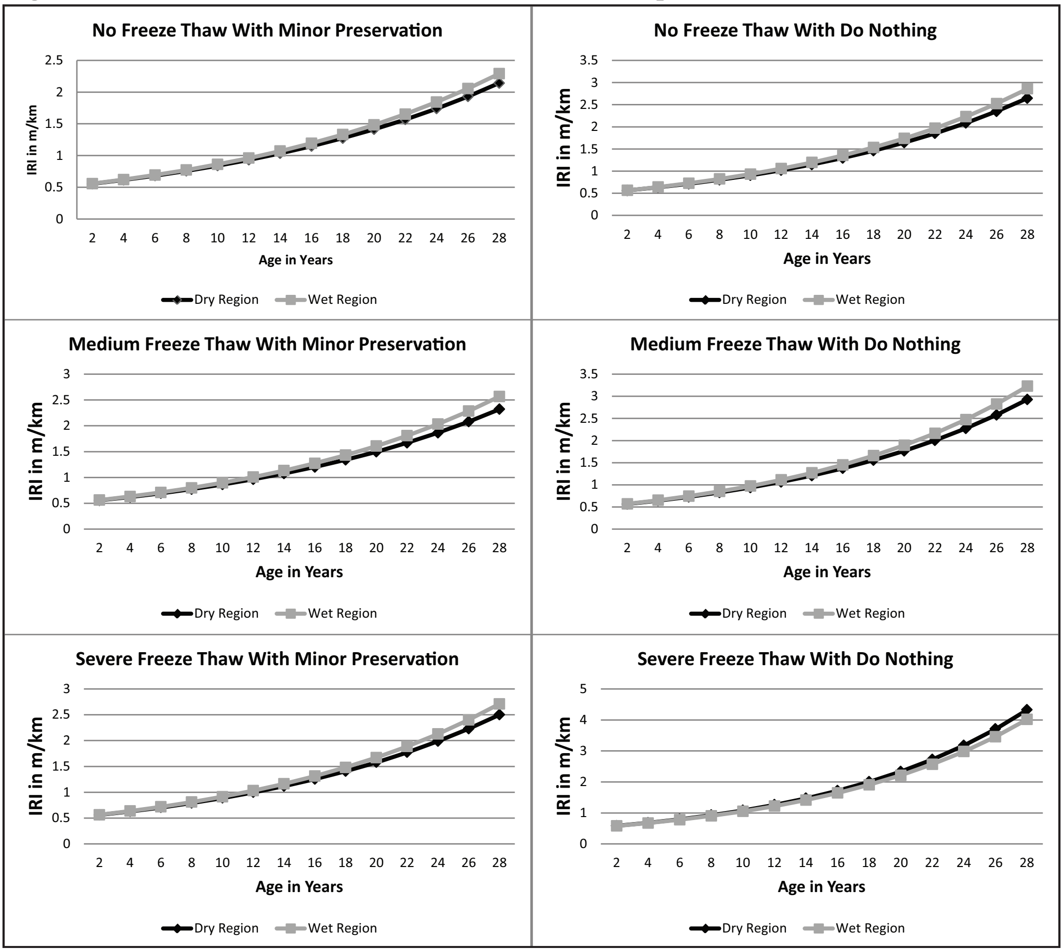


Notice that severe freeze-thaw cycles always have the highest deterioration rates and no freeze-thaw cycles always have the lowest deterioration rates for all four analysis categories. However, the range of differences among three freeze-thaw regions is less for the “minor preservation” strategy than the “do-nothing” strategy. This result shows that minor preservation activities are even more important in severe freeze-thaw regions than in no freeze-thaw regions because such minor preservation activities will decrease the pavement deterioration rate more significantly in severe freeze-thaw regions than in no freeze-thaw regions.

To compare the differences of the precipitation effect, as shown in Figure 3, it is found that, in general, the deterioration rate is higher in wet regions than in dry regions. The result is expected because it is widely accepted that precipitation is an influential factor contributing to pavement deterioration. The more precipitation a region receives within a year, the worse the pavement condition will be or the faster the deterioration rate will be.

The severe freeze-thaw and *wet* region using a do-nothing strategy has a lower pavement deterioration rate than the severe freeze-thaw and *dry* region using a do-nothing strategy. This result is counterintuitive, since the opposite effect is expected: a wet region should have higher deterioration rate than dry region with the same other properties. The reason for the counterintuitive result may be the limited number of observations, nine and 12, respectively. As mentioned before, it is rare to apply absolutely no treatment to a pavement, especially to a pavement located in a severe weather condition region. The result is more like a statistical average than a regression result with an insufficient number of observations.

Figure 3: Pre-treatment Performances for Different Precipitation Levels



POST- TREATMENT PAVEMENT PERFORMANCE MODEL FORMULATION

As mentioned earlier, IRI data in LTPP have shown that IRI over time follows the shape of an exponential functional form (Haider and Baladi 2010; Haider and Dwaikat 2010). Moreover, Irfan, Khurshid, Labi, and Flora (2009) suggested post-treatment pavement performance function has an exponential functional form. In this study, the post-treatment performance curve represented by exponential models as shown in equation (2) was assumed.

(2) $IRI_{post}(t) = \alpha_2 * e^{\beta_2 * t}$

Where

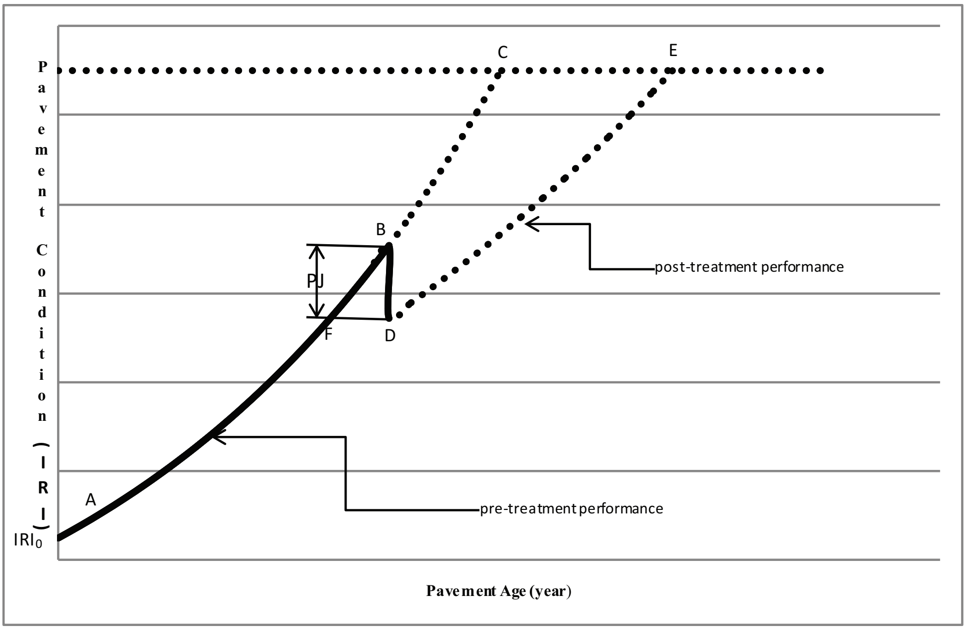
- α_2 = model parameters representing the initial value of IRI for post-treatment performance curve
- β_2 = model parameters representing the deterioration rate in IRI for post-treatment performance curve
- t = pavement age in months since last medium or major preservation activity

Figure 4 shows the relationship between pre-treatment performance curve, short-term treatment performance jump, and post-treatment performance. Figure 4 shows that the pre-treatment IRI follows the pre-treatment performance curve AFBC, unless a treatment is applied at point B, then the IRI value of the pavement will move to point D. The treatment performance jump denoted by PJ is defined as the difference in pavement condition immediately before and immediately after the application of a treatment. Then the post-treatment IRI value will follow the post-treatment performance curve DE.

One can tell from Figure 4 that the post-treatment and pre-treatment performance curves are connected by the treatment performance jump. Then the initial value of a post-treatment curve can be represented by equation (3).

(3) $\alpha_2 = IRI_{pre} - PJ$

Figure 4: Relationship of Pavement Performance and Treatment Performance Jump



With equation (3), pre-treatment performance curve and treatment effectiveness equations, it is not difficult to calculate for a pavement having a treatment applied at any time. Then the next task is to calculate post-treatment deterioration rate. Figure 4 shows the behavior of the post-treatment deterioration rate. A treatment applied to a pavement at point B restores the pavement's surface condition back to the IRI value associated with point D or point F. Treatment should lower the deterioration rate at the moment, which is only the deterioration rate at point D and it is supposed to be no greater than the deterioration rate at point B depending on the effectiveness of the treatment. The treatment should not improve the original condition of the pavement, which is the deterioration rate at point D. It is supposed to be no smaller than the deterioration rate at point F depending on the treatment effectiveness. For example, a reconstruction is considered as the highest effectiveness treatment. If a reconstruction is applied to a pavement, it will restore the IRI value and the deterioration rate back to their original values.

Finding the long-term post-treatment performance deterioration rate, is challenging. If the experimental data after a treatment application at various ages of a pavement are available, it is not difficult to find the post-treatment performance deterioration rate for different treatment application timings. Unfortunately, such data are rarely available and are expensive to obtain. For these reasons, some researchers' studies show unrealistic results.

Haider and Dwaikat (2010) showed that if a treatment is applied early in the pavement's life, it will result a lower treatment performance jump and slower post-deterioration rate. Treatment applied later in the pavement's life results in a higher treatment performance jump and a faster deterioration rate. The conclusion is not always accurate since the higher treatment jump can be combined with slower post-treatment deterioration rate. More specifically, if a treatment has a higher performance jump, it is supposed to lower the deterioration rate at the point instead of increasing the deterioration rate. From Figure 4, the deterioration rate at point D should be no greater than the deterioration rate at point B; and if a treatment has a lower performance jump, it is supposed to lower the deterioration rate at that point, but not lower than the deterioration rate at the after-treatment IRI value point on the pre-treatment curve. From Figure 4, the deterioration rate at point D should be no smaller than the deterioration rate at point F.

The main reason for Haider and Dwaikat (2010) to draw such a conclusion is because of the difficulty to obtain various post-treatment field performance data. They selected one section of pavement post-performance data with one control section of pavement performance data to calculate the treatment effect and deterioration rates. The number of observations is insufficient and the results are very sensitive to the available pavement performance data.

Before the post-treatment performance data become available, one can hardly verify the long-term, post-treatment deterioration rates. However, if researchers only focus on the short term treatment effectiveness, that is researchers only look at treatment performance within one year after a treatment is applied to the pavement, it is realistic to assume that if a treatment can restore the IRI value, it can also restore the deterioration rate based on the new IRI value. As explained in Figure 4, the deterioration rate at point D is the same as the deterioration rate at point F. In this case the pre-treatment deterioration rate, β_1, should be equal to the post-treatment deterioration rate, β_2, because the treatment effect is basically a restoration effect. The mathematical derivation of the relationship is shown in the next paragraph.

If a treatment is applied to a pavement at its age of t_i with a performance jump, PJ_i, the pre-treatment condition can be expressed via Equation (1) as Equation (4) and the pavement age associated with post-treatment IRI condition, t_j, can be expressed via Equation (2) as Equation (5).

(4) $IRI_i^{pre} = \alpha_1 \times e^{\beta_1 \times t_i}$

(5) $t_j = \left[\ln\left(\frac{\alpha_1 \times e^{\beta_1 \times t_i} - PJ_i}{\alpha_1}\right)\right] / \beta_1$

Restored deterioration rate behavior can be expressed as in Equation (6).

(6) $\left.\frac{df_{pre}(\mathrm{t})}{dt}\right|_{t=t_j} = \left.\frac{df_{post}(\mathrm{t})}{dt}\right|_{t=0}$

Where,

f_{pre}(t) and f_{post}(t) are pre-treatment and post-treatment performance functions respectively.

df_{pre}(t) and df_{post}(t) are derivatives of pre-treatment and post-treatment performance functions respectively.

By calculating the derivative of pre-treatment performance curve and post treatment performance functions, one can get Equation (7).

(7) $\alpha_1 * \beta_1 * e^{\beta_1 * t_j} = \alpha_2 * \beta_2$

By substituting Equations (3), (4), and (5) into Equation (7), a relationship between β_1 and β_2 at treatment application time will be obtained as shown in equation (8).

(8) $(\alpha_1 \times e^{\beta_1 \times t_i} - PJ_i) \times \beta_1 = (\alpha_1 \times e^{\beta_1 \times t_i} - PJ_i) \times \beta_2$

Since $\alpha_1 \times e^{\beta_1 \times t_i} - PJ_i = \alpha_2$ and $\alpha 2 \neq$ therefore the post-treatment deterioration rate, β_2, equals the pre-treatment deterioration rate, β_1.

In general, post-treatment performance curve function can be expressed in Equation (9).

(9) $IRI_{jt_i} = (\alpha_1 \times e^{\beta_1 \times t_i} - PJ_{jt_i}) \times e^{\beta_1 \times t}$

Where

IRI_{jt_i} = post-treatment IRI when treatment j applied to a pavement at t_i

α_1 = model parameter representing the initial value of IRI for pre-treatment performance curve

β_1 = model parameter representing the deterioration rate in IRI for pre-treatment performance curve

t_i = pavement age when a treatment applied to the pavement

PJ_{jt_i} = pavement IRI performance jump when treatment j applied to the pavement at age t_i

t = post-treatment pavement age in months

PAVEMENT PERFORMANCE MODEL VALIDATION

Regression model forecast verification is sometimes called validation, or evaluation. The purpose of this process is to help assess the specific strengths and deficiencies of regression models when they are used to forecast values of the dependent variable using values of the explanatory variables that were not represented in the sample dataset used to estimate the model. Ultimately, this process may provide justifications for uses of the model for forecasting and supporting better decision making (Wilks 2006).

Cook and Kairiukstis (1990) state that reduction of error (RE) "should assume a central role in the verification procedure" (p. 181). RE is an example of forecast skill statistic (Wilks 2006). Wilks (2006) defined forecast skill as the relative accuracy of a set of forecasts with respect to some set of standard controls, which are usually the average values of the predictand. The equation used to calculate RE can be expressed in the following Equation (10).

(10) $RE = 1 - \frac{SSE_v}{SSE_{ref}}$

Where SSE_v = sum of squares of validation errors between observed and predicted values over the validation period and SSE_{ref} = sum of squares of validation errors between observed and control values or reference values over the validation period.

Jackknife is a statistical method for systematically computing the statistical estimate leaving out one observation at a time from the sample set (Wilks 2006). One application of the Jackknife procedure is to compare the difference between the omitted observation's value and predicted value for the omitted observation. The difference is defined as validation error. It can be mathematically expressed as Equation (11).

(11) $e_{(i)} = y_i - \widehat{y_{(l)}}$

Where y_i and $\widehat{y_{(l)}}$ are the observed and predicted values of the predictand for validation data set i, and the notation (i) indicates that the validation data set i was not used in fitting the model that generated the prediction $\widehat{y_{(l)}}$.

The sum of the squares of errors for validation, SSE_v, can be expressed as Equation (12) and the sum of squares of errors for reference, SSE_{ref}, can be expressed as Equation (13).

(12) $SSE_v = \sum_{i=1}^{n_v} e_{(i)}^2$

(13) $SSE_{ref} = \sum_{i=1}^{n_v} (y_i - \bar{y})^2$

Where n_v is the validation period or the number of validation tests and $\bar{y}$ is the mean of the predictand, which usually serves as a reference or control value.

Using a leave-one-out Jackknife procedure will select one validation data point to exclude from the original dataset at one time, until all the data points are selected once. The regression calculations are repeated n times, which should be the number of observations in the original data set. Notation i in Equation (11) can then be explained as the ith observation in the original dataset, which is selected as the leave-out validation data. Notation n_v in Equation (12) and (13) then should be n. Equation (12) will then be identical to PRESS (the predicted residual sum of squares) and RE can be expressed as Equation (14) and used as the objective model verification statistic in this study.

(14) $RE = 1 - \frac{\sum_{i=1}^{n} (y_i - \widehat{y_I})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = 1 - \frac{PRESS}{SSE_{ref}}$

Theoretically, the value of RE can range from negative infinity to one, where one indicates perfect prediction for the validation data set. It will only occur when all the residuals for validation data are zero (i.e. PRESS = 0). On the other hand, if PRESS is much greater than SSE_{ref}, RE can be negative and large. As a rule of thumb, a positive RE indicates that the regression model on average has some forecast capability. Conversely, if RE ≤ 0, the model is deemed to have no ability to predict (Cook and Kairiukstis 1990; Wilks 2006). The similarity in form of the equations for RE and regression R^2 expressed as Equation (15) suggests that RE can also be used as validation evidence for R^2. The closer the values of RE and R^2 are to each other, the more the model is accepted as a predictive tool. RE sometimes is referred as Jackknife R^2.

(15) $R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^{n} (y_i - \widehat{y_I})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$

Twelve pre-treatment performance models were developed in this study. The model validation results are shown in Table 4. Difference is calculated by (R-Square – Jackknife R-Square) / R-Square.

Table 4: Forecast Validation of Regression Models

Models	R-Square	Jackknife R-Square	Percent Difference
S-D-Nothing	91.43%	89.87%	1.71%
S-W-Nothing	88.35%	85.17%	3.60%
M-D-Nothing	81.03%	78.37%	3.28%
M-W-Nothing	84.90%	83.94%	1.13%
N-D-Nothing	77.90%	73.03%	6.25%
N-W-Nothing	79.60%	72.05%	9.48%
S-D-Minor	80.87%	79.84%	1.27%
S-W-Minor	86.49%	86.02%	0.54%
M-D-Minor	77.70%	77.26%	0.57%
M-W-Minor	78.63%	78.58%	0.06%
N-D-Minor	85.02%	84.88%	0.16%
N-W-Minor	69.13%	68.98%	0.22%

Table 4 provides several indications about the models: (1) all the models are accepted with consideration of some forecast skills, since all the Jackknife R-Squares are positive values, (2) all the models are accepted as validated models, since most of them have Jackknife R-Square values close to R-Square, and (3) all the models have Jackknife R-Square values higher than 70%. The Jackknife R-Square values shown in Table 4 indicate that most of the models can be accepted as exhibiting higher forecasting ability, since 100% means perfect forecast (Cook and Kairiukstis 1990, Wilks 2006).

SUMMARY AND CONCLUSIONS

This paper presents pre-treatment performance models with exponential form. Moreover, the paper demonstrates a method for determining the post-treatment performance model by using pre-treatment performance models and short-term performance effectiveness.

The pre-treatment models demonstrate several findings. The models reveal that pavement IRI deterioration rate is higher with a do-nothing strategy compared with a minor preservation strategy. However, the magnitudes of the differences for different analysis regions vary. Differences in severe weather condition regions tend to be greater than in less severe weather condition regions. Additionally, deterioration rates increase with a freeze-thaw cycle level increase. However, the range of the differences among three freeze-thaw regions is smaller for a minor preservation strategy region than for a do-nothing strategy region. It shows that minor preservation activities are even more important in severe freeze-thaw regions than in no freeze-thaw regions because such minor preservation activities will lower pavement deterioration rates more significantly in severe freeze-thaw regions than in no freeze-thaw regions. The models also demonstrate that the deterioration rate is higher in wet regions than in dry regions. In most cases, a wet region has higher deterioration rates than a dry region.

