Volume 8 Number 1, 2005
ISSN 1094-8848

JOURNAL OF TRANSPORTATION AND STATISTICS

U.S. Department of Transportation
Research and Innovative Technology Administration
Bureau of Transportation Statistics

JOURNAL OF TRANSPORTATION AND STATISTICS

PEG YOUNG Editor-in-Chief
DAVID CHIEN Associate Editor
CAESAR SINGH Associate Editor
JEFFERY MEMMOTT Associate Editor
MARSHA FENN Managing Editor
JENNIFER BRADY Data Review Editor
VINCENT YAO Book Review Editor
DORINDA EDMONDSON Desktop Publisher
ALPHA GLASS-WINGFIELD Editorial Assistant
LORISA SMITH Desktop Publisher

A PEER-REVIEWED JOURNAL

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 8 Number 1, 2005
ISSN 1094-8848

U.S. Department of Transportation
Research and Innovative Technology Administration
Bureau of Transportation Statistics

The *Journal of Transportation and Statistics* releases three numbered issues a year and is published by the

Bureau of Transportation Statistics
Research and Innovative Technology Administration
U.S. Department of Transportation
400 7th Street SW, Room 4117
Washington, DC 20590
USA
journal@bts.gov

Subscription information

To receive a complimentary subscription:

mail Product Orders
Bureau of Transportation Statistics
Research and Innovative Technology Administration
U.S. Department of Transportation
400 7th Street SW, Room 4117
Washington, DC 20590
USA

phone 202.366.DATA
internet www.bts.dot.gov

Information Service

email answers@bts.gov
phone 800.853.1351

Cover and text design Susan JZ Hoffmeyer
Cover photo Gregg Stansbery

The Secretary of Transportation has determined that the publication of this periodical is necessary in the transaction of the public business required by law of this Department.

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 8 Number 1
2005

Contents

Letter from the Deputy Administrator

Dear JTS Readers,

In February 2005, the Norman Y. Mineta Research and Special Programs Improvement Act created a new U.S. Department of Transportation modal administration, known as the Research and Innovative Technology Administration (RITA). The mission statement for this new organization is straightforward: identify and facilitate solutions to the challenges and opportunities facing America's transportation system.

RITA's staff came from the former Research and Special Programs Administration's Office of Innovation, Research and Education, the Secretary's Office of Intermodalism, the Bureau of Transportation Statistics, the Transportation Safety Institute in Oklahoma City, and the Volpe National Transportation Systems Center in Cambridge, Massachusetts. More than 750 DOT employees make up this unique organization.

RITA will enable DOT to more effectively coordinate and manage the Department's research portfolio and expedite implementation of crosscutting innovative technologies. Secretary Mineta's vision of RITA is part university research lab and part Silicon Valley entrepreneurial company. He wants this administration to foster the exchange of ideas and information in a high-priority incubator committed to research and move these innovative ideas from the laboratory into the field. We look forward to *Journal of Transportation and Statistics* (JTS) readers coming along and even participating in our journey of discovery.

RITA is dedicated to the advancement of DOT priorities for innovation and research in transportation technologies and concepts. These innovations will improve our mobility, promote economic growth and safety, and ultimately deliver a more integrated transportation system.

RITA is not intended to displace the R&D activities of the various DOT operating administrations, neither will it intervene with their associations with particular transportation entities or modal communities. While the offices assigned to RITA continue their current, vital work, they will grow. This growth will afford DOT the opportunity to realize greater collaboration, information sharing, coordination, support, and advocacy for its widespread research efforts.

As the first Deputy Administrator of RITA, I thank all the JTS readers for their continued interest in this publication. Please feel free to share this publication with others. We look forward to expanding our readership through you, our valued readers. I hope you find this publication and our new administration a valued resource.

Please do not hesitate to contact me or the JTS editorial staff with your questions and comments about either this publication or other transportation interests.

Sincerely,

Eric C. Peterson

ERIC C. PETERSON
Deputy Administrator
Research and Innovative Technology Administration

The Dynamics of Aircraft Degradation and Mechanical Failure

LEONARD MACLEAN[1, *]
ALEX RICHMAN[2]
STIG LARSSON[1]
VINCENT RICHMAN[3]

[1] School of Business Administration
Dalhousie University
Halifax, Canada B3H 3J5

[2] AlgoPlus Consulting Ltd.
Halifax, Canada B3H 1H6

[3] Sonoma State University
Rohnert Park, CA 94928-3609

ABSTRACT

This paper looks at the predictability of system failures of aging aircraft. We present a stochastic, dynamic model for the trajectory of the operating condition with use. With failure defined as the operating condition below a critical level, the dynamics of the number of failures with accumulated use is developed. The important factors in the prediction of mechanical failures are the number of previous repairs and the time since last repair. Those factors are related to repair procedures, with the time of repair and the extent of repair (fraction of good-as-new) being variables under the control of the operator. The methodology is then applied to data on non-accident mechanical failures affecting safety that result in unscheduled landings.