The study results can have important implications in the design of pavement preservation strategies decision making. Pavement engineers often need to develop pavement preservation schedules. Knowledge of pre- and post-treatment pavement performance is essential for engineers to decide when and what subsequent preservation activities should be carried out.

Endnotes

1. One might argue that freeze-thawing conditions occur over longer periods than a single day. However, what we defined here is based on FHWA's LTPP's definition, which is based on freeze-thaw days. There is no other way we can obtain the true freeze thaw condition from the available LTPP data.

   We believe the LTPP definition makes sense. Yes, the actual freeze thaw conditions occur over longer periods than a single day. But when the true freeze thaw happens it is always linked to changing temperature from less than 32 to more than 32 degrees. And the definition used here is when freeze-thaw days are less than 70 days in a year, it is considered as no-freeze-thaw category. From 70 freeze-thaw days to 140 freeze-thaw days, it is considered as medium freeze-thaw category. And from 140 to 230 days, it is considered as severe freeze-thaw category. The rationale behind this definition is the more frequent temperature changes from less than 32 to more than 32 in a single day the more likely freeze-thaw condition will happen.

   From the statistical models, the results show a clear difference among the current freeze-thaw level definitions.

2. This is the LTPP definition based on the number of wet days in a year, not the total amount of precipitation within a year. The rationale of the LTPP's definition is the more frequent the wet days within a year the greater the precipitation effect on the pavement.

References

Al-Omari, B. and M.I. Darter. "Relationships Between International Roughness Index and Present Serviceability Rating." *Transportation Research Record* 1435, (1994): 132-136.

American Association of State Highway and Transportation Officials (AASHTO). *AASHTO Guide for Design of Pavement Structures*. Washington, D.C., AASHTO, 1993.

Cook, E. and L. Kairiukstis *Methods of Dendrochronology: Applications in the Environmental Science.* Kluwer Academic Publisher and International Institute for Applied Systems Analysis, Dordrecht, The Netherlands, 1990.

Elkins, G.E. and P. Schmalzer. *Long-term Pavement Performance Information Management System: Pavement Performance Database User Reference Guide* (FHWA-RD-03-088), 2009.

Gibby, A.R. and R. Kitamura. "Factors Affecting Condition of Pavement Owned by Local Governments." *Transportation Research Record* 1344, (1992): 116-123.

Haider, S.W. and G.Y. Baladi. "Effect of Pavement Condition Data Collection Frequency on Performance Prediction." *Transportation Record*, 89th Annual Meeting, 2010.

Haider, S.W. and M.B. Dwaikat. "Estimating Optimum Timing for Preventive Maintenance Treatments to Mitigate Pavement Roughness." *Transportation Record*, 89th Annual Meeting, 2010.

Hein, D. and D. Watt. *Proceedings of the Annual Conference of the Transportation Association of Canada.* "Municipal Pavement Performance Prediction Based on Pavement Condition Data." Calgary, Alberta, Canada, 2005.

Irfan, M., M. Khurshid, S. Labi, and W. Flora. "Evaluating the Cost Effectiveness of Flexible Rehabilitation Treatment Using Different Performance Criteria." *Journal of Transportation Engineering* 135 (10), (2009): 753-736.

Lytton, R. L. *Proceedings of Second North American Conference on Managing Pavements: Concepts of Pavement Performance Prediction and Modeling,* Vol. 2. Toronto, Canada: Public Roads, 1987.

Madanat, S.M., S. Bulusu, and A. Mahmoud. "Estimation of Infrastructure Distress Initiation and Progression Models." *Journal of Infrastructure Systems* 1 (3), (1995): 146-150.

Ozbay, K. and R. Laub. "Models for Pavement Deterioration Using LTPP." *New Jersey Department of Transportation, Division of Research and Technology and U.S. Department of Transportation Federal Highway Administration*, New Jersey, 2001.

Paterson, W.D.O. and B. Attoh-Okine. "Summary Models of Paved Road Deterioration Based on HDM-III." *Transportation Research Record* 1344, (1992): 99-105.

Perera, R.W. and S.D. Kohn. *LTPP Data Analysis: Factors Affecting Pavement Smoothness.* National Cooperative Highway Research Program Project 20-50 (8/13). Washington, D.C.: Transportation Research Board, 2001.

Prozzi, J.A. and S.M. Madanat. "Development of Pavement Performance Models by Combining Experimental and Field Data." *Journal of Infrastructure System* 10 (1), (2004): 9-22.

Ramaswamy, R. and M. Ben-Akiva. "Estimation of Highway Pavement Deterioration From In-service Pavement Data." *Transportation Research Record* 1272, (1990): 96-106.

U.S. Department of Transportation (USDOT) and Federal Highway Administration (FHWA). *Long-term Pavement Performance Information Management System: Pavement Performance Database User Reference Guide.* Publication No. FHWA-RD-03-088. Washington, D.C., U.S. Department of Transportation, 2010.

Wilks, S.D. *Statistical Methods in the Atmospheric Sciences (2nd ed.).* Academic Press, Elsevier, London, 2006.

Pan Lu *has been with the Upper Great Plains Transportation Institute since 2005, and has served as associate research fellow since 2010. Her recent research areas include county and local road analysis, economic impacts of freight transportation in North Dakota, multimodal transportation fuel efficiency analysis, pavement performance modeling, multi-objective pavement preservation decision making, and bridge performance studies. She received her B.S. in engineering economics from North China Electric Power University of Beijing in 2002. She earned her Ph.D. in transportation and logistics from North Dakota State University in 2011.*

Denver Tolliver *is director of the Transportation and Logistics Program and director of the Upper Great Plains Transportation Institute at North Dakota State University. He is also the director of the eight-university Mountain-Plains Consortium, which is one of the USDOT's designated university transportation centers. He holds doctoral and master's degrees from the Virginia Polytechnic Institute in environmental design and planning and urban and regional planning respectively. Since 1989, Tolliver has been awarded $23 million in competitive research grants from federal and state agencies and private industries to conduct highway, railway, waterway, and multimodal analyses. He recently received the 2011 Institute of Transportation Engineers award for innovation in Transportation Education.*

Productivity Improvements in the U.S. Rail Freight Industry, 1980-2010

by Carl D. Martland

Between 1980 and 2008, extensive productivity improvements and changes in traffic mix allowed railroads to become more profitable despite declining prices and stronger competition from motor carriers. These productivity improvements enabled the Class I railroads to halve their real costs per ton-mile, even though costs for fuel and other resources rose faster than inflation. Productivity improvements were greatest for bulk traffic moving in unit trains, containers moving in double-stack trains, and high-volume shipments moving long-distances in specialized equipment. While the rail industry indeed achieved tremendous improvements in productivity following passage of the Staggers Act in 1980, it is incorrect to point to deregulation as the primary reason for these gains. Other factors that were even more critical to productivity growth included technological advances, new labor agreements, improved management, and public policy responses to the Northeast Rail Crisis.

BACKGROUND AND OVERVIEW

In the mid-1960s, although the rail industry appeared to be recovering from the stresses of World War II, the Korean War, and the recessions of the 1950s, it was on the brink of what became known as the Northeast Rail Crisis. The unexpected bankruptcy of the Penn Central Railroad in 1970 led to a decade of restructuring the national rail network, the railroads that operated over that network, and the regulations that determined what those railroads could or could not do. The bankruptcy of the Penn Central and a half dozen smaller railroads threatened severe disruption to transportation and distribution systems in the northeast, the loss of tens of thousands of railroad jobs, the loss of rail service to hundreds of communities, and severe costs to electric utilities, automobile assembly plants, chemical plants, steel mills, and other rail-dependent companies.

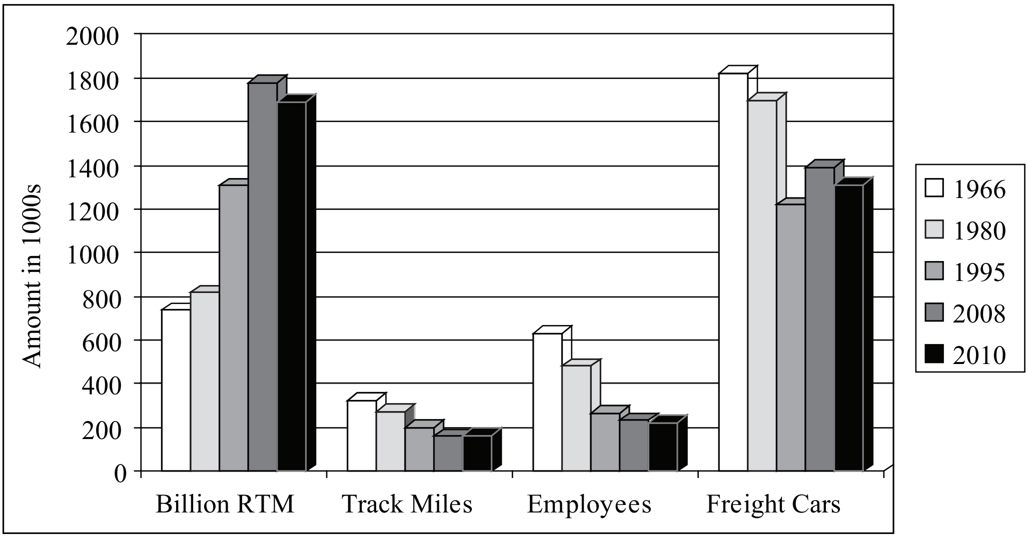
Because of widespread concerns over these possibilities, the federal government took dramatic action. In 1971, Congress created Amtrak, which allowed railroads to exit the unprofitable inter-city passenger business. Next, Congress devised a process that led to the creation of Conrail, a new federally controlled railroad that emerged from the chaos of the bankruptcies in 1975. Although Congress also gave railroads greater pricing freedom and an easier path toward network rationalization,[1] the financial ills of the rail industry continued. In the late 1970s, the Chicago, Rock Island and Pacific Railroad, and the Milwaukee Road filed for bankruptcy, and a Congressionally mandated study highlighted the dire situation facing the industry:

> *Continuation of the trends of the postwar period would result within the next 10 years in an industry facing enormous capital shortages, competing only for bulk shipments of low-value goods, lacking the resources needed for safe operation, and to a very considerable degree, operating under the financial control or ownership of public agencies.* (Secretary of Transportation 1978, p. 3)

Congress considered, but ultimately chose not to extend a Conrail-type solution to that region. Instead, the major legislative actions during the 1970s were enhanced and extended by passage of the Staggers Act in 1980.

Over the next three decades, the rail industry improved its productivity, lowered its prices, and increased its profitability.[2] Despite more than doubling the revenue ton-miles (RTM) that were handled, the industry sharply reduced its track-miles, its employees, and its fleet of freight cars (Figure 1). A series of mergers reduced the number of Class I railroads and ultimately produced two large systems in both the eastern and western portions of the U.S. While tens of thousands of miles of light density lines were abandoned, a similar amount were converted to short line or regional railroads that now play a major role in serving customers.

Figure 1: Output Doubled, but Resource Requirements Declined Between 1966 and 2008



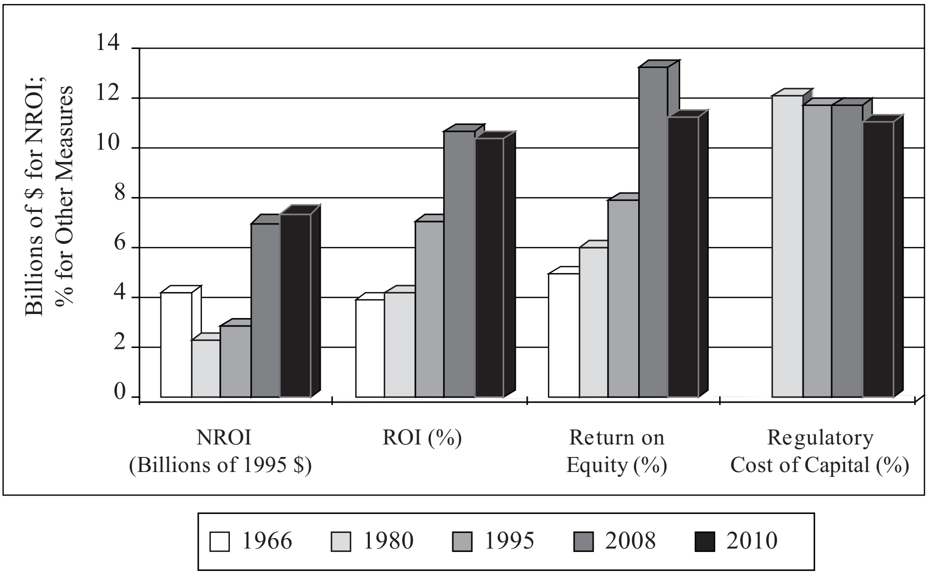
Data Source: AAR, *Railroad Facts*, various editions.

These productivity improvements were instrumental in improving the industry's financial performance (Figure 2). Net railway operating income (NROI), a commonly used measure of railroad profitability, more than tripled (in constant dollars) between 1980 and 2010, while return on investment (ROI) more than doubled. NROI equals operating revenues minus the sum of operating expenses, current and deferred taxes, and rents for equipment and joint facilities; it does not include interest expense or non-operating revenues or expenses. ROI is calculated as the ratio of NROI to average net investment in railroad property. The ability to raise capital is more related to the return on shareholders equity than to ROI; if return on equity is greater than the railroads' cost of capital, then they can more readily raise the funds needed to maintain and improve their services. The Association of American Railroads (AAR) has acknowledged the importance of productivity improvements for the turnaround in the industry's financial performance:

> *Better output per employee, more efficient utilization of infrastructure, and improved locomotive fuel efficiency helped freight railroads attain their best industry operating ratio (78.6%) since World War II. The resulting financial performance, which included a return on equity of 11.3% and a return on investment of 10.2%, was a welcome and long-sought improvement after a disappointing record over the last forty years.* (AAR 2007, p.5)

The deregulation of the rail industry, specifically the passage of the Staggers Act in 1980, is often cited by both economists and rail industry officials as the reason for the productivity improvements the industry achieved (e.g., Morrison and Winston 1999, Gómez-Ibáñez and de Rus 2006, Hamberger 2010). Deregulation certainly facilitated rationalization of the network, encouraged marketing initiatives, and enabled railroads to concentrate on their most profitable lines of business. However, a closer examination of the sources of productivity changes, as summarized in this paper, reveals

Figure 2: Financial Performance Improved Between 1966 and 2010



Data Source: AAR, *Railroad Facts*, various editions.

that it was not deregulation, but changes in traffic mix, rail-related research and development, rapid improvements in communications and computers, new labor agreements, and public investment in railroads that led directly to most of the industry's productivity gains. Furthermore, deregulation was the dominant factor in declining rail rates, which for decades hindered the industry's ability to translate productivity improvements into financial success. The Staggers Act eliminated the regulatory framework that had long allowed railroads to raise their rates to keep pace with inflation. Proponents of deregulation expected the act to allow higher rates, but the dominant long-term results were enhanced competition and decades of declining rates (Lowtan 2004).

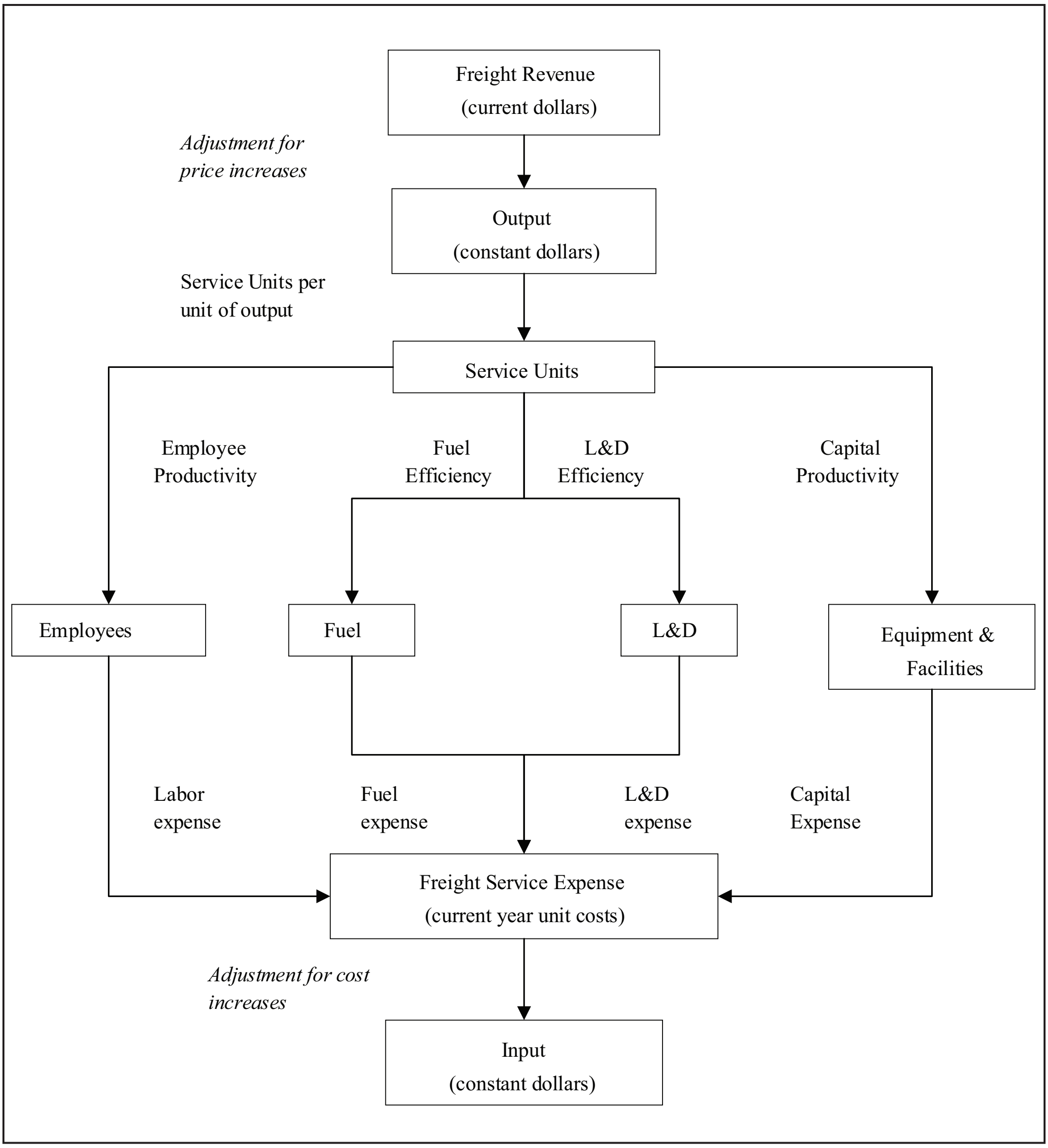
MEASURING PRODUCTIVITY

In basic economic terms, productivity is the ratio of output to input. It is possible to become more productive by producing more output from the same resources or by producing the same output with fewer resources. This simple concept is complicated by the difficulties in measuring outputs and inputs, especially for a complex system with multiple outputs and inputs. Partial measures, such as revenue ton-miles per employee, may be easy to calculate, but they fail to account for the substitutability of inputs, e.g., capital investments that reduce the need for labor. Broader measures are needed to capture overall trends in productivity.

Two quite different approaches have been used to measure and understand trends in railroad productivity. Transportation economists typically employ econometric techniques to calibrate cost functions and to estimate productivity change based upon public data sources. These techniques can be used to investigate important economic concepts such as economies of scale, scope, and density (Bereskin 2009). However, they can only show how costs or productivity change with respect to very aggregate measures such as average length of haul, revenue ton-miles per route-mile, or average route-miles per railroad. For example, Christensen Associates (2010) recently completed an econometric study of railroad costs and productivity for the Surface Transportation Board (STB). They found that the rate of productivity growth declined from 6.8% in 1987 to 1% in 2000 before turning negative in 2008, but their methodology could not explain the causes underlying either the growth or the decline.

Access to railroad information systems enables an analytical approach than can identify the sources of productivity improvement. Earlier studies have measured railroad productivity as the ratio of current year freight revenue deflated by a rail price index to current year freight expense deflated by the AAR's Railroad Cost Recovery Index (Martland and McCullough 1984, Martland 1989, Martland 1999, Martland 2006). With this approach, productivity in the base year is the inverse of the operating ratio, which in fact is commonly cited by railroads as a measure of productivity. Since the measures of both output and input are directly tied to the normal accounting framework used in the industry, it is possible to investigate the causes of productivity changes (Figure 3). Changes in output can be directly related to changes in traffic mix, traffic volume, or revenue per ton, ton-mile or carload. Changes in input can be directly tied to specific aspects of operations, loss and damage (L&D), capital investment, or the makeup of the labor force.

Figure 3: Structure for Analyzing Railroad Productivity



Source: Martland and McCullough, 1984, p. 54.

Moving rail shipment requires various activities known as service units, including train-miles, car-miles, tons, and carloads. The number of service units required per unit of output depends upon actual traffic characteristics and operating conditions, e.g., average length of haul, network structure, operating plans, and train capacity. Resource requirements relate to appropriate service units, e.g., the number of track maintenance employees can be related to gross-ton-miles. Current year freight expense is the summation of the costs of all the inputs. Adjusting for increases in the unit costs for each input provides a measure of inputs for each year. With this framework, changes in overall productivity can reflect changes in rate indices and unit costs as well as changes in traffic characteristics, operating capabilities, and resource utilization. While this approach makes it possible to examine the sources of productivity improvement, it does not as readily deal with issues such as economies of scale, scope, and density that can be addressed within a well-defined economic approach.

OVERALL PRODUCTIVITY IMPROVEMENT 1966 to 2010

Although the primary focus of this paper concerns the post-Staggers period, it is necessary to begin well before 1980 in order to understand trends in performance and in order to have a basis for comparing performance before and after deregulation. This section presents measures related to the overall rate of productivity growth and trends in financial performance. Table 1 shows output, input, productivity, and the ratio of prices to costs for selected years, beginning in 1966. A measure of rail freight output (Row O3) was calculated by dividing the current year freight revenue (Row O2) by a rail price index based upon data published by the Surface Transportation Board (Row O1).[3] This output measure was then converted to an index in which 1985 = 100 (Row O4). Since prices fell steadily as a result of deregulation, this measure of output grew faster than freight revenues. A measure of rail freight input (Row I3) was calculated by dividing current year freight expense (Row I2) by the railroad cost recovery index published by the Association of American Railroads (Row I1). This input measure was then converted into an index in which 1985 equals 100 (Row I4). Although freight operating expenses were much higher in 2010 than in 1985, the input index was much lower. Productivity (Row P1) is the ratio of the output (Row O3) to input (Row I3). Two indices of productivity are shown, one with a base year of 1985 and one with a base year of 1995. The final row of the table shows the compounded annual rate of change:

(1) Rate of Change = (Productivity Year N)/(Productivity Year 0)$^{1/N}$ – 1

For example, productivity increased from 230 in 1995 to 310 in 2000, an overall increase of 34.8%. The annual rate of increase was 6.166% (rounded off in the table to 6.2%). Over a five-year period, this rate of increase, compounded annually, would increase productivity by 34.8% (i.e., $(1.06166)^5$ -1 = .348).

During this 44-year period in which output nearly quadrupled, input declined by nearly two thirds and productivity increased more than seven-fold, from 62 to 476. Productivity improvement was much greater after 1980. The annual rate of change, which was 2% between 1966 and 1980, accelerated to 9.4% in the late 1980s, then began a gradual decline, with productivity improvement of 2.9% per year between 2008 and 2010.

The ratio of the price index to the cost index (Row R1) declined throughout this period, although the ratio was fairly stable from 2005 to 2010. Prior to 1980, when railroad productivity growth was slow, inflation was a national problem. The railroads at that time complained that the Interstate Commerce Commission would not let them adopt general increases that would allow them to recoup their rising costs (Secretary of Transportation 2008; Lowtan 2004). However, rail rates rose by a factor of three between 1966 and 1980, and average constant dollar revenue per ton-mile in 1980 was just 3% less than in 1966 (Row R2). The problem for railroads in 1980 was that rail costs had been rising even faster than prices, and the ratio of the price index to the cost index fell from 1.64 to 1.23. After 1985, the price index actually declined, while the cost index continued to rise. In 2010, the ratio of the price index to the cost index (Row R1) was down to 0.277.

In general, when rates fail to keep pace with costs, profitability suffers unless productivity improvements can offset rising cost. Because the railroads were able to achieve rapid productivity growth over a prolonged period, they were able to increase their profitability despite rising costs and falling prices. If rail revenue per ton-mile in 2010 had been the same as in 1980, revenue would have been $112 billion rather than $56 billion ($56.3 billion x 4.86 cents/ton-mi / 2.45 cents/ton-mi = $112 billion). Rail revenue per ton-mile fell in part because of changes in traffic mix and in part because of reductions in rates.

Using a rail rate index eliminates the effects of traffic mix and inflation. If prices had remained at 1980 levels, and if there had been no change in traffic volume or traffic mix, then freight revenues would have been higher:

(2) Revenue at 1980 rates = Revenue Year N *(Price Index 1980/Price Index Year N)

For example, if rail rates had simply managed to remain at 1980 levels, then rail revenue for the traffic moved in 2010 would have been $75.7 billion ($56.3 billion x 94.2 / 70 = $75.7 billion), an increase of $19.4 billion. Revenue would have been higher if rates had risen to offset some of the effects of inflation.

While shippers have certainly benefited from lower prices, railroads have also benefited by having higher NROI. In current dollars, NROI rose from approximately $1 billion in 1980 to $10 billion in 2010 (Table 2, Row 1). As a percentage of freight revenues, NROI rose from 5% in 1980 to 17.7 % in 2010 (Table 2, Row 2). If NROI had remained at 5% of freight revenue, NROI would have been much lower in 1985 and in subsequent years (Row 3). For example, NROI would have been $2.9 billion rather than $10 billion in 2010, and the difference of $7.1 billion can be interpreted as the railroads' net gain from productivity improvements.

Row 6 in Table 2 shows the combined benefits to railroads and to their customers, while Row 7 shows that the greatest share of the net benefits went to customers in terms of lower prices[4] rather than to railroads in terms of higher NROI.

This section has documented the overall changes in outputs, inputs, productivity, and profitability. Productivity improvements have enabled the railroads to become more profitable while at the same time providing service at lower rates. Despite increases in output and unit costs, railroads were able to reduce the resources required to move their freight by more than a third over the past 25 years. In 2008 and 2010, when compared with a base year of 1980, benefits from productivity improvements averaged just over $25 million, and approximately three-quarters of these benefits were passed on to rail customers in the form of lower rates.

Table 1: Rail Freight Productivity, 1966 to 2010

		1966	1980	1985	1990	1995	2000	2005	2008	2010
Output										
O1	Rail Rate Index	31.7	94.2	100	80.8	67.2	58.8	60.8	71.7	70.0
O2	Freight Revenue, Class I (Billions)	$9.3	$26.4	$26.7	$27.5	$31.4	$33.0	$44.5	$59.4	$56.3
O3	Output (O2/O1)	.293	.280	0.267	0.340	0.467	0.562	0.731	0.828	0.804
O4	Output Index 1985 = 100	84	105	100	127	175	210	274	310	301
Input										
I1	RR Cost Recovery Index	19.4	76.7	100.0	120	138	161	205	257	253
I2	Operating Expense, Class I (Billions)	$9.2	$26.4	$25.2	$24.7	$27.9	$29.0	$37.8	$47.4	42.7
I3	Input (I2/I1)	.474	.344	0.252	0.205	0.203	0.181	0.185	0.184	0.169
I4	Input Index, 1985 = 100	188	136	100	81	80	72	73	73	67
Productivity										
P1	Productivity (O3/I3)	62	81	106	166	230	310	396	450	476
P2	Productivity Index, 1985 = 100	58	77	100	156	217	293	374	425	450
P3	Productivity Index, 1995 = 100	27	35	46	72	100	135	172	195	207
P4	Annual Rate of Change	n.a.	2.0%	5.4%	9.4%	6.8%	6.2%	5.0%	4.2%	2.9%
Relative Decline in Rail Rates										
R1	Ratio of Price Index to Cost Index (O1/I1)	1.64	1.23	1.00	0.67	0.49	0.37	0.297	0.279	0.277
R2	Revenue per Ton-Mile, (Constant 1995$, cents)	4.99	4.86	4.03	2.99	2.40	2.07	2.13	2.49	2.45

Data Sources: AAR, *Railroad Facts,* Various Editions for Rows O2, I1, and I2. For Row O1, see endnote 3.

Table 2: Sharing the Benefits Since 1980: Higher NROI and Lower Prices (Current Dollars)

		1966	1980	1985	1990	1995	2000	2005	2008	2010
Impact on Railroad Profitability										
1	NROI (billions)	$1.04	$1.3	$1.7	$2.6	$2.9	$3.9	$6.1	$9.2	$10.0
2	NROI as % of Freight Revenue	11.2%	5.0%	6.5%	9.6%	9.1%	11.9%	13.7%	15.3%	17.7%
3	Additional NROI, 1980 base	N.A.	0	$0.4	$1.3	$1.3	$2.3	$3.8	$6.2	$7.1
Savings to Shippers										
4	Operating Revenues, 1980 rates (billions)	N.A	$27.3	$25.1	$32.0	$44.0	$52.9	$68.9	$78.0	$75.7
5	Savings to shippers due to lower rates (billions of current dollars)	N.A	0	-$1.5	$4.6	$12.6	$19.9	$24.4	$18.6	$19.4
Total Benefits										
6	Total Benefits Realized (Additional NROI plus Savings to Shippers, billions)	N.A	0	-$1.1	$5.8	$13.9	$22.1	$28.3	$24.9	$26.5
7	Additional NROI as % of Total Benefits	N.A	N.A	N.A.	11.5%	5.0%	7.1%	9.9%	25.1%	26.8%

SOURCES OF PRODUCTIVITY IMPROVEMENT

The previous section used an aggregate approach to documenting the magnitude of productivity improvements that have been achieved by railroads. That approach simply compared output to input, where output was measured as freight revenue adjusted by a rate index and input was measured as freight service expense adjusted by a cost index. The results from the aggregate analysis indicate that the industry made substantial gains, but do not explain how those gains were achieved. This section uses a more detailed methodology to identify the sources of the improvements in productivity. It examines three major sources of improvement, namely fewer service units per unit of output, fewer resources per service unit, and network rationalization. The objective of this section is to demonstrate that there are quantifiable sources of productivity improvement that together account for the aggregate benefits of more than $25 million per year that have been gained by railroads and their customers despite rising unit costs.

Fewer Service Units per Unit of Output

If there had been no changes in traffic mix, length of haul, equipment, tons/load, or trip distances, then output measured as freight revenue adjusted by a rate index would track changes in carloads, tons, and ton-miles. In fact, all of these factors changed substantially, allowing output to grow faster than any of the service units (Table 3).

Table 3: Indices of Various Measures of Railroad Activity (1995 = 100)

	1980	1984	1990	1995	2000	2005	2008	2010
Revenue Ton-Miles (billion)	70	71	79	100	112	130	136	130
Freight Train-Miles (million)	93	81	83	100	110	120	115	104
Freight Car-Miles (billion)	96	86	86	100	114	124	122	117
Carloads (million)	95	85	90	100	117	131	129	123
Tons (million)	96	92	92	100	112	123	125	119
Output	60	61	73	100	121	157	177	172

Data Source: service units for Class I railroads as reported in AAR, *Railroad Facts*, various editions.

The ratios of service units to output declined dramatically from 1984 through 2008 (Table 4). From 1984 to 1995, the steepest decline was in tons/unit of output; from 1995 to 2008, the steepest decline was in freight train-miles per unit of output.

Table 4: Ratio of Service Units per Unit of Output (1995=1.00)

	1980	1984	1990	1995	2000	2005	2008	2010
Revenue Ton-Miles (billion)	1.17	1.16	1.09	1.000	0.93	0.83	0.77	0.75
Freight Train-Miles (million)	1.56	1.32	1.14	1.000	0.91	0.76	0.65	0.60
Freight Car-Miles (billion)	1.60	1.42	1.18	1.000	0.94	0.79	0.69	0.68
Carloads (million)	1.59	1.40	1.24	1.000	0.97	0.84	0.73	0.72
Tons (million)	1.16	1.51	1.26	1.000	0.93	0.78	0.70	0.69

Data Source: service units for Class I railroads as reported in AAR, *Railroad Facts*, various editions.

To understand why the output measure grew faster than any of the service units, consider the differences between typical coal and intermodal trains in 2008:

- Coal train: 110 cars, each with 114 tons of coal; revenue was $1841/car or $202,510 per train.
- Intermodal train: 200 containers, with an average of 14 tons each; revenue was $1,013 per container or $202,600 per train.

These two trains had essentially the same revenue, so they would have had essentially the same output if 2008 were the base year for a productivity study. However, the coal train hauled 12,500 tons of coal in 110 carloads, whereas the intermodal train hauled only 2,800 tons of freight in 200 loads. When intermodal grows faster than coal, then there are fewer tons and ton-miles, but more loads per unit of output. Changes in traffic mix, equipment type, average length of haul, and other factors have also affected the service units required per unit of output.

Table 5 shows the total number of service units that would have been needed if the ratio of service units to output were the same as in 1995. By comparing the numbers in this table with the actual numbers of service units, it is possible to estimate how many more or how many fewer units were actually used. The annual savings were estimated by assuming unit costs of $5/1,000 revenue ton-miles (track expenses), $5/train-mile (train crew and locomotive expense), $0.06/car-mile (car maintenance and ownership), and $150/Carload (administration and overhead). These same unit costs were used previously to represent typical costs per service unit in 1995 (Martland 1999).[5] In 1995, the total annual savings relative to 1980 was $2.8 billion (Table 6). In 2010, the annual savings relative to 1995 was $7.1 billion. The savings are additive, so the savings in 2010 relative to 1980 was nearly $10 billion. The largest portion of these savings came from the reduction in RTM per unit of output, which is closely related to the rise in intermodal traffic. The next largest portion came from the reduction in carloads per unit of output, which is related to increases in car capacity. Savings in freight-train-miles reflect the operation of longer, heavier trains. The following sub-sections provide more detail concerning the changes in traffic mix, length of haul, and train capacity.

Table 5: Service Units at 1995 Level of Service Units per Unit of Output

	1980	1984	1990	1995	2000	2005	2008	2010
Revenue Ton-Miles (billion)	552	562	753	1306	1574	2046	2318	2250
Freight Train-Miles (million)	275	279	333	458	552	718	813	789
Freight Car-Miles (billion)	18.3	18.6	22.2	30.5	36.8	47.8	54.1	52.5
Carloads (million)	14.2	14.5	17.3	23.7	28.6	37.2	42.1	40.9

Table 6: Estimated Savings in Billions of 1995 $ (negative numbers indicate higher costs relative to 1995)

	1980	1984	1990	1995	2000	2005	2008	2010
Revenue Ton-Miles	-$1.4	-$1.8	-$1.4	$0	$0.5	$1.8	$2.7	$2.8
Freight Train-Miles	-$0.4	-$0.4	-$0.2	$0	$0.2	$0.9	$1.4	$1.6
Freight Car-Miles	-$0.4	-$0.5	-$0.2	$0	$0.1	$0.6	$1.0	$1.0
Carloads	-$0.7	-$0.9	-$0.6	$0	$0.1	$0.9	$1.7	$1.8
Total Savings	-$2.8	-$3.6	-$2.5	$0	$1.0	$4.1	$6.9	$7.1

Traffic Mix. One of the most important trends affecting aggregate measures of productivity was the general shift in traffic mix toward commodities best suited for transportation by rail (Table 7). In 1969, a third of rail carloads consisted of coal, metallic ore, grain or other bulk commodities that could move in multi-car shipments or unit trains. Intermodal carloads, which do not require local switching, were less than 6% of carloads in 1969. Motor vehicles and chemicals, which can be effectively moved in highly specialized rail equipment, were less than 10% of rail freight carloads in 1969. The majority of rail traffic moved almost entirely in general freight service that involved complex terminal operations, extensive intermediate classification work, and complicated operating plans. Over time, bulk commodities remained a third of total carloads, while intermodal, automobiles and chemicals together accounted for another third. The remaining general freight traffic declined from more than half of the total prior to 1970 to barely a third in 2010. Railroad operations are no longer dominated by traditional carload services.

Table 7: Traffic Mix: Carloadings by Major Commodity Group, Selected Years

Commodities	1969	1980	1995	2003	2008	2010
Suitable for Unit Trains (Coal, Metallic Ore, and Farm Products)	33.1%	39.8%	34.8%	30.8%	33.0%	33.1%
Intermodal[6]	5.5%	7.4%	15.2%	19.7%	20.7%	20.6%
Specialized (Motor Vehicles and Chemicals)	9.3%	10.0%	12.8%	12.5%	11.0%	11.5%
All Other (General Freight)	52.1%	42.8%	37.2%	35.2%	35.4%	34.9%

Data Source: traffic data from AAR, *Railroad Facts,* various editions.

Length of haul. There was a steady increase in the length of haul from 503 ton-miles/ton in 1965 to 615 in 1980, to 843 in 1995, and to 919 in 2008 (AAR, *Railroad Facts*, various editions). Increasing the average length of haul appears to improve system productivity for several reasons. First, there are fixed costs unrelated to distance associated with each trip: obtaining an empty car, filing a waybill, and obtaining payment as well as the relatively high costs of pickup and delivery. Increasing trip length simply adds additional car-miles moving in long, heavy trains over the high-density main line network. Hence rail costs per car-mile decline with distance, allowing a similar reduction in rates. According to a study conducted by the Surface Transportation Board (Office of Economics 2009), average revenue per ton-mile for coal and grain declined by at least 25% between 1985 and 2003 for each of three distance categories (less than 500 miles, 500-1,000 miles, and greater than 1,000 miles). For both commodities, the average revenue per ton-mile for long distance trips was at least 50% lower than the average revenue per ton-mile for short distance trips. Thus, even if there are no actual changes in rates or operations, an increase in average trip length is likely to result in decreases in both average cost and revenue per ton-mile.