INTRODUCTION

An aircraft is a complex machine composed of many interrelated parts, components, and systems. Electrical and mechanical systems are designed with an expected life length, where length refers to time units (hours) of use. As the aircraft and systems age and their use accumulates, they gradually

Email addresses:
*Corresponding author—L.C.MacLean@dal.ca
A. Richman—arichman@algoplusaviation.com
S. Larsson—S.O.Larsson@dal.ca
V. Richman—Vincent.richman@sonoma.edu

KEYWORDS: Aircraft failures, aircraft degradation and repair, airline schedule reliability.

degenerate until they are no longer able to perform the functions for which they were designed; that is, the system is in a failed state.

A nonfunctional part, component, or system can be upgraded through replacement or repair, in which case the condition of the aircraft is restored to some degree. Maintenance can be based on the condition; that is, items are repaired when they fail. However, failure during operation can have serious consequences, so detection of items with a high probability of failure through periodic inspection becomes a major component of maintenance.

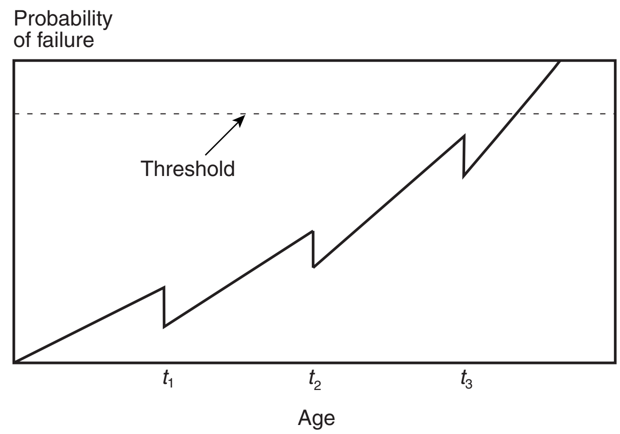
The failure rate (the probability of failure at a point in time) for a degenerating system increases with use and age. Figure 1 depicts alternative patterns of failure rates for an aircraft that undergoes periodic maintenance (a similar figure appears in Lincoln 2000). In case A, the aircraft has an increasing failure rate with age and reaches an acceptability threshold, at which point the aircraft would need to be replaced. The failure rate declines with periodic maintenance, but the improvement through maintenance diminishes over time. The threshold is not reached in case B, likely because of increased effort and cost put into maintaining the aircraft.

The cost of maintenance required to keep aircraft airworthy (below the threshold) is a major concern of operators. Although replacement time was set by manufacturers at 20 years for many aircraft models, this life length was extended by operators. An assumption has been made that aircraft operating condition can be kept at an acceptable level beyond the intended life through maintenance, but costs are high.

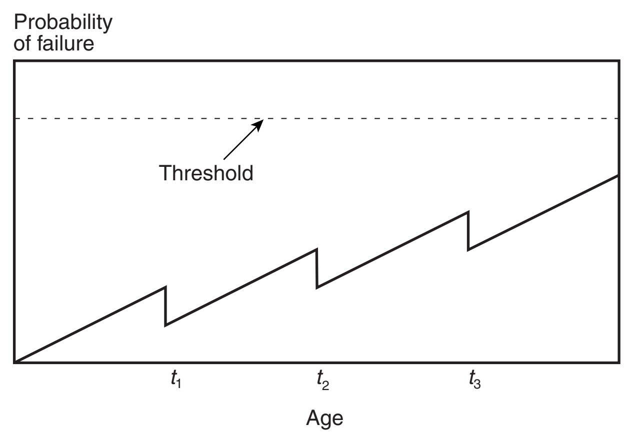
In 1997, 46% of U.S. commercial aircraft were over 17 years of age and 28% were over 20 years. In 2001, 31% of the U.S. commercial fleet were over 15 years of age, and those aircraft accounted for 66% of the total cost of maintenance per block hour.[1] Although aging (the degeneration in operating condition with accumulated use) inevitably occurs, it is modified by a number of factors: quantity and quality of repair work; intensity of use; deferral of the schedule for planned maintenance;

FIGURE 1 Probability of Failure of a Single Aircraft

Case A: Increasing Failure Rate



Case B: Constant Failure Rate



Sources: AlgoPlus, 2004, available at http://www.cislpiloti.org/Technical/GAIN/WGB; and AvSoft, 2004, available at http://www.avsoft.co.uk/.

and the environment (Alfred et al. 1997). It should be noted, however, unscheduled maintenance accounts for up to 60% of the overall maintenance workload (Phelan 2003).