Train Capacity. Train productivity (net tons/train-mile) improved by 1%-2% per year from 1980 to 2006 as a result of increases in train length and increases in the average loading density. Increases in siding length, modifications to freight yards, and higher capacity facilities for loading and unloading bulk trains enabled railroads to increase average train length. R&D efforts helped produce stronger track components that allowed the industry to increase axle loads, while investments that increased clearances along major routes allowed operation of double stack container trains. The ability to run longer trains of heavier cars resulted in very significant increases in train capacity and the amount of freight handled by main line trains.

Table 8 shows that the gross tonnage of a typical general merchandise train increased by two-thirds over this period.[7] General merchandise trains carry a mixture of loaded and empty cars, so the table shows the average gross tons per train rather than net tons. The 50% increase in tons/car reflected the introduction of 50' and 60' boxcars as well as the increase in axle load limits. The increase in train length reflects the combined effect of two factors. First, the maximum train length increased as a result of railroad investments to extend the length of sidings and to modify the layout of classification yards to allow them to receive and assemble longer trains. Second, as part of ongoing efforts to reduce crew costs and to make better use of line capacity, the railroads adjusted their operating plans so that the average train length was closer to the maximum train length for general merchandise trains.

Table 8: Characteristics of Typical General Merchandise Trains, 1980 and 2006

	1980	2006	% Change
Gross tons per car (average of load and empty)	61	92	50%
Train length	5100	6400 feet	25%
Gross tons per train	4500	7500 tons	67%

Data Source: see End Note 7.

Table 9 shows that the net tonnage carried by a typical loaded bulk unit trains increased by about a quarter between 1980 and 2006. These trains were already quite efficient in 1980, as their train length was generally close to the maximum allowable train length. Further increases in train length therefore reflected increases in siding length rather than a change in operating philosophy. The increases in tons/car reflected the shift from the 263,000 lb. GVW of the early 1980s to the 286,000 lb. GVW that was standard by the beginning of the 21st century. The increase in the maximum GVW allowed an extra 12.5 tons per car to be loaded in existing cars beginning in 1990. An additional boost in load capacity resulted from constructing new equipment with lower tare weight; a 286,000 lb. car with an aluminum rather than a steel body can carry more than 120 net tons, a 20% increase over what could be carried in the steel 263,000 lb. cars used in 1980.

Table 9: Characteristics of Typical Loaded Bulk Trains, 1980 and 2006

	1980	2006	% Change
Eastern Railroads			
Net tons per loaded car	80	86	7.5%
Train length	100	112	12%
Net tons per loaded train	8,000	9,700	21%
Western Railroads			
Net tons per loaded car	100	112.5	12.5%
Train length	100	112	12%
Net tons per loaded train	10,000	12,600	26%

Data Source: see End Note 7.

The improvement in line haul productivity was greatest for intermodal trains. In 1980, the typical intermodal train used 89-foot flat cars for trailer-on-flat-car (TOFC) or container-on-flat-car (COFC) service. Double stack trains, which were first introduced in the early 1980s and which can handle twice as many containers as a COFC train, are now widely used for domestic as well as for international intermodal freight.

Sources of Productivity Improvement: Fewer Resources per Service Unit

The next major source of productivity improvements came from improvements in resource utilization. Analyses are given in this section for fuel, freight cars, labor, and track, because these are areas that made major contributions to productivity improvement.

Fuel Efficiency. In 1966, when the industry moved 744 billion revenue ton-miles (RTM), their fuel efficiency measured was 188 RTM/gallon. By 1995, efficiency had increased to 373 RTM/gallon; and by 2010, when the industry moved 1,691 billion RTM, efficiency had grown another 29% to 484 RTM/gallon (AAR, *Railroad Facts*, various editions). These benefits largely resulted from better management of operations and better locomotive technology. Total fuel costs are a function of fuel price, traffic volume, and fuel efficiency:

(3) Fuel cost = (Price per gallon) x RTM/(RTM per gallon)

If fuel efficiency had been at 1995 levels throughout the entire period from 1966 to 2010, then less fuel would have been needed in the earlier years and more fuel would have been needed in later years. Fuel costs would have been $600 million higher in 2010, assuming the price per gallon remained at the 1995 level of $0.60/gallon (1691 billion RTM/373 RTM per gallon – 1691 billion RTM/484 RTM per gallon = 4.5 billion gallons at 1995 levels of efficiency minus 3.5 billion gallons at 2010 levels of efficiency = 1 billion gallons saved, which means that $600 million would have been saved if the average fuel price had still been $0.60 per gallon). The benefits of higher fuel efficiency were actually $2.3 billion in 2010 because of the increase in fuel price from $0.60 to $2.24/gallon ($2.3 billion = $600 million savings at $0.60 per gallon x $2.24 per gallon / $0.60 per gallon). However, these efficiency savings were more than offset by the extra $5.7 billion in fuel costs related to the increase in fuel prices (3.5 billion gallons x ($1.64 per gallon - $0.60 per gallon) = $5.7 billion).

The cost of fuel has a marked effect on overall rail costs. In 2008, when diesel fuel cost peaked at an average of $3.12 per gallon, fuel accounted for 19.9% of operating revenue. In 1980, when fuel costs were also high in real terms, fuel was 11.8 % of revenue. In 1995, when costs declined to the levels of the early 1970s, fuel costs were only 6.5% of operating revenues.

Freight Car Utilization. Several trends stand out concerning freight cars (Table 10). First, the fleet is much smaller than it was in 1980, despite the increase in carloads from 22.2 million in 1980 to a peak of 30.6 million in 2008. Second, ownership shifted from railroads, who now own less than half of the fleet, toward customers and car supply companies. Third, average capacity increased dramatically. There are several main reasons for these trends. Changes in demand favored the growth of intermodal shipments and bulk commodities moving in unit trains. Equipment used for these rapidly growing market segments is highly utilized, with cycle times less than 10 days per load. Cycle times for general freight equipment exceed 20 days, because of the additional time required for loading, intermediate classifications, unloading, and empty distribution.

Table 10: Characteristics of the Freight Car Fleet, Selected Years 1966-2008

	1966	1980	1985	1990	1995	2000	2005	2008
Freight Cars (millions)	1.826	1.711	1.422	1.212	1.218	1.38	1.317	1.393
Ownership								
Class I Railroads	81.5%	68.3%	61.0%	54.4%	47.9%	40.6%	36.0%	32.3%
Other Railroads	1.9%	6.0%	7.8%	8.6%	7.0%	9.6%	9.1%	7.8%
Shippers & Other	16.6%	25.8%	31.2%	37.1%	45.2%	49.9%	54.4%	59.8%
Utilization								
Annual Carloads per car	16.2	13.2	13.7	17.7	19.5	20.1	23.6	22.0
Average Car Cycle (car-days per carload)	22.5	27.6	26.6	20.7	18.7	18.1	15.4	16.6
Average Capacity (Tons)	61.4	79.4	84.3	88.2	89	92.7	97.2	100.5

Source of Data: AAR, *Railroad Facts*, various editions.

The more than ten-day reduction in the average car cycle time reduced car costs per carload by more than $200 between 1980 and 2008, assuming ownership costs of $20 per day.[8] If cycle times had remained at 1980 levels, annual equipment costs would have been $6 billion/year higher by 2008 ($200 per load x 30.6 million loads > $6 billion). Most of these savings reflect changes in traffic mix rather than improvements in cycle times for specific types of equipment, as studies of rail freight service have not documented any long-term trends toward improvements for general service freight since the 1970s (Martland and Alpert 2007).

The shift in ownership also relates to traffic mix. High-volume customers, such as electric utilities, can justify owning or leasing their own equipment. Other customers, notably chemical companies, prefer to own or lease equipment to ensure equipment quality and cleanliness. If railroads' share of the fleet had remained at 1980 levels, they would have owned an additional 474,000 cars that were actually owned by customers and car supply companies in 2008 (Table 11). At $20/car day, the annual costs associated with these cars would be $3.46 billion. Shippers point to this as an example of costs that they have incurred that offset some of the lower rates that they have received since 1980.

Table 11: Impacts of Improvements in Car Cycle Time and Changes in Freight Car Ownership

	1980	1985	1990	1995	2000	2005	2008
Car Cost/Carload ($20/car day)	$553	$532	$413	$375	$363	$309	$332
Annual savings related to reduction in average cycle time since 1980 (Billions)	$0.0	$0.6	$3.9	$4.8	$4.8	$6.4	$6.0
Additional Cars Owned by Car Companies & Shippers After 1980 (millions)		0.078	0.138	0.237	0.332	0.378	0.474
Cost shifted to Customers (Billions)		$0.57	$1.00	$1.73	$2.43	$2.76	$3.46

Source of Data: analysis based upon information in prior table.

The percentage of shipments that can use general purpose, free-running equipment is a declining portion of rail traffic. The change in traffic has affected the composition of the car fleet. Box cars declined from a third of the fleet in 1966 to a quarter in 1980 and a twelfth in 2008, largely because much of the traffic base that supported the use of box cars and other general purpose equipment disappeared (Lewis 2012).

Labor Productivity. As depicted above in Figure 1, total Class I railroad employment dropped 22% from nearly 600,000 in 1966 to 460,000 in 1980, a decline of 1.6% per year. Employment dropped another 64% to 164,000 in 2008, a decline of 1.8% per year, despite increases in output. The fact that railroads are highly unionized did not prevent this very significant reduction in the work force, and labor agreements, which were re-negotiated periodically, shared productivity gains with employees in the form of higher wages. In constant 2005 dollars, wages increased from $16.02 in 1968 to $26.40 in 2008, which is an average rate of 1.25% per year compounded over the entire 40-year period.

The number of employees in each major category dropped dramatically between 1968 and 2008 according to data published annually by the AAR (AAR, *Railroad Facts*, various editions). The largest drop—nearly 90%—occurred for professional and administrative employees, because of the tremendous advances in communications and computers. A 70% reduction in maintenance of way employees resulted from the introduction of better track materials and better maintenance equipment. Similar reductions were achieved for maintenance of equipment, reflecting the use of larger cars with better components and more modern maintenance facilities, but also the shift of fleet ownership away from the Class I railroads. The number of train and enginemen (i.e. engineers, firemen, conductors, and brakemen) declined 64% because of the loss of merchandise traffic, the reduction in yard activity, labor agreements that reduced crew consist and changed work rules, the operation of longer and heavier trains, and the shift of many light density lines and smaller terminals to short lines and regional railroads.

A measure of labor productivity can be obtained by relating the total number of employees in each category to the number of service units that are most closely related to that category. Table 12 shows indices of employees per service unit, with 1995 = 100. The improvement in employee productivity was even greater than the sharp decline in the number of employees. The index of executives, officials, and assistants per carload fell by more than 50% between 1980 and 2008. Professional and administrative employees per carload dropped by 90% largely because of advances in computers and communications, while maintenance employees per service unit, both for track and for equipment, dropped to a quarter of their 1980 level by 2008. Transportation T&E (train and enginemen) declined less rapidly than the other operating categories, particularly between 1995 and

2008. Other transportation, which includes yard employees, continued its steady decline throughout the entire 40-year period.

Table 12: Employees Per Service Unit, by Category (1995 =100)

	Service Unit	1968	1980	1995	2008
Executives, Officials and Assistants	Carloads (millions)	124	170	100	74
Professional & Administrative	Carloads (millions)	397	362	100	39
Maintenance of Way & Structures	Revenue ton-miles (billions)	390	300	100	65
Maintenance of Equipment & Stores	Car Miles (millions)	361	280	100	68
Transportation, Other than T&E	Train Miles (millions)	678	325	100	61
Transportation, T&E	Train Miles (millions)	277	228	100	93

Source of data: AAR, *Railroad Facts*, various editions.

The continuing improvements in employee productivity had a significant effect on the workforce and the payroll of the Class I railroads. The number of employees in any category can be shown as a function of workload and a measure of labor productivity:

(4) Employees = Service Units x (Employees per Service Unit)

The total payroll can be calculated as a function of the size of the work force, the average wage, and the benefits percentage (i.e., average benefits expressed as a percentage of the average wage):

(5) Payroll plus benefits = Employees x Average Wage x (1 + Benefits Percentage)

These two equations make it possible to estimate the benefits associated with improvements in labor productivity using wage rates and benefits percentage for each major category of employees for any desired base year. Table 13 shows rail employment by category for 1980, 1995 and 2008 along with two additional columns: 1) the number of employees that would have been required in 1995 assuming that labor productivity (service units per employee) had remained at 1980 levels for each category of employment and 2) the number of employees that would have been required in 2008 assuming that productivity had remained at 1995 levels.

If labor productivity had not improved, then total 1995 employment would have increased by 328,000 to 516,000, and total wage payments would have added $15.0 billion to railroad operating costs. These estimates are based upon the average wage rates and the changes in employment for each category of employees. The final row of Table 13 calculates the average savings per employee as the total savings in wages divided by the total reduction in employees.

Additional costs would also have been incurred related to railroad retirement and other employee benefits (which equaled 38.5% of wages in 2008). If labor productivity had remained at the 1995 levels until 2008, the workforce in 2008 would have included 236,000 people, adding $3.2 billion to the payroll. In other words, if labor productivity had not improved, the work force would have expanded to keep up with the increases in carloads, RTM, car-miles, and train-miles. By 2008, the added payroll would have been over $18 billion higher relative to 1980, and the total additional cost including benefits would have been on the order of $25 billion per year. And these numbers are all calculated using 1995 wage levels, when the average wage was $19/hour. In 2008, when the average wage was $30/hour, annual savings from labor productivity were $39 billion in current dollars.

Table 13: Savings Resulting from Improvements in Labor Productivity (Constant 1995 $)

	1980 Employees	1995 Employees, 1980 Productivity	1995 Employees	2008 Employees, 1995 Productivity	2008 Employees
Executives, Officials and Assistants	17,328	18,193	10,708	13,979	10,183
Professional & Administrative	92,780	97,411	26,940	36,522	13,637
Maintenance of Way & Structures	84,390	120,058	40,033	55,405	35,664
Maintenance of Equipment & Stores	99,614	103,724	37,106	45,866	30,659
Transportation, Other than T&E	29,141	31,184	9,597	11,373	6,664
Transportation, T&E	135,741	145,256	63,831	73,131	67,632
Total	458,994	515,826	188,215	236,277	164,439
Reduction in employees			327,611		71,838
Total Savings			$15.0 billion		$3.2 billion
Average wages saved per position			$46,000		$44,000

Data Source: AAR, *Railroad Facts*, various editions.

Improved Track Materials and Track Maintenance Technology. When 100-ton cars were introduced in the 1960s, the railroads reduced operating costs because more tons could be carried in each car and heavier trains could be handled without having to increase siding lengths or changing yard layouts. However, the heavier cars put much greater stress on the track structure at a time when many railroads were dealing with financial problems by deferring maintenance. As a result, track costs rose during the 1970s, spurring the industry toward the development of better rail, stronger track structures, and more effective maintenance practices, which ultimately led to longer rail and tie lives and reduced maintenance requirements (Table 14). Despite the fact that wheel loads increased again in 1990, the tons of rail replaced and the number of new ties installed declined. Using approximate 1995 costs of $300,000 per mile of rail and $40 per tie,[9] the annual cost of track maintenance dropped by 50% between 1980 and 1995. The cost per 10,000 revenue ton miles dropped even more. After increasing 18% between 1966 and 1980, costs declined 65% by 1995, dropping from $39.13 to $13.61. If rail and tie replacement had continued at the 1980 level, costs would have been $2.3 billion higher in 1995 and $2.1 billion higher in 2010. For more detailed information concerning track technology, track costs, and track productivity, see Chapman and Martland (1997 and 1998).

Table 14: Track Maintenance Costs for Materials and Equipment (estimated 1995 cost of $300,000 per mile of rail and $40 per tie)

	1966	1980	1995	2008
Revenue Ton-Miles (billion)	738	918	1306	1777
Maximum Gross Vehicle Weight (1000 lbs.)	263	263	286	286
Estimated Cost for Rail and Tie Programs (1995 $)	$2.5 billion	$3.6 billion	$1.8 billion	$2.2 billion
Cost/10,000 RTM (1995 $)	$33.22	$39.13	$13.61	$12.53
Extra Cost Relative to 1995 (1995 $)	$1.5 billion	$2.3 billion	0	($0.2 billion)

Data Sources: C.D. Martland, P. Lewis, and Y. Kriem, 2011.

Sources of Productivity Improvement: Network Rationalization

Line Abandonment. The U.S. rail network has been shrinking since the 1920s, when railroads began to shed underutilized lines and restructure their systems to meet changing demands for rail transportation. In 1950, there were approximately 225,000 miles in the U.S. Rail System. Over the next 60 years, 65,000 route-miles were abandoned and another 19,000 were converted to rail trails, leaving 140,000 route-miles in operation in 2009 (Table 15).

Table 15: Changes in the Size of the U.S. Rail System 1950 to 2010

Decade	Initial Route Miles (1000s)	Miles Abandoned During Decade	Route-Miles, End of Decade (1000s)	Route-Miles Operated, End of Decade (1000s)	Open Rail Trails (1000s)
1950-59	225	8,776	216	N.A.	N.A.
1960-69	216	12,640	203	N.A.	N.A.
1970-79	203	19,770	183	N.A.	N.A.
1980-89	183	18,920	165	158	N.A.
1990-99	165	4504	160	144	N.A.
2000-09	160	315 (2000-2004)	159	140	19

Source of data: miles abandoned from ICC and STB Annual Reports, as summarized by Ozment and Spraggins (2008); route-miles calculated by subtracting miles abandoned from initial route miles; route-miles operated from summary of miles operated by state in *Railroad Facts*, 1990, 2000, and 2010; rail trails from website of Rail-to-Trails Conservancy website (accessed July 7, 2011).

When lines are no longer operated, abandoned, or converted to rail trails, the minimum annual maintenance to keep a line operational is no longer needed. The greatest savings will usually relate to ties, because ties deteriorate no matter how few trains are run. There are on the order of 3,000 ties per mile on a light density line; if 100 are replaced each year at a cost of $60 each for labor and materials, the annual cost would be $6,000 per mile. The annual savings in tie replacement and all other routine maintenance activities would be in excess of $10,000 per mile. If unprofitable operations were ceased, annual savings could be twice as large. For example, a study of abandonment applications between 1951 and 1960 found average annual savings of $4,600 in 1973 dollars, which would be $18,000 in 2009 dollars (Sloss, Humphrey and Krutter 1975). Using this estimate of savings per mile abandoned, the benefits to the rail industry of eliminating 70,000 miles of light density lines would exceed $1 billion per year.

Short Line and Regional Railroads. Between 1980 and 2008, the Class I railroads transferred nearly 30,000 route-miles to short-line and regional railroads, many of which were newly formed as a means of maintaining service on these lines. By the end of this period, approximately 550 small railroads operated a third of the industry's route-miles, originated or terminated more than a third of the rail industry's general merchandise traffic, and earned more than $3 billion in revenue.

Table 16: Route-Miles Operated

Category	1967	1980	1987	1995	2008
Class I	198,603	164,822	147,568	125.072	94,082
Regional			15,100	18,815	16,690
Local			14,534	26,546	28,554
Switching & Terminal			4,011		
Total Non-Class I	11,223	18,255	33,645	45361	45244
Total	209,826	183,077	181,213	170,433	139,887
Non Class I as % of Total	5%	10%	19%	27%	32%

Source of data: *Railroad Facts*, 1982, p. 42 (totals for 1967 and 1980); *Railroad Facts*, 1984, p. 42 (Class I mileage for 1967 and 1980); *Railroad Facts*, 1988, 1996, and 2009.