In addition to the possibility of maintaining an airworthy operating condition, other reasons exist for not retiring an aircraft from the fleet: the high cost of new aircraft; the increase in demand requiring an expanded fleet; and an earlier shortage of production capacity for new planes (Friend 1992). These and other factors result in a large number of aircraft in use beyond their planned retirement. A claim could be made that commercial aircraft are being strained to perform well beyond their intended operating life. Of course, if this is true we would expect that either the rate of aircraft failure

[1] A *block hour* refers to flying time in hours, including takeoff and landing.

would show a corresponding increase or the operating hours per aircraft would decrease because planes would be out of service for repairs more frequently.

In the United States, Service Difficulty Reports (SDRs) contain records of the safety problems an aircraft experiences during operation. This database, maintained by the Federal Aviation Administration, is considered a potential source of important information on aircraft failures (Sampath 2000). A study comparing failure rates by carrier identified significant factors that explain the differences in the rate of SDRs across carriers (Kanafani et al. 1993). Because accumulated aircraft use (age) was not included in that study, degradation with age and differences over time in the safety of individual aircraft were not considered.

A THEORY OF DEGRADATION AND REPAIR

With age and accumulated use, the many interrelated parts and components in an aircraft can be assumed to degrade. The operating condition or airworthiness of the aircraft is based on the status of individual parts, components, and systems, with the items that are most degenerated being the main determinants. A certain level of degeneration implies failure; that is, the item is no longer operational. As well, failure of certain components or combinations of components may render the aircraft not airworthy, which means the aircraft is in a failed state. To address the failure of operating systems, airline management undertakes a program of maintenance, with scheduled preventive maintenance and unscheduled repair/replacement of failed parts, components, and systems. This section presents a conceptual model for the degradation and repair of aircraft. The model provides a foundation for hypotheses about operations that can be tested with field data.

Degradation

To characterize the degradation process, consider that the operating condition of an aircraft is captured by an unobserved health status index. The value of the index is derived from the condition of the various parts, components, and systems in the aircraft. Let t be the age of an aircraft, defined by the accumulated hours of use, and let

$Y(t)$ = the health status of an aircraft at age t.

The status is a dynamic stochastic process, with the change in status at any age being a random variable. Assume that the average condition declines with age, but at any point variation in the status, based on environmental factors and operating characteristics, can occur. The dynamics of degradation at a point in time can be represented by a stochastic differential equation as

$$dY(t) = \mu_t dt + \sigma_t dZ_t \tag{1}$$

where

$\mu_t < 0$ is the degradation rate,

$\sigma_t > 0$ is a scaling factor, and

dZ_t is an independent random process.

For example, if the random process is white noise, the stochastic differential equation defines a Wiener process, and the distribution of the health status at a given age is Gaussian (Aven and Jensen 1998). So, with starting state y_0 and constant parameters μ and σ, the distribution of status after t time periods is Normal,

$$Y(t) \propto N\left(y_0 + \mu t, t\sigma^2\right) \tag{2}$$

Mechanical Failure

In this degradation framework, at any age (hours of use) the possibility exists that the status of a item during operation will drop below the critical level for functionality and the component reaches a failure state. Degradation and failure of components lower the value of the health status index Y. In particular, failure of parts and components included on a minimum equipment list (MEL) indicates the aircraft remains airworthy. Beyond the MEL, moderate mechanical failures that occur during aircraft operation would render the aircraft not airworthy. Assume that the critical health status level y^* defines airworthiness. Then an aircraft failure occurs when $Y(t) < y^*$.

Based on the stochastic model, the many parts, components, and systems have a probability of failure during operation and, therefore, the aircraft has a probability of failure. For an airworthy aircraft, the important variable is the time to failure. Let T be the length of life (hours of use before failure) of an aircraft, with the probability distribution $F(t) = Pr(T \le t)$, and the corresponding density $f(t)$. Then

$$\lambda(t) = \frac{f(t)}{1 - F(t)}$$

is the failure rate at time t (Aven and Jensen 1998). The failure time distribution is determined by the failure rate because

$$F(t) = 1 - \exp\left\{\int_0^t \lambda(s)ds\right\}.$$

In the example, where the state dynamics are defined by a Wiener process, the failure time is inverse Gaussian with density

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2 t^3}} \exp\left(-\frac{((y_0 - y^*) - \mu t)^2}{2\sigma^2 t}\right). \quad (3)$$

Repair

Failure during operation may precipitate unscheduled maintenance, particularly when items beyond the MEL fail and consequently the aircraft is not airworthy. The repair/replacement of failed items is called on-condition repair. On-condition repair brings the system back to the operating status expected of the system given its age, that is, the same status as just prior to failure. With these *minimal repairs* (Block et al. 1985), the aircraft failure rate is unchanged since other parts, components, and systems are still in the degraded state attained just before repair. Typically, moderate mechanical failures result in such minimal repair.