Mergers. The number of Class I rail systems declined from more than 40 during the 1970s to just seven in 2010. The role of mergers in improving productivity is difficult to quantify, and the most recent major mergers all resulted in lengthy periods of confused operations and poor customer service. Nevertheless, mergers create larger railroads with more opportunities for line and terminal consolidation, and they facilitate comprehensive, integrated planning for operations and investments. With a broader perspective concerning network structure and the ability to route traffic over multiple routes, a large railroad can implement coordinated marketing, operating, and investment strategies, tasks that were much more difficult when multiple railroads owned small pieces of the network.

Terminal Consolidation and Transformation. During the 1960s and 1970s, many railroads invested in hump yards in order to reduce terminal costs, increase capacity and improve service. Most were built to handle 2,000 to 3,000 cars per day, and each allowed multiple small yards to be closed or downgraded. When general merchandise traffic collapsed in the early 1980s, this spurt of yard construction ended abruptly. Mergers of the Class I railroads made it feasible to eliminate redundant facilities, while changes in traffic mix and the competitive environment made it necessary to restructure old class yards as intermodal terminals.

SUMMARY OF PRODUCTIVITY IMPROVEMENTS

The rail industry in recent years has been more productive and more profitable than at any other time since World War II. More traffic moves over the system, rates are lower for most customers than they were in 1980, and resources are used more effectively. Mainline operations have been consolidated to the extent that seven Class I major railroads dominate the industry. Longer, heavier trains operate over tracks that handle traffic volumes unimagined in the 1970s.

Productivity improvements have been essential to the current prosperity of the industry. Changes in traffic mix, better railroad technology, vastly improved information technology and the growing role of short line railroads allowed Class I railroads to cut their workforce from nearly 600,000 in 1968 to 164,000 in 2008. The industry was able to achieve very substantial improvements in labor productivity despite the fact that it remained highly unionized. Changes in work rules and wage rates resulted from negotiations between management and labor, which limited the pace, but not the ultimate extent of improvements. The workforce shrank largely through attrition, and wages grew faster than inflation.

Between 1980 and 2010, productivity improvements ultimately led to annual cost savings of approximately $60 billion per year. The greatest savings came from two independent sources[10]:

- Reductions in service units per unit of output. If the traffic mix and operating conditions in 2010 had been the same as in 1980, there would have been many more RTM, train-miles, car-miles, and carloads, and annual operating costs would have been $10 billion higher.
- Reductions in resources per service unit ($49 billion per year). If resource utilization had remained the same in 2010 as in 1980, the industry would have needed many more employees, much more diesel fuel, more freight cars, more rail and more ties, and annual operating costs would have been $49 billion higher.
    - Labor productivity measured as service units per employee for major categories of rail employment ($39 billion per year)
    - Fuel efficiency, measured as RTM per gallon ($2 billion per year)
    - Equipment utilization, measured as average car-days per load ($6 billion per year)
    - Track costs (2 billion/year)

Additional savings came from network rationalization, including $1 billion annually related to the reduction in the number of track-miles operated. The total savings, which reached $60 billion per year by 2010, are indeed sufficient to explain how the Class I railroads were able to increase their profitability even though they faced rising unit costs for labor and materials and falling rail rates because of increased competition within the freight transportation market.

The great majority of the productivity benefits went to rail customers. Until recently, railroads retained well under 20% of the benefits in terms of increases in net railway operating income. Customers benefited because rates not only did not keep pace with inflation, they actually declined. In constant dollars, there was more than a 50% decline in average revenue per ton-mile from 1980 to 2010. If average revenue per ton-mile had kept pace with inflation, then rail revenues would have increased by more than $50 billion in recent years. However, even following several years of rail rate increases related to rising fuel prices, rail rates remained lower than they were in the early 1980s. Average revenue per ton-mile was 3.18 cents in 1981, the first full year after the Staggers Act; between 2006 and 2010, average revenue per ton-mile was just 3.12 cents. Based upon the Surface Transportation Board's rail rate index, rail customers saved $20-$25 billion per year in recent years because rates were lower than they were in the early 1980s.

Productivity improvements have helped the railroads to overcome the difficulties of the 1970s and to rationalize their network. In the 1970s, the prospect of rail line abandonments caused controversy among railroads, their customers, and state and local governments. Federal legislation made it easier for railroads to abandon lines, but at the same time encouraged the revitalization of remaining light density lines by transferring them to short-line or regional railroads, many of which were created after 1980. Today, the remaining light density lines are in better physical condition, they have managers focused on local conditions, and most enjoy prospects for further growth.

The traffic base of the railroads changed dramatically over the past 50 years. In 1968, general freight traffic dominated the system, unit trains of coal and grain were secondary, and piggyback service was negligible. Over time, intermodal traffic became much more important, while traditional single-car shipments declined; both trends reflect the globalization of the economy. Most intermodal traffic consists of international containers, a market more suited to railroad capabilities than general domestic freight, which has steadily shifted to truck. Small shipments, shorter-distance shipments, and shipments from customers with smaller annual volumes are now much less likely to move by rail.

Productivity improvements in the rail industry can be traced to specific technological, managerial, or institutional causes. For example, labor productivity improved largely because of technological and institutional innovations. The widespread use of computers and communications allowed reductions in clerical and professional staff. Labor negotiations led to reductions in crew consist, changes in the basis of pay, and more flexible operating rules. Research by the railroads, suppliers, and the federal government led to better materials and maintenance techniques that reduced the need for track maintenance; new track machinery reduced the time and labor required to maintain and upgrade track.

Resource utilization improved largely because of technological and operational innovation coupled with investment decisions concerning railroad equipment and infrastructure. Suppliers built more powerful, reliable, and fuel efficient locomotives. Railroads and suppliers strengthened the track structure to allow heavier loads and introduced larger, more efficient freight cars. Double-stack container trains reduced the cost of intermodal transportation by nearly half, thereby making intermodal much cheaper than trucks for moving containers. Railroads, at times with financial support from public agencies, lengthened sidings, increased clearances and modified yard layouts in order to allow longer trains and larger freight equipment.

CONCLUSION

By the mid-1990s, the long, slow process of rationalization was essentially complete. Rationalization eliminated hundreds of unnecessary lines and terminals that were left over from the 19th century, when railroads dominated intercity transportation. The underutilized facilities and unprofitable branch lines that plagued the railroads during the Northeast Rail Crisis of the 1970s were gone or revitalized. The system was smaller, but it had more capacity and it was better suited to the long-haul, high-volume traffic that moves most efficiently by rail. The organizational structure of the industry was stronger, and the largest railroads offered single-line service to vast regions of the country. Numerous small railroads continued to serve thousands of customers on light density lines and offered sites for industrial development that had competitive access to multiple Class I railroads.

Although the Staggers Act is frequently cited as the dominant factor causing productivity improvements in the rail industry, this act was only one element of federal actions taken to buttress the rail industry during the 1970s and 1980s. Legislation resulting from the Northeast Rail Crisis created Amtrak and Conrail, helped accelerate the rationalization process, and introduced more flexibility in pricing, including contract rates. The U.S. Department of Transportation constructed the Transportation Test Center in Pueblo, supported a great deal of research related to track quality and railroad safety, and worked with the railroads and suppliers to build the research underpinnings of a new generation of track materials and maintenance techniques. While Staggers provided much more pricing freedom, it led to two decades of declining prices. Staggers was undoubtedly good for most rail customers, who enjoyed lower rates, but the railroads survived only because of their success in improving productivity.

Many of the rail industry's actions that led to productivity improvements simply continued efforts that were begun well before 1980. Railroads have been pursuing mergers, facility consolidation, and other network rationalization activities since the 19th century (Locklin 1966). The R&D efforts that enabled improvements in track and equipment also continued ongoing efforts since the 19th century. The labor agreements that allowed crew consist reductions for essentially all the Class I railroads by the early 1990s culminated negotiations that began in the 1950s, that resulted in major work stoppages in the 1960s, and that only began to be resolved in the mid-1970s when a few small railroads agreed to share the potential cost savings with their employees (Martland 1983).

Information technology was another vital factor that had nothing to do with deregulation. The potential of computers was evident in the 1960s, when railroads began introducing information systems that ultimately eliminated rooms full of clerical employees, enabled consolidation of offices, and extended planning and control capabilities. Railroad managers deserve credit for recognizing the potential of computers and communications for managing their systems.

In the future, continued emphasis on productivity improvement could have negative implications related to capacity and public policy. In 1980, the rail industry was far more concerned about over-capacity than under-capacity, and railroads could advance by cutting back. Today, railroads must invest to increase capacity. Between 2006 and 2010, the Class I railroads invested more than $9.4 billion per year in track and equipment, which was 18% of their total freight revenue. Whether or not the industry will continue to be willing and able to make such investments remains to be seen, as railroads may continue to focus on traffic best suited to rail, leaving a growing volume of shorter haul, lower volume freight shipments on the highways.

Although rail's share of total intercity ton-miles recently returned to where it was in the late 1960s (42.7% in 2007 vs. 41.4% in 1967), truck's share increased much more rapidly, from 22% to 31% over that same period (BTS 2011). Many public agencies would like to reverse this trend by moving freight from truck to rail, which will require a coordinated effort involving railroads, suppliers, and public agencies. John Horsley, executive director of the American Association of State Highway and Transportation Officials, summarized the challenges in his foreword to his organization's assessment of investment needs for the rail freight industry:

> *Given the forecasts of substantial increases in freight over the coming years, it will be a challenge for the freight-rail industry to maintain its share of freight movement, and an even greater challenge to increase it.* (AASHTO 2003, p. i)

Endnotes

1. The Regional Rail Reorganization Act of 1973 led to the creation of Conrail and the Regional Revitalization and Regulatory Reform Act of 1976, eased abandonment procedures, allowed rate flexibility, authorized contract rates, and also required intensive studies of the fundamental causes of and potential solutions to the "railroad problem."

2. This paper estimates changes in productivity by comparing peak years to avoid interpreting factors related to the business cycle as changes in productivity. Peak years, based upon trends in profitability and return on investment, were 1966, 1980, 1995, and 2008. Data for 2010 for intervening years are sometimes included to give additional insight to trends in performance. Unless otherwise noted, data were obtained from *Railroad Facts*, published annually by the Association of American Railroads.

3. The STB rail rate index was used where possible (Office of Economics, Environmental Analysis & Administration, 2009). For earlier and later years, the index was extrapolated using the change in the average constant dollar revenue per ton-mile as reported by the AAR in *Railroad Facts.*

4. Some of these savings were offset by the fact that substantial costs related to equipment were transferred to customers, some of whom had to invest in higher-capacity loading facilities to take advantage of unit train rates.

5. These unit costs per service are similar to the unit costs used in various studies conducted by MIT Rail Group in cooperation with the AAR and individual railroads during the mid-1990s. They were intended to represent the costs associated with moving freight over the main-line railroad system in the United States and Canada circa 1995. As in the much more detailed

service unit costing used in regulatory proceedings, double-counting is avoided by allocating specific categories of expenses to specific service units. For example, in the simplified approach used in this paper, a) equipment maintenance costs are allocated to car-miles and equipment ownership costs to fleet size, b) maintenance of way labor costs are allocated to ton-miles, while materials and equipment cost are allocated to the level of rail and tie replacements, and c) crew costs are allocated to train miles, while fuel costs are allocated to ton-miles. Some costs are left out of this simplified approach, but none are double counted.

Because of the rail industry's productivity gains, unit costs per service have been fairly constant despite marked productivity growth. The STB publishes a quarterly index of railroad inflation (the Rail Cost Adjustment Factor or RCAF) that includes both an unadjusted measure that is equivalent to the AAR's Railroad Cost Recovery Index (RCR) and a measure that is adjusted for productivity. Both the RCAF and the RCR are published annually by the AAR in *Railroad Facts*. The unadjusted RCAF rose 104% from 1987 to 2010, which is similar to the 117% increase in the RCR over that period. When adjusted for productivity improvements, the RCAF fluctuated, but the trend was stable, averaging 0.521 in 1987, 0.515 in 2008, and 0.520 in 2011. Thus, the costs per service unit in 1995, which are based upon 1995 costs and productivity, are believed to be representative of costs per service unit for the entire period from 1987 to 2011.

6. Intermodal carloads as reported by the AAR for 1969 and 1980; 60% of the total of "Miscellaneous Mixed Freight" and "Other" carloads for later years.

7. Source of data for Tables 8 and 9: C.D. Martland, "Productivity Improvements Related to Train Length and Tonnage," unpublished memo dated February 8, 2010. This memo documented typical train characteristics by synthesizing data from studies conducted for individual railroads, FRA, AAR, and TTCI between 1971 and 2007.

8. The cost of freight cars varies with the type of car; new box cars, gondolas, and hoppers acquired since 2000 have cost an average of $71,000; the average new car in 2010 cost $75,000 (AAR, *Railroad Facts* 2011). The car cost savings are based upon average ownership costs of 10% of the purchase price, rounded off to $7,300 per year or $20 per car-day.

9. These approximate numbers assume renewal of ballast and replacement of turnouts as well as replacement of the rail. The cost includes the cost of materials and track equipment, but not the cost of labor, which is captured elsewhere in the analysis.

10. The reduction in service units per unit of output and the reduction of employees per service unit are two different aspects of productivity improvement whose combined effects, along with changes in output and unit costs, will determine the magnitude of labor cost savings. For example, if train length increases, then fewer train-miles and fewer train crews will be needed, i.e., there will be a reduction in service units per unit of output. If the size of the train crew decline from more than four to less than three, and if the length of the average crew district increases, then there will be fewer train crew members per train-mile, i.e., there will be a reduction in employees per service unit. The labor costs associated with train crew members, or any other class of employees, is calculated as follows:

    (a) Payroll(t) = Employees(t) x Average Wage(t) (1 + benefit percentage(t))
    (b) Employees = Service Units(t) x Employees per Service Unit(t)
    (c) Service Units = Output(t) x Service Units per Unit of Output(t)

    Solving these equations for a base year and another year will show how much payroll increased, and it will be possible to determine the extent to which changes in any of the factors resulted in a change in total labor cost, i.e., in the payroll for this category of employees. In this paper, the savings related to service units per unit of output (i.e., equation (b)) are discussed in relation to Table 6, assuming 1995 labor productivity, wage rates, and unit costs per service unit; as discussed in Endnote 5, the costs per service unit have been stable going back at least to 1987. The savings related to employees per unit of output (i.e., equation (c)) are discussed in relation to Table 13, which shows savings based upon average wages for 1995. The paper then adjusts for the changes in wage rates to estimate current savings.

References

AASHTO. *Freight-Rail Bottom Line Report*. American Association of State Highway and Transportation Officials, 2003.

Association of American Railrods. *Railroad Facts*. Association of American Railroads, Washington D.C., various editions.

Bereskin, C. Gregory. "Railroad Economies of Scale, Scope, and Density Revisited." *Journal of the Transportation Research Forum* 48 (2), (2009): 23-38.

Bureau of Transportation Statistics. *National Transportation Statistics* 2011. U.S. Department of Transportation, 2012.

Chapman, J.D. and C.D. Martland. "Railroad Track Productivity: A Historical Perspective." *Transportation Quarterly* (Summer 1997): 105-118.

Chapman, Jeffrey D. and Carl D. Martland. "The Effects of Fixed Plant and Rolling Stock Technology on Railroad Track Costs." Presentation to the American Railroad Engineering and Maintenance Association (AREMA), Chicago, IL, Sept. 15, 1998.

Christensen Associates. *An Update to the Study of Competition in the U.S. Freight Railroad Industry.* Final Report for The Surface Transportation Board. L.R. Christensen Associates, Inc., Madison, WI, January 2010.

Gómez-Ibáñez, J. A. and G. de Rus. *Competition in the Railway Industry.* Edward Elgar, Northampton, MA, 2006.

Hamberger, Edward R. "The Anniversary of Our American Express Story." *The Huffington Post,* October 8, 2010.

Lewis, Paul. "Analysis of Rail Freight Traffic: Gaining a Better Understanding of the Industry and the Factors that Influence Traffic." *Proceedings, Transportation Research Forum.* Transportation Research Forum, Fargo ND, 2012.

Locklin, Philip. *Economics of Transportation.* Richard D. Irwin Inc, Homewood, IL, 6th ed, 1966.

Lowtan, Donovan. "Assessing the Impact of Regulation and Deregulation on the Rail and Trucking Industries." Thesis (MST). Massachusetts Institute of Technology, February 2004.

Martland, C.D. "Workload Measurement and Train Crew Consist Adjustments: The Boston & Maine Experience." *Transportation Journal* 21 (3), (1983): 34-57.

Martland, C.D. "Improving Railroad Productivity: Implications of U.S. Experience for Canadian Railroads." *Journal of the Transportation Research Forum* 29 (2), (1989): 427-436.

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.

Martland, C.D., "Productivity, Pricing and Profitability in the U.S. Rail Freight Industry, 1995-2004." *Journal of the Transportation Research Forum* 45 (3), (2006): 93-108.

Martland, C.D. and Steven Alpert, "Origin-to-Destination Performance for General Merchandise Traffic Moving to or from Short Line Railroads." *Journal of the Transportation Research Forum* 46 (3), (2007): 51-68.

Martland, C.D., P. Lewis, and Y. Kriem. *Productivity Improvements in the US Rail Freight Industry 1980 to 2010.* Final Report. MIT School of Engineering, August 2011.

Martland, C.D. and Gerard McCullough. *Improving Railroad Productivity*. MIT Center for Transportation Studies, 1984.

Morrison, S.A. and C. Winston. "Regulatory Reform of U.S. Intercity Transportation." Gómez-Ibáñez, J. A., W.B. Tye, and C. Winston, eds. *Essays in Transportation Economics and Policy*. Washington, D.C.: Brookings Institution Press (1999): 469-492.

Office of Economics, Environmental Analysis & Administration. *Study of Railroad Rates: 1985-2007*. Surface Transportation Board, January 16, 2009.

Ozment, John and Barry Spraggins. "A Study of Long Term Economic Impact of Rail Line Abandonments." *Journal of Academy of Business and Economics* 8 (4), (2008).

Secretary of Transportation. *A Prospectus for Change in the Freight Railroad Industry*. U.S. Department of Transportation, 1978.

Sloss, James, Thomas J. Humphrey, and Forrest Krutter, *An Analysis and Evaluation of Past Experience in Rationalizing Railroad Networks.* Studies in Railroad Operations and Economics 16. MIT Department of Civil and Environmental Engineering, 1975.

Acknowledgement

This paper summarizes some of the major results of a study of railroad productivity conducted at MIT (Martland, Lewis, and Kriem 2011).

***Carl D. Martland** retired from MIT in 2007, but continues to work with the rail industry and public agencies on research related to transportation systems. A senior research associate and lecturer in the MIT Department of Civil and Environmental Engineering, he was the director of the MIT Rail Group from 1978-2001 and the program manager of the Association of American Railroad's Affiliated Research Laboratory at MIT from 1983-2001.*

Martland has taught project evaluation, engineering systems design, freight transportation management, transportation systems analysis, and transportation demand & economics at MIT. The author of more than 120 papers and research reports, he recently published a textbook "Toward More Sustainable Infrastructure: Project Evaluation for Planners and Engineers." In 1991, he was a co-author of the paper that won TRF's Outstanding Paper Award; in 1989, 1990, 1991, 1993, and 1994 he won the Conrail Award for the Best Paper on Railroads presented to TRF. In 1997, TRF selected Martland as the recipient of the Distinguished Transportation Researcher Award "in recognition of his pioneering the planning and costing techniques that are now commonly used by many U.S. railroads; his research has aided the revitalization of America's railroads, improving their efficiency, productivity, and service quality."