In addition to unscheduled maintenance, the whole system is subject to time-based or *block repair,* where items are inspected and replaced/refurbished before failure. This scheduled preventive maintenance improves the operating condition to a status greater than expected for its age and correspondingly reduces the system failure rate (Brown and Proschan 1983). To incorporate repair into the degradation model, the age variable is partitioned into intervals based on the block repair times. Assume that the first scheduled block repair is at age (hours of use) τ, and subsequent block repairs are at regular intervals of δ hours of use, where $\delta \le \tau$. Then age t can be written as

$$t = I\tau + k\delta + r, \quad (4)$$

where $I = 1$ if $t \ge \tau$, $I = 0$ if $t < \tau$;

$k = \left]\frac{t-\tau}{\delta}\right[$ if $t \ge \tau$, $k = 0$ if $t < \tau$; and

$r = t - I\tau - k\delta$.

The notation $]x[$ defines the greatest integer less than x. Equation (4) gives the age in terms of $(I + k)$ = the number of repairs, and r = the use since the last block repair. The intervention with a block repair improves the health status of an aircraft above the level expected for its age. Let the improvement level from a block repair at age t be up to the line

$$y(t) = \alpha + \beta t, \quad (5)$$

where $\alpha > y^*$ and $\beta \le 0$.

The repair line is theoretical and the important parameter is β, which describes the repair policy to return the aircraft to a fraction of good-as-new at scheduled times. If $\beta = 0$, then repair always brings the plane to the same health status regardless of age.

To simplify the presentation, repair policies that are equivalent in terms of the total repair effort will be considered. Let

L = the expected length of the operating life of an aircraft.

Assume that all feasible repair policies have the same total repair over the expected life of the aircraft. That is,

$$\int_0^L (\alpha + \beta t)dt = \alpha^* L,$$

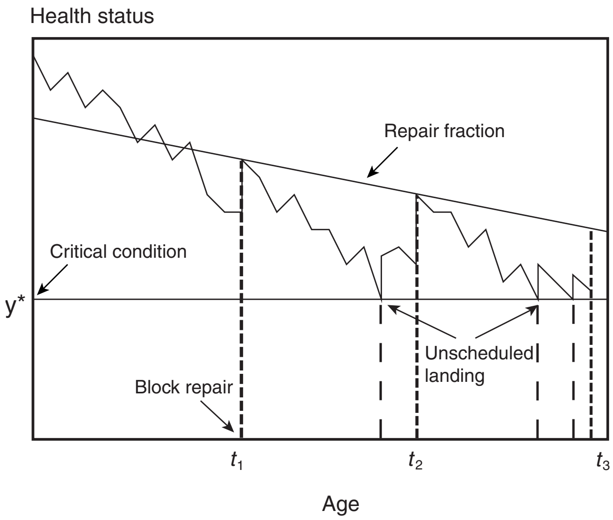
for some constant α^*. With this condition, the repair policy is determined by β, which also determines the distribution of repair over the aircraft's lifetime.

The partition of age at block repairs generates renewal cycles for the degradation process, with the first cycle starting at the initial status y_0, and subsequent cycles beginning at the status defined by the repair line:

$$y_j(\tau, \delta, \beta) = \alpha + \beta t_j, \quad (6)$$

where $t_j = \tau + (j-1)\delta, j = 1, ..., k$. The repair policy is defined by: τ — the hours of use until the first block repair; δ — the hours of use between subsequent block repairs; and β — the repair fraction. Using the definitions of t_j and y_j, the increase in the health status at each block repair from β can be calculated. The policy determines the starting state

FIGURE 2 Aircraft Degradation and Repair



Sources: AlgoPlus, 2004, available at http://www.cislpiloti.org/Technical/GAIN/WGB; and AvSoft, 2004, available at http://www.avsoft.co.uk/.

and length of renewal phases or cycles for the degeneration process. Figure 2 illustrates the cycles of degradation and repair for an aircraft.

NUMBER OF FAILURES

In each renewal phase of the degradation model, there is a chance that the aircraft fails; that is, its status drops below the critical level y^*. Let T_j = time to critical condition y^* in cycle j starting from y_{j-1}, $j = 1,2,\ldots$. The failure time distribution for T_j is written as $F_j(s|\tau, \delta, \beta)$, with density $f_j(s|\tau, \delta, \beta)$, where the repair policy (τ, δ, β) determines the starting status. Given the failure time distribution, the failure rate in the jth cycle is

$$\lambda_j(s|\tau, \delta, \beta) = \frac{f_j(s|\tau, \delta, \beta)}{1 - F_j(s|\tau, \delta, \beta)}. \tag{7}$$

Consider

$N(t)$ = the number of aircraft failures to age t for repair policy (τ, δ, β). (8)

Because

$$\int_0^t \lambda(s|\tau, \delta, \beta) ds = -ln(1 - F(t))$$

the expected number of failures is

$$E(N(t)) = I(-ln(1 - F_1(t))) - \sum_{j=1}^{k} ln(1 - F_{j+1}(\delta)) - ln(1 - F_{k+2}(r)). \tag{9}$$

In the failure rate for each renewal phase, the probability distribution for the time to failure has the same form, but the starting state in each phase declines if $\beta < 0$. With $t_j = \tau + (j-1)\delta$, and the starting state in phase $j + 1$ as

$$y_j = \alpha + \beta t_j = \alpha^* + \beta\left(\tau + (j-1)\delta - \frac{L}{2}\right),$$

define

$$\psi(x, y_j) = -ln(1 - F_{j+1}(x)). \tag{10}$$

Thus, $\psi(x, y_j)$ is the expected number of failures between times 0 and x in phase $j + 1$, with failure time distribution F_{j+1} and starting state y_j. Then

$$E(N(t)) = I \cdot \psi(\tau, y_0) + \sum_{j=1}^{k} \psi(\delta, y_j) + \psi(r, y_{k+1}). \tag{11}$$

The degradation process and repair policy are determined by parameter values, and those policies determine the properties of the expected number of failures over time. Let ψ'_x and ψ'_y denote first derivatives of ψ with respect to x and y, respectively. Thus, ψ'_x is the change in expected failures with use (degradation) within a phase, and ψ'_y is the change with respect to the phase starting state, determined by the block repair policy. The following general results establish the expectations for mechanical failures when the degeneration model applies.

Proposition 1 (degeneration): If the health status of an aircraft degenerates with use, then between block repairs, the failure rate with use increases, as does the expected number of failures in a fixed-width use interval.

In the dynamic model, degeneration follows from $\mu < 0$. With degradation, $\psi'_x = \lambda > 0$, and $\psi''_x > 0$, which implies an increasing failure rate between repairs.

Proposition 2 (imperfect repair): If the block repair is imperfect, then the failure rate with use since the last repair is nondecreasing with the number of previous block repairs, and the expected number of failures in a fixed (use since last repair) interval is nondecreasing with the number of repairs. If the repair fraction decreases over time, then the expected number of failures is increasing.

In the model, $\psi'_y > 0$. If $\alpha < y_0, \beta = 0$, then the expected number of failures in a fixed interval is constant after the initial block repair. If $\beta < 0$, then the starting state y decreases, with increasing failure rates in successive phases between block repairs.

Proposition 3 (repair interval): If the imperfect repair fraction is decreasing over time, then the expected number of failures in a fixed-use interval increases/decreases as the block repair interval increases/decreases.

Block repair increases the health status above that expected for accumulated use, so more block repairs (shorter times between block repairs) raise the expected value of y and decrease the failure rate and number of failures.

FAILURE MODEL

The link between the latent state model for degradation/repair and the model for the number of failures shows how the operating practices of airlines can manifest themselves in mechanical failures, safety problems, and unscheduled maintenance. Historical data on failures and maintenance will have that complex relationship embedded. The information on failures and block repairs is available, but the degradation rate and extent of repair (fraction of good-as-new) are unknown. However, from Proposition 1, the time since the last block repair reflects degradation, and from Proposition 2, the extent of the repair is directly related to the number of block repairs. The transformation of equation (11) for the expected number of failures to an expression in terms of the number of block repairs and the time since the last block repair is achieved by a series approximation to the function for $E(N(t))$.

From the model, the average level of repair is α^*. Consider the first order approximation to $\psi(\delta, y)$ around (δ, α^*):

$$\psi(\delta, y_j) \approx C_0(\tau, \delta, \alpha^*) + C_1(\tau, \delta, \alpha^*) \times j. \qquad (12)$$

In the last (incomplete) phase, a second order approximation to the number of failures around $(0, \alpha^*)$ is reasonable, assuming the failure rate is increasing monotonically with use. Then

$$\psi(r, y_{k+1}) \approx D_0(\tau, \delta, \alpha^*) + D_1(\tau, \delta, \alpha^*)k + D_2(\tau, \delta, \alpha^*)r + D_3(\tau, \delta, \alpha^*)r^2. \qquad (13)$$