Martland has been active in professional organizations, including TRF, INFORMS, and TRB. He was TRF's program vice president in 1984 and president in 1986; in 1991, he was selected to be the second recipient of the Herbert O. Whitten Award for lifetime service to TRF. In 1998, Martland served as the president of the Rail Applications Special Interest Group of INFORMS.

Stopping Behavior of Drivers at Stop-Controlled Intersections: Compositional and Contextual Analysis

by Mintesnot Woldeamanuel

This research examines how drivers conduct themselves at stop signs by looking at the effect of different compositional variables (socio-demographic attributes) and ecological variables (physical attributes that affect people's behavior) on drivers' decisions to make a complete stop, as required by law. Observational study was designed to collect data at different parts of an urban area, and the binary logit model is used for the analysis. The modeling results show that five variables (age of the driver, number of passengers in the vehicle, presence of law enforcement officers within a block radius, using headlights, and time of the day the trip took place) are statistically significant in explaining relationships between those variables and the stopping behavior of drivers.

INTRODUCTION

Stop signs at intersections, predominantly used in the U.S. and Canadian cities, play a crucial role in regulating the movements of traffic and pedestrians and, thus, avoiding traffic accidents. The Fatality Analysis and Reporting System of the U.S. National Highway Traffic Safety Administration (NHTSA) shows that there were 34,017 fatal crashes related to junction and traffic control devices in 2008, out of which 8% are at stop-controlled intersections (NHTSA 2010). According to the 2009 NHTSA Annual Report on Traffic Safety Facts, 36% of fatal crashes at intersections are at stop signs, whereas 33% are at traffic signals, 25% at no traffic control devices, and the remaining at traffic control devices labeled as "other/unknown." The same data source reports that injury crashes at intersections are also higher at stop signs than other traffic control devices (NHTSA 2011). A similar report revealed that there was an increase in urban accidents, although a 7.1% decline in intersection crashes were observed between 2009 and 2010 (NHTSA 2012a). However, statistical projection of traffic fatalities for the first half of 2012 shows that an estimated 16,290 people died in motor vehicle traffic crashes (not location specific). This represents an increase of about 9.0% as compared with the estimated 14,950 fatalities that occurred in the first half of 2011 (NHTSA 2012b).

Several other studies have reported the fact that many injuries and fatalities occurred at intersections (with signals or stop signs) (Campbell et al. 2004; Retting et al. 2003; NHTSA 2001; Van Houten and Retting 2001). As 50% of all urban accidents occur at intersections, studying drivers and pedestrians' behavior is of vital importance for public safety (Neuman et al. 2003). At unsignalized stop-controlled intersections, drivers who fail to stop or after stopping, proceed without looking for traffic on the major road, create a substantial crash risk (Van Houten and Retting 2001). According to a survey by the National Safe Kids Campaign (2003), nearly half of the 25,660 vehicles surveyed at intersections marked with stop signs violated the stop signs by not coming to a complete stop at intersections. The same study shows that more than a third of motorists rolled through the stop signs, whereas nearly a tenth of motorists did not even slow down for the stop signs.

There are various explanations to the question of why drivers are not complying with the stop signs, as required by law. Although ecological issues such as the built environments, stop sign visibility, and road design played a significant role, the compositional variables, such as drivers' behavior, explained in terms of carelessness, lack of attention, or unnecessary overconfidence in controlling their surroundings, and the driver's socio-demographic background, cause a failure to comply to the law of making a complete stop. The compositional variable may also include age, gender, and hand-held cell phone use while approaching a stop sign. There are studies that attempted to create relationships between the socio-economic backgrounds of the travelers, their trip making circumstances, and their stopping behavior. For example, Kishore et al. (2009) conducted an observational study to determine the percentage of vehicles completely stopping at stop-sign intersections. According to the results of the study, the greatest contributing factors that caused most drivers to completely stop were the presence of conflicting vehicle movement, followed by movement of vehicles, vehicle arrival sequences, and the driver's age group. The study also found that drivers, during off-peak periods, have a higher probability of not completely stopping than those during peak periods because of less conflicting movement (either pedestrians or vehicles from other directions).

Other studies identified speeding as an influencing factor. For example, according to a study by NHTSA (2004), speeding was the dominant factor in the vehicle fatal crashes in which the driver violated the traffic signal/stop sign, while inattention ranked second. Van Houten and Retting (2001) assessed several studies on the subject and documented that poor compliance at stop signs, characterized by failure to stop or to look adequately for oncoming traffic, improper lookout, and stop-sign visibility, to be the leading cause of crashes. Age also played a significant role in influencing the behavior of drivers at stop signs. A study by Preusser et al. (1998) shows that drivers aged 65 and above are 2.3 times more at risk of being involved in accidents at all-way stop sign intersections when compared with being 1.3 times more at risk in other situations. Drivers aged 85 years and older are 10.6 times more at risk at stop-sign intersections. Braitman et al. (2007), in their comparative analysis between groups of drivers ages 35-54 and drivers ages 70 and older, found that crashes where drivers failed to yield the right of way increases with age and occurred mostly at stop-controlled intersections. The stopping behavior of the drivers can also be influenced by pedestrian movements at intersections. According to the National Safe Kids Campaign (2003), motorists were more likely to stop when pedestrians were present. However, the same study shows that nearly a third of motorists violated the stop signs when child pedestrians were present. Nearly half of motorists violated the stop signs when no pedestrians were present. Drivers were more likely to stop for pedestrians who were crossing than for those waiting to cross, although a significant percentage of drivers did not come to a complete stop at intersections where pedestrians were crossing. Traffic volumes, urban settings, and the behavior of other drivers can also have an influence on the way drivers behave at stop signs (Keaya et al. 2009; Kishore et al. 2009).

This paper reports findings from an observational study of stopping behavior by including additional new explanatory variables that have not often appeared in previous studies, such as cell phone use, number of passengers in the vehicle, and presence of a law enforcement officer that could have an influence on the stopping behavior and decisions of drivers when approaching a stop sign at intersections. The study is based on a field observation in St. Cloud, Minnesota, and application of probabilistic models. Unlike previous studies, the model is based on utility theory. The paper aims to investigate the distractions that may inhibit St. Cloud drivers from making a complete stop in order to provide an insight on variables that are responsible for influencing stopping behaviors of drivers.

St. Cloud is located about 60 miles northwest of the Twin Cities (Minneapolis-St. Paul and metro area), the largest populated region in Minnesota. Located in three different counties and on the Mississippi River, it is located in the heart of central Minnesota. With a population of 63,000 and a metro population of 167,392 residents, St. Cloud is also the transportation hub of Central Minnesota. St. Cloud has high manufacturing employment relative to the rest of the state and the nation.

This paper is organized into five sections. Following the introduction, the second section presents the data used and the methodology. The third section explains the analysis results, focusing on the effect of compositional and ecological/contextual variables on the stopping behavior of motorists. Section four explains and discusses the results, and the final section offers concluding remarks and policy recommendations.

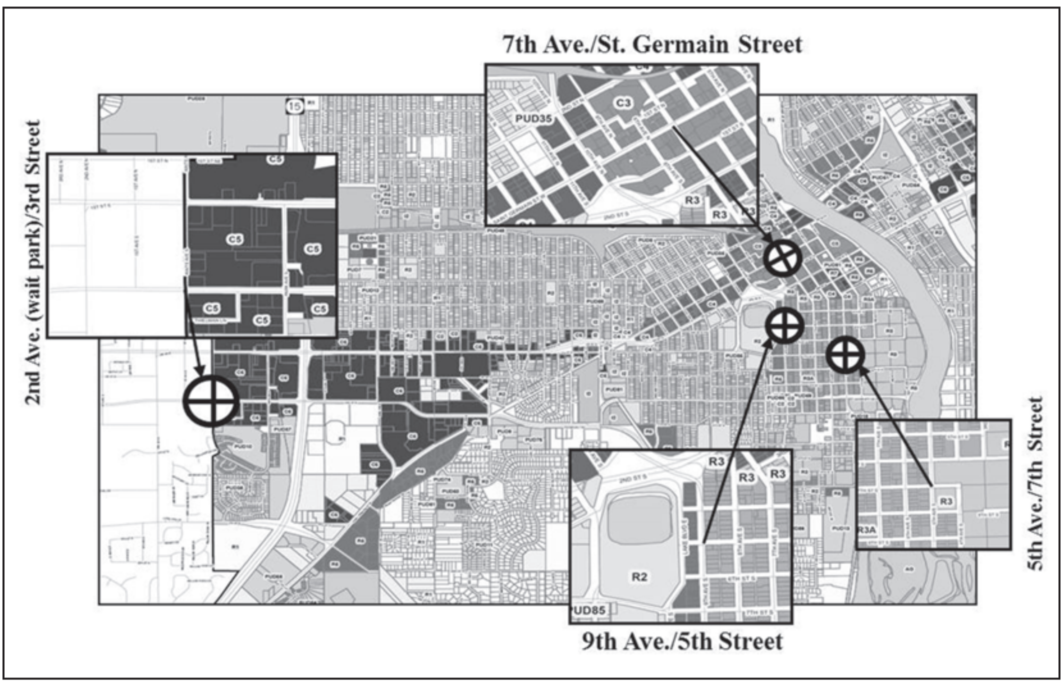
MATERIAL AND METHODS

The research approach involved modeling drivers' stopping behavior using data obtained from an observational survey. Observations were undertaken from March 25 to March 30, 2010, to understand drivers' behavior when they approach a two-way or four-way stop-controlled intersection and produce certain stopping types (complete, rolling, or no stop)[1]. A binary logit model was used to explain the relationship between the selected explanatory variables (compositional and contextual) and stopping types. The premise of the selection of variables is that behavior is a function of the individual's characteristics as well as contextual/environmental variables. Therefore, the independent variables were divided into two major categories: compositional (such as gender, age, number of passengers in the car, cell phone use while driving) and ecological (such as urban characteristics of the area [density and land use], day and time of the observation). A regulatory variable (availability of law enforcement) and a mechanical variable (headlight) are also added as independent variables. Indicator coefficients were estimated using the model in order to examine whether those variables really affect the decisions of the driver when approaching intersections regulated by stop signs.

Data Collection

All data were gathered at four different intersections: two two-way stops and two four-way stops. All intersections were located in different urban settings (density and land use). These included high-density commercial, medium-density residential/university, low-density commercial, and low-

Figure 1: Observation Locations



Source of map: City of St. Cloud

density residential. Intersections observed include that of 7th Avenue & St. Germain Street (high-density commercial; a four-way intersection), 5th Avenue & 7th Street (medium-density residential/University; a four-way intersection), 2nd Avenue & 3rd Street (low-density commercial; a two-way intersection), and 9th Avenue & 5th Street (low-density residential; a two-way intersection). All of the observed intersections are located within the municipal boundaries of the City of Saint Cloud, Minnesota. To gather a consistent set of data, all intersections were observed 12 times for a total of 2,400 vehicle observations (50 vehicles per observation; 600 observations per intersection). In order to ensure the consistency of observations, the data collectors were given a set of guidelines for each variable. Observations with any kind of ambiguity were eliminated from the sample.

To examine whether there is a relationship between *the days of the travel* and stopping type, weekday and weekend observations were conducted. Times chosen for observation on weekdays were 8:00 A.M., 1:00 P.M., and 6:00 P.M. These times were selected to examine whether there are varying stopping behaviors at different times of the day. The weekend observations included 12:00 P.M., 6:00 P.M. and 12:00 A.M. as we believed that such times would be necessary for observation, because most motorists do not work on the weekend and, therefore, start their day later and stay out for extended hours.

Variables

The dependent variable is the observed stopping types at intersections. A driver's stopping type is classified into three different types. The first is a *complete stop,* which would mean that the automobile would reach a velocity of exactly zero mph. The second type of stop, and most common among most motorists, was that of a *rolling stop,* which would mean that the automobile was traveling *about* a rate of five mph or less and moved forward without making a complete stop. The third stopping type is that of a *no stop,* which, to happen, would require the motorists be traveling at above five mph past the stop sign. The descriptive statistics showed that out of the 2,400 vehicle observations, 35% of the drivers made a complete stop, whereas 65% of them did not comply with the law of making a complete stop (52% made a rolling stop and 13% did not make any stop at all).

Table 1: Descriptive Statistics of the Dependent Variable

The Dependent Variable	Cases	Percentage
Stopping types		
0- No stop/rolling stop	1570	65%
1- Complete stop	830	35%

The independent variables were chosen to include the following. This study focuses on variables whose influence on the drivers' decision to make or not to make a complete stop is not well tested in previous research endeavors on the subject.

Compositional Variables:

- Gender: a binary variable with 1 = female and 0 = male
- Age: taking into consideration the difficulty of recording age in an observational study, the variable is categorized into three ordered variables based on the observer's judgment; 1 = young, 2 = middle-age, 3 = old. "Young" is considered to be less than 30 years old, "middle-age" is between 30 to 60 years old and "old" is more than 60 years old. If the age of the driver could not be determined during nighttime, the observation was deleted from the sample. However, this rarely occurred since all the intersections observed have very good street lighting.
- Cell phone use: a binary variable of 1 if the driver is using a *hand-held* cell phone while approaching the stop sign and 0 otherwise

Table 2: Descriptive Statistics of the Independent Variables

<table>
<tr><th rowspan="2">Independent Variables (predictors)</th><th colspan="2">No stop</th><th colspan="2">Rolling stop</th><th colspan="2">Complete stop</th><th colspan="2">Total Cases</th></tr>
<tr><th>Mean</th><th>SD</th><th>Mean</th><th>SD</th><th>Mean</th><th>SD</th><th>Count</th><th>Percentage</th></tr>
<tr><td rowspan="3">Gender of the driver
1= Female
0= Male</td><td>0.47</td><td>0.50</td><td>0.48</td><td>0.50</td><td>0.46</td><td>0.50</td><td></td><td></td></tr>
<tr><td colspan="2">47%</td><td colspan="2">48%</td><td colspan="2">46%</td><td>1138</td><td>47%</td></tr>
<tr><td colspan="2">53%</td><td colspan="2">52%</td><td colspan="2">54%</td><td>1262</td><td>53%</td></tr>
<tr><td rowspan="4">Age of the driver
1= Young
2= Middle-age
3= Old</td><td>1.92</td><td>0.77</td><td>1.74</td><td>0.74</td><td>1.70</td><td>0.72</td><td></td><td></td></tr>
<tr><td colspan="2">34%</td><td colspan="2">44%</td><td colspan="2">45%</td><td>1031</td><td>43%</td></tr>
<tr><td colspan="2">40%</td><td colspan="2">39%</td><td colspan="2">39%</td><td>940</td><td>39%</td></tr>
<tr><td colspan="2">26%</td><td colspan="2">17%</td><td colspan="2">16%</td><td>429</td><td>18%</td></tr>
<tr><td rowspan="5">Number of passengers in the vehicle
0
1
2
3+</td><td>0.94</td><td>1.14</td><td>0.68</td><td>0.96</td><td>0.81</td><td>0.97</td><td></td><td></td></tr>
<tr><td colspan="2">51%</td><td colspan="2">57%</td><td colspan="2">50%</td><td>1288</td><td>54%</td></tr>
<tr><td colspan="2">19%</td><td colspan="2">26%</td><td colspan="2">27%</td><td>621</td><td>25%</td></tr>
<tr><td colspan="2">17%</td><td colspan="2">10%</td><td colspan="2">16%</td><td>305</td><td>13%</td></tr>
<tr><td colspan="2">13%</td><td colspan="2">7%</td><td colspan="2">7%</td><td>186</td><td>8%</td></tr>
<tr><td rowspan="3">Cell phone use at intersections
1=Yes
0= No</td><td>0.27</td><td>0.45</td><td>0.23</td><td>0.42</td><td>0.24</td><td>0.43</td><td></td><td></td></tr>
<tr><td colspan="2">27%</td><td colspan="2">23%</td><td colspan="2">24%</td><td>577</td><td>24%</td></tr>
<tr><td colspan="2">73%</td><td colspan="2">77%</td><td colspan="2">76%</td><td>1823</td><td>76%</td></tr>
<tr><td rowspan="3">Law enforcement within one block
1=Yes
0= No</td><td>0.09</td><td>0.29</td><td>0.05</td><td>0.22</td><td>0.11</td><td>0.32</td><td></td><td></td></tr>
<tr><td colspan="2">9%</td><td colspan="2">5%</td><td colspan="2">89%</td><td>2215</td><td>92%</td></tr>
<tr><td colspan="2">91%</td><td colspan="2">95%</td><td colspan="2">11%</td><td>185</td><td>8%</td></tr>
<tr><td rowspan="3">Headlight
1=On
0=Off</td><td>0.38</td><td>0.49</td><td>0.39</td><td>0.49</td><td>0.43</td><td>0.49</td><td></td><td></td></tr>
<tr><td colspan="2">38%</td><td colspan="2">39%</td><td colspan="2">43%</td><td>958</td><td>40%</td></tr>
<tr><td colspan="2">62%</td><td colspan="2">61%</td><td colspan="2">57%</td><td>1442</td><td>60%</td></tr>
<tr><td rowspan="3">Time of observation
1=Day time
0= Night time</td><td>0.57</td><td>0.50</td><td>0.50</td><td>0.50</td><td>0.47</td><td>0.50</td><td></td><td></td></tr>
<tr><td colspan="2">57%</td><td colspan="2">50%</td><td colspan="2">47%</td><td>1200</td><td>50%</td></tr>
<tr><td colspan="2">43%</td><td colspan="2">50%</td><td colspan="2">53%</td><td>1200</td><td>50%</td></tr>
<tr><td rowspan="3">Day of observation
1= Weekdays
0= Weekends</td><td>0.55</td><td>0.50</td><td>0.49</td><td>0.50</td><td>0.50</td><td>0.50</td><td></td><td></td></tr>
<tr><td colspan="2">55%</td><td colspan="2">49%</td><td colspan="2">50%</td><td>1200</td><td>50%</td></tr>
<tr><td colspan="2">45%</td><td colspan="2">51%</td><td colspan="2">50%</td><td>1200</td><td>50%</td></tr>
<tr><td rowspan="5">Urban setting
1= low density residential
2= low density commercial
3= medium density residential/univ.
4= high density commercial</td><td>2.34</td><td>1.13</td><td>2.58</td><td>1.10</td><td>2.45</td><td>1.14</td><td></td><td></td></tr>
<tr><td colspan="2">30%</td><td colspan="2">22%</td><td colspan="2">27%</td><td>600</td><td>25%</td></tr>
<tr><td colspan="2">18%</td><td colspan="2">29%</td><td colspan="2">22%</td><td>600</td><td>25%</td></tr>
<tr><td colspan="2">30%</td><td colspan="2">24%</td><td colspan="2">26%</td><td>600</td><td>25%</td></tr>
<tr><td colspan="2">23%</td><td colspan="2">26%</td><td colspan="2">25%</td><td>600</td><td>25%</td></tr>
</table>

Contextual/Ecological Variables:

- Number of passengers: This is a continuous value of the number of passengers observed; 0 is given if it is a "drive alone" situation
- Time: although there are four different times of day chosen for observation (morning, afternoon, evening, and night), for analytical convenience the nominal nature of the variable is converted into a binary variable of 1 = day time (morning and afternoon) and 0 = night time (evening and night)
- Day: although there are four different days of the week chosen for observation (Tuesday, Thursday, Saturday, and Sunday), for analytical convenience the nominal nature of the variable is converted into a binary variable of 1 = weekdays (Tuesday, Thursday) and 0 = weekends (Saturday and Sunday)

- Urban setting: to account for traffic volume on the major street, four zonal types (urban settings based on activity types and buildings per square meter) were selected for observation, and the variable is arranged in an ordered fashion based on density, i.e.,
  - 4= high density commercial (downtown area: a four-way intersection);
  - 3= medium density residential/university (residential neighborhoods around campus: a four-way intersection);
  - 2= low density commercial (areas in suburban shopping centers: a two-way intersection);
  - 1= low density residential: a two-way intersection

Regulatory Variable:

- Law enforcement is a binary variable of 1 if there is a police officer or patrol within one block radius from the stop sign and 0 otherwise

Mechanical Variable:

- Headlights: a binary variable of 1 if the driver used headlights and 0 otherwise. The use of headlights was observed during daytime and nighttime and in both cases, there were drivers with headlights on and off.