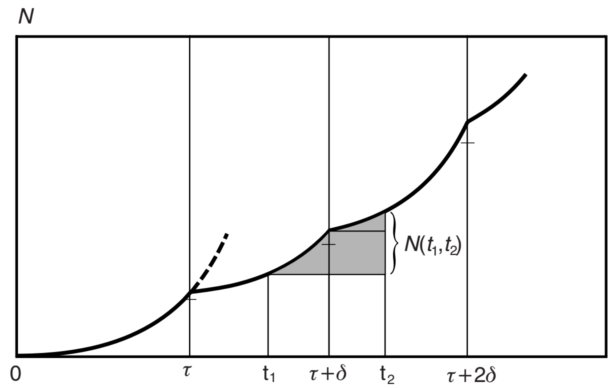
The coefficients in the approximating functions are defined by derivatives of ψ, evaluated at (τ, δ, α^*). Substituting the approximations in equations (12) and (13) into equation (11), the expected number of failures has the form

$$E(N(t)) \approx B_0 + B_1 I + B_2 k + B_3 k^2 + B_4 r + B_5 r^2. \qquad (14)$$

(Note that $\sum_{j=1}^{k} j = \frac{k + k^2}{2}$.)

A representation of the number of failures is shown in figure 3.

FIGURE 3 Number of Failures in an Observation Window



Sources: AlgoPlus, 2004, available at http://www.cislpiloti.org/Technical/GAIN/WGB; and AvSoft, 2004, available at http://www.avsoft.co.uk/.

Thus, B_1 is the expected number of failures in the first phase. $\{B_2, B_3\}$ capture the expected number of failures in subsequent phases, and $\{B_4, B_5\}$ capture the expected number in the last (incomplete) phase. The coefficients in the expected number function that relate to the propositions are B_3 and B_5. If the block repair is imperfect, then $B_3 > 0$. Figure 3 shows this effect with the failure function starting above the origin at block repair times. An accelerated failure rate between block repairs implies $B_5 > 0$, which is shown with a steeper slope in successive phases between repairs.

The approximating equation for the expected number of failures is in a very suitable form for analysis. Consider an observation window (interval) (t_1, t_2), where $t_2 - t_1 < \delta$. With $t_1 = I_1\tau + k_1\delta + r_1$, and $t_2 = I_2\tau + k_2\delta + r_2$, the number of failures in the interval (t_1, t_2) is approximately

$$\eta = E(N(t_1, t_2)) = E(N(t_2)) - E(N(t_1)),$$

so that

$$\eta = B_1(I_2 - I_1) + B_2(k_2 - k_1) + B_3(k_2^2 - k_1^2) + B_4(r_2 - r_1) + B_5(r_2^2 - r_1^2). \quad (15)$$

This change model relates the number of failures in an observation window to the degeneration and the block repairs in the window. In equation (15), $(I_2 - I_1) = 1$ if the first repair is in the interval, and zero otherwise; $(k_2 - k_1) = 1$ if a later repair occurs in the window, and zero otherwise.

MODEL TESTING

We used data from AlgoPlus (2004) to test the failure model on operating failures and AvSoft (2004) on aircraft use. For the purposes of this study, an operating failure is defined as an unscheduled landing due to mechanical problems affecting safety. Thus, an unscheduled landing is a record of an operating condition at or below a critical or intervention level. In figure 2, the unscheduled landings (failures) occur when the health status drops to the critical level, where airworthiness fails. It is also possible that components fail and the event does not lead to an unscheduled emergency landing. As mentioned earlier, a minimum equipment list details which components may fail without the need for an unscheduled landing. In terms of the degradation/repair model, the critical condition line is below the condition for failures on the MEL, so that reaching the critical line implies unsafe operation and a need to interrupt the flight of an aircraft.

Data

The record of unscheduled landings over time provides an information base for analyzing the degeneration in the operating condition of an aircraft. The AlgoPlus data contain detailed records on all unscheduled landings as reported in the Service Difficulty Reports for all commercial aircraft in the United States. The AvSoft data maintain records on departures and flying hours for all commercial aircraft in North America. Both datasets have the serial number, chronological age, model, and carrier/operator for each aircraft.

An observation window from 1990 to 1995 inclusive was chosen, and all aircraft operating during that time for three operators and two models were selected for this study. For each aircraft, the following information was recorded: 1) model; 2) operator; 3) age on December 30, 1995; 4) use (block hours, cycles) by month from January 1990 to December 1995; 5) dates out of service for at least one month between 1990 and 1995; and 6) number of unscheduled landings between 1990 and 1995. We interpreted the out-of-service period in the observation window as a time when scheduled repair was undertaken. The identification of these periods is within a record of otherwise continuous use. Outside the observation window, the block repair (preventive maintenance) cycle was set at 10 years for the first block repair and 8 years for subsequent block repairs. This is based on the recommendations for D-check cycles.[2] Of course, in practice the time of block repairs would be variable across aircraft and using a fixed value (outside the window) could reduce the power of the fitted models.

Table 1 presents a brief description of the aircraft in the dataset. For the aircraft in the study group, table 1 shows substantial differences across models and operators in the age of aircraft as of December 1995 and the number of unscheduled landings between January 1990 and December 1995.

TABLE 1 Description of the Study Sample

Operator	Model	Number	Average age (yrs)	Average unscheduled landings
O_1	M_1	230	10.43	3.15
	M_2	34	5.17	1.82
O_2	M_1	221	7.74	1.35
	M_2	0	—	—
O_3	M_1	73	10.83	0.96
	M_2	83	6.93	0.77

The definition of age in the degradation of aircraft refers to hours of use rather than chronological age. However, an aircraft operator might make little distinction between airworthy aircraft of varying ages when making decisions on use. To consider this point, we looked at the relationship between

[2] A *D-check* refers to the major maintenance and overhaul programs in which the aircraft is completely stripped down and inspected, with many parts and components replaced or refurbished.

flying hours per month and chronological age in the data for the period 1990 to 1995. The correlation in the data between monthly flying hours and age is $r = 0.07$. The intensity of use appears almost constant across age, indicating that aircraft are not being used less as they age. With constant use per unit time, the accumulated hours of use are almost a scalar multiple of chronological age. So, calendar time was used in the model for predicting the number of failures; that is, the time between block repairs and the time since the last repair will be measured in calendar time rather than accumulated hours of use.

Regression Model

The formulation of a change model for the number of failures creates a framework suitable for observation and statistical analysis. Based on the model in equation (15), consider the regression model

$$N(t_1, t_2) = \beta_1^q X_1 + \beta_2^q X_2 + \beta_3^q X_3 + \beta_4^q X_4 + \beta_5^q X_5 + \varepsilon \,, \tag{16}$$

where

$N(t_1,t_2)$ = the number of failures between ages t_1 and t_2,

X_1 = the indicator for the first τ-repair in the interval,

X_2 = the indicator for the kth δ-repair in the interval, $k \geq 1$,

X_3 = the difference between the squared number of repairs, $k_2^2 - k_1^2$,

X_4 = the difference in residual times, $r_2 - r_1$,

X_5 = the difference in squared residual times, $r_2^2 - r_1^2$, and

ε = the random error.

In the regression model, assume that the unscheduled landings and item failures from degradation are directly related to the number of block repairs and the time since the last block repair. There are also other factors such as repair skill level, maintenance philosophy, and operational environment involved in unscheduled landings (Phelan 2003). We will assume that these other factors are associated with the operator. As well, the aircraft model is a factor in failure rates. So, the coefficients in the regression model depend on the aircraft model and the aircraft operator. This is reflected in the regression model with a superscript q on the coefficients.

The coefficients in the regression model are counterparts of the coefficients in the failure model, and appropriate tests characterize the role of degradation and repair on failures for a particular model and operator combination. Table 2 displays the relevant research hypotheses.

TABLE 2 Research Hypotheses

Regression model hypothesis	Interpretation
$H\colon \beta_2^q - \beta_1^q > 0$	Block repair is incomplete; not to good-as-new.
$H\colon \beta_3^q > 0$	The repair fraction of good-as-new is decreasing.
$H\colon \beta_4^q > 0$	The failure rate increases with the time since block repair.
$H\colon \beta_5^q > 0$	There is an accelerated failure rate with increasing time since repair.

A comparison of the coefficients for different model and operator combinations would reveal differences in model degeneration rates and/or differences in operator maintenance practices. To include comparisons, an expanded regression equation is defined. Consider the indicator variables:

$U = 1$ for model M_1 and 0 otherwise

$V_1 = 1$ for operator O_1 and 0 otherwise

$V_2 = 1$ for operator O_2 and 0 otherwise.

The regression equation for defining model and operator effects is

$$N(t_1, t_2) = \sum_{j=1}^{5} \beta_j X_j + \left[\sum_{i=1}^{2} \sum_{j=1}^{3} \lambda_{ij} V_i X_j\right] + \left[\sum_{j=1}^{2} \gamma_j U X_{3+j}\right] + \varepsilon \,. \tag{17}$$

An equivalent formulation, which reveals the effect on coefficients, is

$$N(t_1, t_2) = \sum_{j=1}^{3} \left(\beta_j + \sum_{i=1}^{2} \lambda_{ij} V_i\right) X_j + \sum_{j=1}^{2} (\beta_{3+j} + \gamma_j U) X_{3+j} + \varepsilon \,. \tag{18}$$

To simplify notation, consider the vectors

$$\beta = \begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\\ \beta_4\\ \beta_5\end{pmatrix}; \lambda_1 = \begin{pmatrix}\lambda_{11}\\ \lambda_{12}\\ \lambda_{13}\\ 0\\ 0\end{pmatrix}; \lambda_2 = \begin{pmatrix}\lambda_{21}\\ \lambda_{22}\\ \lambda_{23}\\ 0\\ 0\end{pmatrix}; \gamma = \begin{pmatrix}0\\ 0\\ 0\\ \gamma_1\\ \gamma_2\end{pmatrix}.$$

Table 3 shows the variations on the regression equation.