Model Structure: Binary Logit Model

The model to be estimated in this study is the effect of compositional, contextual, regulatory, and mechanical variables on stopping behavior of motorists. Since the dependent variable (stopping behavior) is a binary variable with 1 = making a complete stop and 0 = making no stop or rolling stop, a Binary Logit Model is chosen for the analysis. Binary models are widely used in economic, marketing, transportation, and other fields to represent the choice of one among a set of mutually exclusive alternatives. When drivers are faced with two choices [a choice of making a complete stop (i) over not making a stop (j)], the probability that j is equal to [1-P(i)]. The general form of the binomial logit model is:

(1) Prob [Y_i=1| making a complete stop] = Exp $(\alpha+\sum\beta_i x_i)$ / [1+ Exp $(\alpha+\sum\beta_i x_i)$]

The model application is based on the utility theory, which assumes that the decision maker's choice to stop or not to make a complete stop is captured by a value called utility (U). The decision maker selects the alternative in the choice set with the highest utility.

(2) $U = \alpha + \beta_i x_i$

Where β_i is the coefficient associated with the independent variables; x_i is the value of the independent variables; α is the constant estimated by the model (Greene 2000).

RESULTS

Table 3 reports the estimated utility coefficients (β) of the explanatory variables with their t-values and p-values as a test of statistical significance. Figures 2 to 5 show the estimated likelihood of making or not making a complete stop. The model fits the data set as the chi-square, and the log likelihood ratio are within acceptable range. Results show that five variables (out of the selected nine variables) have a likely influence on stopping behavior of drivers. The positive and the negative signs attached to the coefficients describe the functional relationships between the dependent variable (stopping behavior) and the independent variables. The null hypothesis of relationships between the dependent and independent variables was rejected with p-value above 0.05 for 95% confidence level. The marginal effect (Table 4) is also estimated to show the percentage increase in the independent variable either to increase or reduce the probability of making a complete stop by one percentage point. According to the modeling result, the gender of the driver, cell phone use, the day of observation, and urban setting have no significant influence on how drivers behave when approaching the stop sign. The remaining statistically significant variables are discussed in the following sections.

Table 3: Modeling Results

Independent variables	β	Exp (β)	Standard error	t-value	p-value
Constant (α)	-0.591		0.181	-3.272	0.001
Gender of the driver	-0.081	0.922	0.087	-0.933	0.351
Age of the driver	-0.133	0.875	0.059	-2.239	0.025*
Number of passengers in the vehicle	0.085	1.089	0.044	1.931	0.050*
Cell phone used while approaching intersections	-0.035	0.965	0.103	-0.343	0.732
Law enforcement within one block radius	0.754	2.125	0.156	4.830	0.000*
Headlight	0.211	1.235	0.094	2.253	0.024*
Time of observation	-0.222	0.801	0.092	-2.411	0.016*
Day of observation	0.053	1.054	0.092	0.577	0.564
Urban setting	0.040	1.041	0.042	0.970	0.332

Maximum Likelihood Estimates

Number of observations	2400
Log likelihood function	-1527.366
Restricted log likelihood	-1547.59
Chi-squared	40.44
Degrees of freedom	9

* statistically significant variables for 95% confident level

Table 4: Marginal Effects on Probability of Making a Complete Stop

Independent variables	β	Standard error	t-value	p-value
Constant (α)	-0.133	0.040	-3.303	0.001
Gender of the driver	-0.018	0.020	-0.933	0.351
Age of the driver	-0.030	0.013	-2.240	0.025
Number of passengers in the vehicle	0.019	0.010	1.932	0.053
Cell phone used while approaching intersections	-0.008	0.023	-0.343	0.732
Law enforcement within one block radius	0.170	0.035	4.831	0.000
Headlight	0.048	0.021	2.254	0.024
Time of observation	-0.050	0.021	-2.413	0.016
Day of observation	0.012	0.021	0.577	0.564
Urban setting	0.009	0.009	0.970	0.332

Age

As well documented in previous research papers, age is one of the demographic variables that could have an influencing effect on the stopping behavior of drivers. The modeling result in this study reflects the existence of age's influence on stopping behavior. The variable is statistically significant with a negative β coefficient (means a less than 1 exp(β) value). The negative coefficient associated with the utility value of variable age (β_{age}) explained that older drivers have a lower likelihood of making a complete stop. On the other hand, young drivers have the tendency of making a complete stop (Figure 2). The marginal effect -0.030 also shows that a 3% increase in age reduces the probability of making a complete stop by one percentage point.

Number of Passengers in the Vehicle

The number of passengers in the car greatly influences the stopping behavior of the driver. People behave differently when they are alone as opposed to when they are with people. When they are alone, there is an extended freedom in their mind to “break” the law. The result in this study shows that the likelihood of making a complete stop increases with the number of passengers in the car.

Presence of Law Enforcement Officers

Law enforcement presence is defined in this study as the availability of law enforcement officers within a *one block radius* of the intersections being observed (with the visual reach of the field observer). Drivers have more incentive to obey traffic laws when law enforcement officers or a police car is around. The modeling result in this study proves that a positive relation does exist between law enforcement presence and making a complete stop. It is found that if an officer of the law is present, motorists are far more likely to refrain from making a rolling or no stop. The probability estimate, as well as the descriptive statistics, (only 9% did not make a stop in the presence of the law enforcement) show that the likelihood of making complete stops increase in the presence of police officers (refer to Table 2 and Figure 4).

Headlights

Operating without headlights at night and during periods of low visibility is a common cause of traffic crashes. The result of this study indicated that there is indeed a significant influence of having headlights on for the stopping decisions of drivers. According to the modeling result, there is a positive relationship between the use of headlights and making a complete stop. This implies that the visibility of the surrounding area might have an effect on drivers’ judgment while approaching intersections. The probability estimate shows that drivers with no headlights on at appropriate times have a lower likelihood of making a complete stop. Figure 5 shows that there is a slight drop of the likelihood of drivers making a complete stop without their headlights on. It is worth noting that the use of headlights was observed during daytime and nighttime and in both cases there were drivers with headlights on and off. The distribution of observation was as follows: nighttime-on= 451 drivers; nighttime-off= 749 drivers; daytime-on=507 drivers; and daytime-off=693 drivers.

Time of the Day

The time of the day when the trip is occurring is believed to affect the stopping behavior of drivers. The result in this study shows that there is a notable functional relationship reflecting the effect of the time of the day on the stopping behavior of drivers. Time by itself may not have any effect. However, spatial activities (activities at different places at different times) and the driver’s disposition or reaction to activities could vary with time. Although there were four time categories used to observe drivers at stop-controlled intersections (morning, afternoon, evening, and night), for analytical convenience, the nominal nature of the variable is converted into a binary variable (1 = daytime and 0 = nighttime). The negative sign attached to the utility coefficient of the variable (β_{time}) shows that there is high probability of making a complete stop during nighttime.

Figure 2: Probability of Stopping vs Age

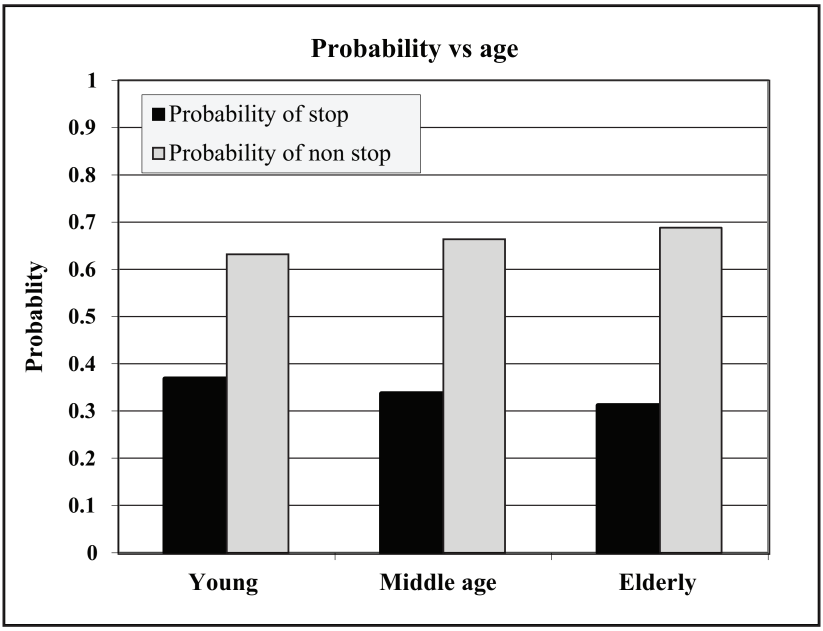


Figure 3: Probability of Stopping vs. No. of Passengers

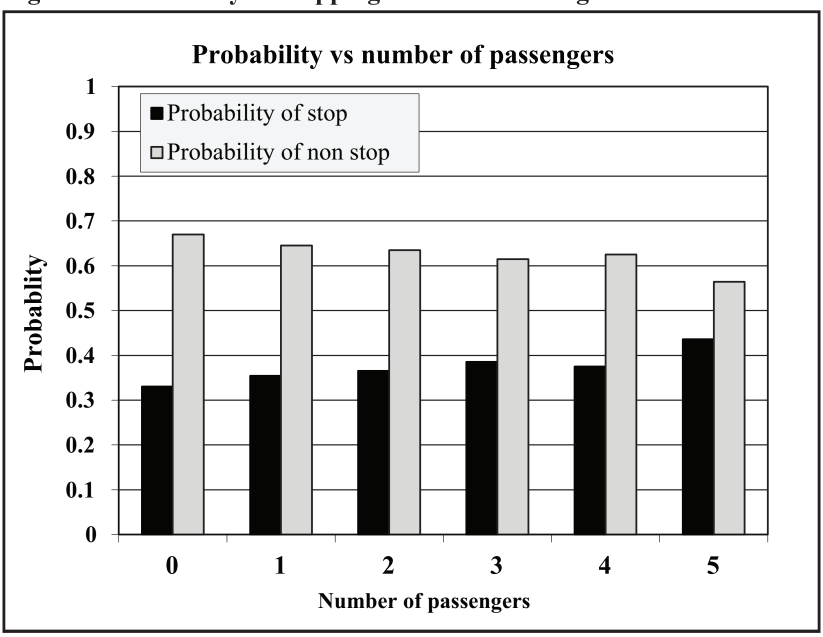


Figure 4: Probability of Stopping vs. Law Enforcement

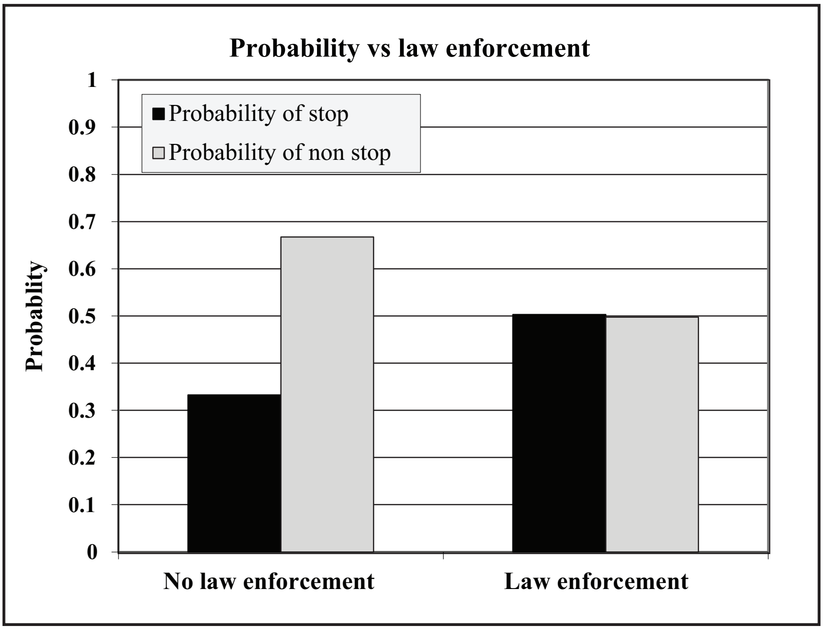


Figure 5: Probability of Stopping vs. Headlights

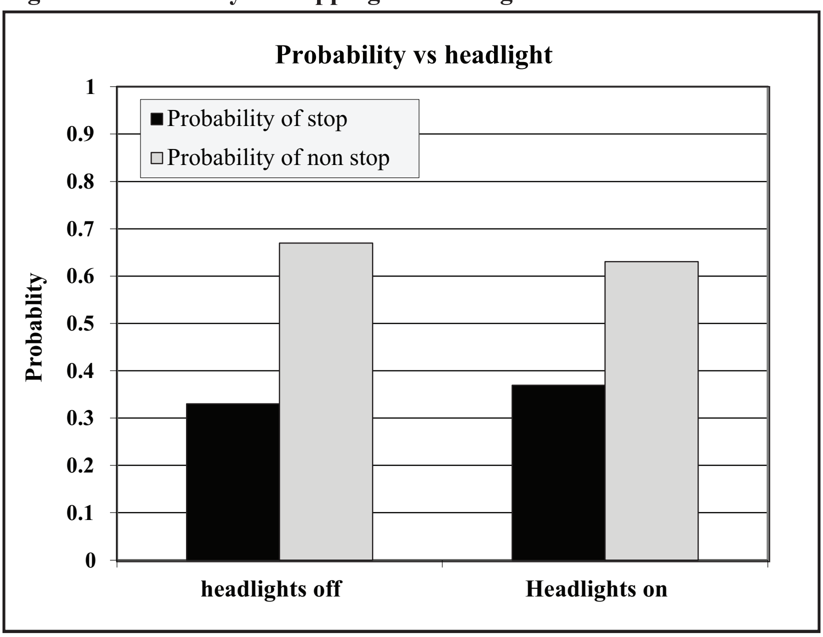
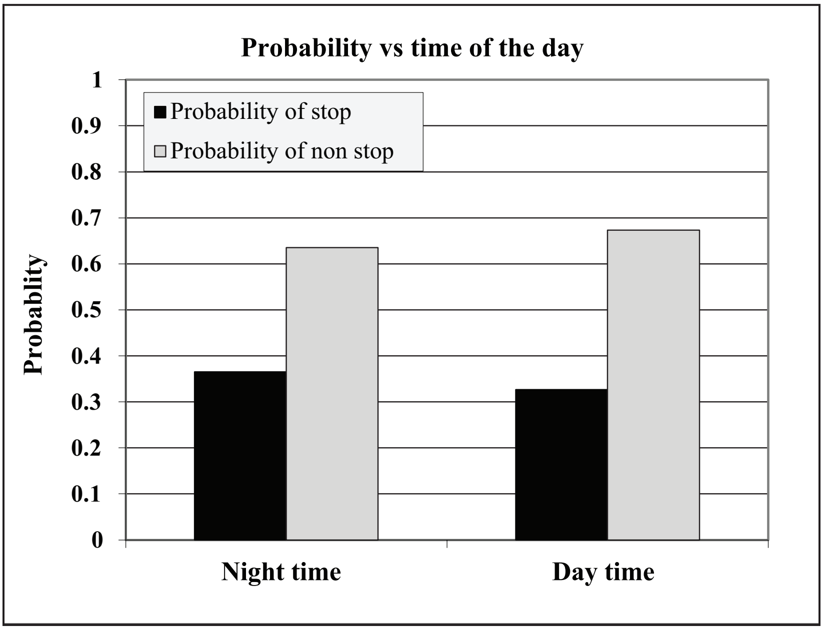


Figure 6: Probability of Stopping vs. Time of the Day



DISCUSSION

This study focused on examining compositional and contextual predictors such as demographic variables, time, availability of law enforcement, and number of passengers with the driver and their influence on the way drivers make decisions while approaching stop signs (whether to make a complete, a rolling, or no stop). The result yielded that age (among compositional variables), presence of law enforcement officers, the use of headlights during appropriate hours, time of the day when the driving takes place, and number of passengers in the car (contextual/ecological variables) have a significant influence on how drivers comply with the law of stopping at stop-controlled intersections.

Among the compositional variables chosen, gender seems to have no effect on the stopping behavior of drivers. The descriptive statistics showed that the majority of both male and female drivers make a rolling stop. On the other hand, age is found to be the compositional predictor to have a significant influence on the decision of the driver. Using the probability estimate equation 1, the likelihood of making a complete stop by a different age group was calculated and presented in Figure 2. The graph shows that the probability of making a complete stop decreases with age. The likelihood estimate on age shows that there is a 29% probability of making a complete stop by a driver of age 60+ compared with 38% likelihood by the young and 34% probability by the middle-aged drivers. The "age" parameter is an interesting predictor as stop sign visibility (size and color of the sign) could have an effect on older drivers. Other studies also indicated similar concerns. For example, Keay et al. (2009) described the old drivers' failure to stop at stop signs as a visual and cognitive failure. Braitman et al. (2007), in their comparative analysis between groups of drivers aged 35-54 and drivers aged 70 and older, found that crashes where drivers failed to yield the right of way increases with age and occurred mostly at stop-controlled intersections.

Number of passengers also showed an effect on the drivers' decision and behavior at stop-controlled intersections. Interestingly, the model results indicate people behave better when they are accompanied by passengers. Hand-held cell phone use, which was selected as an explanatory variable, hoping to see its effect on stopping behavior, has not exhibited significant influence.

Interestingly though, 24% of the observed drivers were using cell phones while approaching the stop-controlled intersections. However, their stopping behavior is proportionate to those who are not using cell phones while attempting to stop. A study on mobile phone use and traffic accidents suggested that braking reaction time is slower during a telephone conversation, so phone users come to a standstill closer to the vehicle in front of them, a stopping line, or an intersection (Dragutinovic and Twisk 2005).

Turning the headlights on at appropriate times is positively and significantly related to making a complete stop. While many cars have automatic headlights or daytime running lights (DRL) as a safety feature, there are thousands of cars on the street with manual headlights that drivers, at times, forget to turn on during appropriate times. DRL as a road safety measure is often difficult to understand for the road user because he or she "knows" that with sufficient attention every road user can be seen in daylight. Nevertheless, studies show that visual perception in daytime traffic is far from perfect and it is worse in conditions of low ambient illumination (where the natural light from the surrounding environment is obstructed). In a striking example, 8% of cars in an open field in broad daylight were not visible from relevant distances without the use of DRL (Hörberg and Rumar 1975, 1979; Allen and Clark 1964; Koornstra et al. 1997; Wang 2008). On shady roads or those with backgrounds which mask objects in the foregrounds, the visibility and contrast of cars in popular colors is greatly reduced. It is known from in-depth accident studies that failing to see another road user in time (or at all) is a contributing factor in 50% of all daytime accidents, and for daytime intersection accidents this increases to as much as 80% (Koornstra et al. 1997). According to the NHTSA (2008b), the passenger vehicle occupant fatality rate during nighttime is about three times higher than the daytime rate. This has been a concern for practitioners for a long time. Therefore, awareness programs and technological interventions (such as illuminating color marks that enhance visibility) are essential.

Although previous studies (such as Keaya et al. 2009; Retting et al. 2003) found that there are differences between high- and low-density areas when it comes to failure to stop at stop-controlled intersections, the variable is statistically insignificant in this study. However, the positive sign attached with the β coefficient shows that drivers residing in low-density neighbourhoods or rural areas were less likely to stop than those in high-density areas. Although there is better visibility in low density areas, it is believed that a higher failure to stop rate may be related to lesser traffic, more visibility around the stop-sign area, and a perception that it is safe to proceed through or turn in the intersection without stopping. Besides, lower density areas have a two-way stop sign, whereas the higher-density areas have a four-way stop sign, making drivers stop at four-way signs more that two-way signs.

While time of the day has a significant influence on what type of stop the drivers make, days of the week (whether it is weekdays or weekends), on the other hand, seems to have no influence on the stopping behavior of the drivers.

Although most of the variables mentioned above are conventional variables, there are three variables that are added value to this study, which give a new perspective on improving safety at stop-controlled intersection. Variables such as cell phone use, number of passengers in the vehicle, and availability of law enforcement officers, which are introduced in this study, have an influence on the stopping behavior and decisions of drivers when approaching a stop sign at intersections.

CONCLUSION AND RECOMMENDATIONS

This research presented the effect of some compositional and contextual variables on the stopping behavior of drivers at stop-controlled intersections in St. Cloud, Minnesota. The variables chosen to have an influence on the stopping behavior of the drivers include socio-demographic characteristics of the drivers (compositional) and the built environment-related (contextual/ecological) variables. Drivers' behavior at stop signs was investigated by using data gathered from observational studies.

The influence of different variables on drivers' decisions to make a complete stop, as required by law, or going against the law by making a rolling or no stop was examined. The binary logit model found that more contextual/ecological variables are statistically significant in explaining relationships between those variables and drivers' decisions to make or not to make a complete stop than the compositional variables. Variables that showed statistical significance (especially those new variables introduced in this study, such as number of passengers in the vehicle and availability of law enforcement officers) could be predictors for policy analysis and strategies to improve stop-controlled intersections and introduce drivers' safety awareness programs, thus reducing traffic accidents occurring at intersections. The presence of law enforcement officers and using headlights indicate that enforcement and safety measures could be an effective mechanism to reduce accidents at intersections. Police departments could use the results to deploy enforcement resources to the most accident prone intersections.

Supporting the finding in this study, a NHTSA report indicated that, for personal cars and light truck vehicles, Daytime Running Lights (DRL) would reduce injury crashes by 3.9 % (NHTSA 2008a). Therefore, increasing the current automatic DLR and phasing out manual DRLs would be considered as one of the policy options to decrease crashes at intersections.

Regarding cell phone use while driving, although the variable is statistically insignificant, the negative sign attached to it indicates that it is negatively correlated with making a complete stop. Thus, policies that restrict cell phones (unless drivers employ hands-free devices) are crucial as distracted driving is a factor in one out of four vehicle crashes in Minnesota, and text messaging and Internet use is outlawed for all drivers in the state. Therefore, policies that outlaw hand-held cell phone use, as other states do, would be essential to improve traffic safety.

Endnotes

1. Each observation was made at the intersection. Therefore, it is assumed that every driver has the same visibility distance from the point of observation.

References

Allen, M.J. and J.R. Clark. "Automobile Running Lights-A Research Report." *American Journal of Optometry* 331, (1964): 1-23.

Braitman, K.A., B.B. Kirley, S. Ferguson, and N.K.Chaudhary. "Factors Leading to Older Drivers' Intersection Crashes." *Traffic Injuries and Prevention* 8, (2007): 267–274.

Campbell, B.N., J.D. Smith, and W.G. Najm. *Analysis of Fatal Crashes Due to Signal and Stop Sign Violations*. U.S. Department of Transportation, DOT-VNTSC-NHTSA-02-09 DOT HS 809 779, Washington D.C., 2004.

Dragutinovic, N. and D. Twisk. *Use of Mobile Phones While Driving-Effects on Road Safety; A Literature Review.* R-2005-12 SWOV, Leidschendam, The Netherlands, 2005.

Greene, W.H. *Econometric Analysis*, 4th ed. Prentice-Hall, Inc., New York, 2000.

Hörberg, U. and K. Rumar. *Running Lights Conspicuity and Glare*. Rep. 178. Univ. of Uppsala, Uppsla, Sweden, 1975.

Hörberg, U. and K. Rumar. "The Effect of Running Lights on Vehicle Conspicuity in Daylight and Twilight." *Ergonomics* 22, (1979): 165-173.

Keay, L., J. Srichand, M. Beatriz, A.T. Kathleen, A.M. Cynthia, D.D. Donald, B. Kevin, Karen J. Bandeen-Rochee, W.G. Emily, and K.W. Sheila. “Urban and Rural Differences in Older Drivers’ Failure to Stop at Stop Signs.” *Accident Analysis and Prevention* 41, (2009): 995–1000.

Kishore, R., N. Kronprasert, and V. Reddy. *Observational Study for Determining the Percentage of Vehicles Completely Stopping at a Stop-Sign Intersection*: *Transportation Safety Final Report.* Group 6, Virginia Tech, National Capital Region, 2009.

Koornstra, M., F. Bijleveld, and M. Hagenzieker. *The Safety Effects of Daytime Running Lights; A Perspective on Daytime Running Lights (DRL) in the EU: The Statistical Re-Analysis And A Meta-Analysis of 24 Independent DRL-Evaluations as Well as an Investigation of Possible Policies on A DRL-Regulation in The EU*. R-97-36 SWOV, Leidschendam, The Netherlands, 1997.

National Highway Traffic Safety Administration. *Traffic Safety Facts, 2000*. DOT HS-809-337, Washington D.C., 2001.

National Highway Traffic Safety Administration. *Analysis of Fatal Crashes Due to Signal and Stop Sign Violations*. DOT-VNTSC-NHTSA-02-09, Washington D.C., 2004.

National Highway Traffic Safety Administration. *Traffic Safety Facts, Research Note*. DOT HS 810 963, Washington D.C., 2008.

National Highway Traffic Safety Administration. *The Effectiveness of Daytime Running Lights for Passenger Vehicles*. DOT HS 811 029, Washington D.C., 2008.

National Highway Traffic Safety Administration. *Fatality Analysis and Reporting System (FARS)*. http://www-fars.nhtsa.dot.gov/Crashes/CrashesLocation.aspx, U.S. Department of Transportation. Accessed July 6, 2010.

National Highway Traffic Safety Administration. *Traffic Safety Facts 2009: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and General Estimation System*. DOT HS 811 402, Washington D.C., 2011.

National Highway Traffic Safety Administration. *2010 Motor Vehicle Crashes: Overview*. DOT HS 811 552, Washington D.C., 2012.

National Highway Traffic Safety Administration. *Early Estimate of Motor Vehicle Traffic Fatalities for The First Half (January–June) of 2012*. DOT HS 811 680, Washington D.C., 2012.

National Safe Kids Campaign. *Stop Sign Violations Put Child Pedestrians at Risk.* A National Survey of Motorist Behavior at Stop Signs in School Zones and Residential Areas, 2003.

Neuman, T.R., R. Pfefer, K.L. Slack, K.K. Hardy, D.W. Harwood, I.B. Potts, D.J. Torbic, and E.R.K. Rabbani. *A Guide for Addressing Unsignalized Intersection Collisions*. NCHRP Report 500, Transportation Research Board, Washington, D.C., 2003.

Preusser, D.F., A.F. Williams, S.A. Ferguson, R.G. Ulmer, and H.B. Weinstein. “Fatal Crash Risk for Older Drivers at Intersections.” *Accident Analysis and Prevention* 30 (2), (1998): 151-159.

Retting, R.A., H.B. Weinstein, and M.G. Solomon. “Analysis of Motor Vehicle Crashes at Stop Signs in Four U.S. Cities.” *Journal of Safety Research* 34, (2003): 485-489.

Van Houten, R. and R.A. Retting. “Increasing Motorist Compliance and Caution at Stop-Signs.” *Journal of Applied Behavior Analysis* 34, (2001): 185-193.

Wang, J.S. *The Effectiveness of Daytime Running Lights for Passenger Vehicles*. NHTSA Technical Report. USDOT, Washington. D.C., 2008.

Dr. Mintesnot Woldeamanuel *is an assistant professor at the department of Urban Studies and Planning at California State University Northridge. Prior to joining Cal State Northridge, he worked as an assistant professor at the Planning and Community Development Program of Saint Cloud State University and as a research fellow at the Transportation Research Institute of German Aerospace Center (DLR). He has a Ph.D. in urban and environmental planning/engineering. He has conducted extensive research and written on the relationship between travel behavior, the environment, and urban form, focusing on transportation mode choice and spatial analysis.*

Book Review

Martland, Carl D. Toward More Sustainable Infrastructure: Project Evaluation for Planners and Engineers. *Hoboken, NJ: John Wiley & Sons, Inc., 2012. ISBN 9780470448762.*

Toward More Sustainable Infrastructure

by Michael E. Smith

Carl Martland has introduced an excellent textbook designed to provide undergraduate students in civil engineering with a high-level understanding of where projects come from, how they should be evaluated, what we should expect to consider when selecting a project or project variant, how projects fit within an overall program, and what the characteristics are of projects that are more likely to be chosen and are more likely to be successful. Each of the 16 chapters in the textbook begins with a superb list of concepts to be covered followed by detailed discussion of the topics, including exemplary illustrative problems. Each chapter is followed by a well thought out summary of topics discussed in the chapter. Following each chapter summary is a set of exercises for the student; these include essays and problems. Students working their way through the essays and problems will find the summaries, topics lists, and example problems to be exactly the kind of resources that they require to quickly locate the specific material needed to assist them in creating first-class essays and answers. The student will find those items very useful when studying for exams.

The example problems contained in each chapter are well explained and illustrate important concepts. In Chapter 2 we see a well-constructed example illustrating how investment in a canal in the nineteenth century could be justified based on the fact that moving goods via water would be considerably less expensive than moving goods using horses and wagons. But the investment in creating the canal is considerable and the reach of the canal is not anywhere near as complete as the reach of roadways and paths that can be traversed by the horses. The result is a classic tradeoff between large investments, lower operating costs, and a less complete level of service that we see today in an untold number of instances.

In a number of cases, the text could be improved by presenting more balanced economic information that would help the nascent engineer avoid a few pitfalls. First, there are a number of engineers who view the world in terms of linear cost functions. That is, they see the provider of a good or service that faces total costs consisting of a fixed start-up cost and a linear cost per unit. The reality, though, is that all providers of goods and services in a free market will continue to create their product until the marginal costs of that product are an increasing function of the number of units. Were they not to do this, then we would have horizontal supply curves and would be blessed with infinite quantities of things. As we are not so blessed, we know that supply curves are not horizontal. I think that a useful addition would be to provide an understanding of cost curves that are non-linear, having a rising marginal cost of production.

The book also contains some references to economic issues that are currently in dispute. These include the multiplier effect (many economists dispute its existence) and the belief that monopolists will always charge excessive prices that can only be held in check by appropriate government regulation. Studies of electric utilities have found that regulation generally fails to hold down prices in that manner. Also, although the text covers benefit-cost analysis very well, there could be usefulness in introducing the concept of Pareto optimality. With that approach, a solution can be searched for that helps everybody instead of the majority. One of the useful things about private-sector solutions is that they tend to approach Pareto optimality while government-based solutions work toward improvements for the majority.

Problem 4.5 discusses the tragedy of the commons. This concept is a very useful one for engineers to understand. One of the best examples of this tragedy that I know of is the urban freeway. Examining how to solve the problem of the over-crowded freeway without burdening government would make for an excellent additional problem.

The lesson of Section 5.5 is excellently stated and is quite true. Since decisions and investments that are undertaken by a government body involve all the complications of the political process in addition to financial and economic evaluation, decisionmaking for such projects can become quite belabored. While private companies still need to evaluate the financial aspects of a project, they do not concern themselves with some of the more arcane political aspects except to the extent required by law. At this point in the text, it could be useful to suggest that moving as many decisions as possible into the private sector, while retaining appropriate regulatory oversight, could be superior to having governments invest in such projects themselves.

Chapter 6 provides a marvelous history of constructing the Panama Canal. At first, the Canal was largely not a commercial success. Instead, the military found it very useful. I think that a great opportunity exists here to compare and contrast how the government pursued the investment and how a private company would handle it.

The essays and problems themselves cover all of the topics in their respective chapters quite well. And some of the characters in the problems have rather clever names. In Chapter 9 we are introduced to the developer Canwy Bildem, who suggests that construction on a new skyscraper should proceed immediately. Perhaps Mr. Bildem is represented by Dewey, Cheatham & Howe! Also, one of the problems in Chapter 8 introduces us to Bonnie and Clyde, who consider bank robbery as a possible way to finance their dream of pursuing a particularly risky investment in South America.

I was particularly impressed with the example problem on page 233 where Martland introduces the concept of incremental investment. Engineers will frequently look at a problem based on totals and averages, ignoring the fact that an incremental investment of I will provide a marginal return of R. Thus, the rate of return for that increment is R/I, which may be much too low, even if the rate of return for the total investment is quite acceptable. That is how we can tell that we should, perhaps, stop with a smaller project.

The chapter on Public Private Partnerships is very impressive. I note that the analyses were all based on well-defined cash flows that resulted in each case. That is, the analysis proceeded as though the government were simply another corporation seeking to maximize its return. The case of the Kansas City flyover project is an excellent example of this. The text discusses the private benefits of reducing delays to the freight trains and the public benefits of reducing delays to travelers on the public roadways. Martland does not discuss how this problem would look were the highways to be privately owned. But if we consider that possibility for a moment, we could see that the entire project, if it were analyzed in the same manner as presented in the text, could be considered entirely from a private-sector view.

At the end of Chapter 12, the student is presented with a case study on Positive Train Control (PTC). The railroads have considered many incarnations of PTC over the years and have concluded that the technology is one in which they do not wish to invest. The government has suggested that the safety benefits of PTC are sufficient that the public should require the railroads to make that investment. The PTC case helps to illustrate quite graphically the public-private tradeoffs and considerable array of approaches that project advocates can use to ensure implementation of a particular investment plan.

The chapter on risk does a thorough job of enumerating the risks that are faced in the pursuit of any project. I might add that project promoters will very likely underestimate the risks, underestimate the costs, and overestimate the benefits. That is the reason why hurdle rates can significantly exceed minimum acceptable rate of return. It is this risk of poor estimation that may be the largest of all risks.

One way to address the risks of poor estimation of cash flows is to estimate them better. And the discussion of using simulation to estimate cash flows is on target. I was particularly impressed with the detailed discussion of how spreadsheets can be used for that process. Computer programming languages, such as FORTRAN, have been used for such purposes for decades. But that approach is difficult to understand and takes a very long time to get right. Spreadsheets are intuitive and easy to understand, making them a good tool for many kinds of analysis. And Martland shows us how the ability to generate random numbers in a spreadsheet adds to the capability of the tool.

Overall, I find the textbook to be an excellent one for instructing undergraduates who want to know how projects should be conceived, analyzed, selected, and managed. Most engineering textbooks look at only the very detailed concepts of project design. As Martland's text points out, though, there is a real need for engineers to go beyond such a narrow perspective. Indeed, I find that this textbook does a great job of answering many of the questions posed by a textbook in engineering economics that formed the basis of a course I took many years ago. In that text, the authors suggested that while engineers focus on project design, it is at times far more important to be able to answer such basic questions as, "Why do the project now?" or "Why do the project in this way?" or even "Why do this project at all?" Much of the time engineers are focused on doing things just because it is technically feasible. But it is just as important that any proposed project be politically and economically feasible as well. Martland's text provides the nascent engineer a marvelous grounding in how to make sure that all the bases are covered.

Michael E. Smith *is an economist for the United States Surface Transportation Board, providing insights into railroad operations based on the Board's annual waybill sample and serving on teams that investigate complaints about railroad rates and services. He received a B.S.(1974) and an M.S.(1975) in civil engineering from Ohio State University and an M.B.A. (1981) from the University of Maryland. He began work in 1975 with the Federal Highway Administration as a transportation planner, then served at the Federal Railroad Administration, assisting with the analysis of the benefits and costs of subsidizing small rail branch lines. Next, at the Association of American Railroads, he led the development of models to understand the economics of introducing technology improvements in the railroad industry. Burlington Northern then hired him to perform a business case analysis for their Advanced Railroad Electronics System, now more widely recognized in the railroad industry as Positive Train Control. After that, he served as manager at Reebie Associates. He is author of more than a dozen published papers in the field of transportation and is pursuing a Ph.D. in economics at George Mason University. He is a registered professional engineer in the District of Columbia.*

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

2011 Alan R. Bender
2010 B. Starr McMullen
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

2012 Genevieve Giuliano
2011 Martin Wachs
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

2012 Kofi Obeng
2011 Anthony Pagano
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

2012 Rafael Costa & C. Parr Rosson, III, *The Impacts of the Panama Canal Expansion on World Cotton Trade*

2011 Alexander Bigazzi and Miguel Figliozzi, *A Model and Case Study of the Impacts of Stochastic Capacity on Freeway Traffic Flow and Benefits Costs*

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
Email: 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.

The Transportation Research Forum gratefully acknowledges the contributions
of sponsoring and sustaining members.

2011-2012 Supporting Members

Michael Babcock, Kansas State University
Marcus Bowman, 3G Mobility, LLC
Michael Bronzini, George Mason University
Paul Gessner, Gessner Transportation Consulting LLC
Andrew Jennings
Eric Jessup, University of Freiburg
Elizabeth Ogard, Prime Focus LLC
Jack Ventura
John Wells, U.S. Department of Transportation

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 individual files at 250 dpi or higher resolution (.jpg or .tif format).
    - Images taken from the web are 72 dpi and are not acceptable for print. Secure high resolution
    - image when possible. Obtain permission for use of images taken from web.
    - Do not embed graphics in Word – submit as individual files (.jpg or .tif 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
1025 Chiswell Lane
Silver Spring, MD 20901
Phone: (301) 593-1872
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