TABLE 3 Model Variations

q	V_1V_2U	Parameters
1	000	$\beta^1 = \beta$
2	100	$\beta^2 = \beta + \lambda_1$
3	010	$\beta^3 = \beta + \lambda_2$
4	001	$\beta^4 = \beta + \gamma$
5	101	$\beta^5 = \beta + \lambda_1 + \gamma$
6	011	$\beta^6 = \beta + \lambda_2 + \gamma$

Table 4 presents the hypotheses for testing the effect of differences in models and operators.

TABLE 4 Comparison Hypotheses

Hypotheses	Interpretation
$H: \lambda_1 \neq 0$	Repair effect for operator O_1 differs from others
$H: \lambda_2 \neq 0$	Repair effect for operator O_2 differs from others
$H: \gamma \neq 0$	The failure rate since repair depends on the aircraft model

Fitted Model

The maintenance policies of a carrier as well as the particular design (components and systems) in an aircraft model are major factors in the operating characteristics of an aircraft. Selecting a single aircraft model from a single carrier removes the complication of varying models and carriers, and thus the assumption of constant degradation and repair parameters is reasonable. This experimental setting is ideal for focusing on the degradation of the operating condition with accumulated use. As such, the data for operator O_2 were used with degradation model (16), because the O_2 fleet consisted only of B737 aircraft.

Because the number of failures is a counting variable, the error variance is not likely to be constant. Therefore, an iteratively reweighted least squares estimation method was used, where the weights were reciprocals of the fitted values (McCullagh and Nelder 1989). The effect of weighting was minimal, so the unweighted sums of squares are reported. The results from fitting the degradation model to the operator O_2 data are given in table 5.

TABLE 5 Fitted Failure Model

Variable	Parameter	Estimate	T	P
X_1	β_1	–1.54	–6.52	0.000
X_2	β_2	6.39	5.03	0.000
X_4	β_4	0.12	6.02	0.000
X_5	β_5	0.02	10.86	0.000

ANOVA

Source	SS	df	F	P
Regression	615.79	4	96.22	0.000
Error	347.21	217		
Total	**963.00**	**221**		

Clearly the overall fit of the change model is strong (F = 96.22). Furthermore, the individual components in the model are highly significant. Maintenance in the observation window and time since maintenance are important factors in predicting the number of unscheduled landings that occur in the window. Of particular significance is the acceleration in the number of repairs (increasing failure rate) as time since repair increases $(\hat{\beta}_5 = 0.02, T = 10.86)$. With reference to Proposition 1, the regression provides the following result.

Result 1 (degradation): The rate of unscheduled landings increases with time since the last block repair.

Furthermore, there is evidence that the block repair is not as good-as-new, because the sign on $X_1 = I_2 - I_1$ is negative and the sign on $X_2 = k_2 - k_1$ is positive. A test on the difference $\hat{\beta}_2 - \hat{\beta}_1 = 7.93$ is highly significant $(P < 0.001)$. However, the data did not allow for a test of diminishing repair fraction. The aircraft in the operator O_2 fleet are relatively new and the maximum is $k = 1$. The regression gives the analogous result for Proposition 2.

Result 2 (imperfect repair): The rate of unscheduled landings decreases after a block repair, with the decrease greater for the initial repair than for subsequent repairs.

To consider the issue of differential effects for the aircraft model and operator, the additional terms with indicator variables were included. Table 6 presents the regression results. The additional sum of squares for models and operators indicated in table 6 are considered after including other effects. That is, the outcome (number of unscheduled landings) was adjusted for the model effect when considering the operator effect, and it was adjusted for the operator effect when considering the model effect. In both cases, the effects are statistically significant (i.e., there is a differential effect for operators and models). In the context of the regression equation, the effect of maintenance on unscheduled landings was not the same for the operators in the study. As well, the effect of the time since maintenance was not the same for the models selected.

TABLE 6 Comparison of Models and Operators

Variable	Parameter	Estimate	T	P
X_1	β_1	–1.46	–4.73	0.000
X_2	β_2	2.84	2.19	0.029
X_4	β_4	0.18	7.51	0.000
X_5	β_5	0.014	7.13	0.000
V_1X_1	γ_{11}	2.44	6.04	0.000
V_1X_2	γ_{12}	1.88	1.39	0.166
V_2X_2	γ_{22}	3.52	1.57	0.116
UX_4	λ_1	–0.14	–3.06	0.002
UX_5	λ_2	0.004	0.99	0.325

ANOVA (OPERATOR)

Source	Additional SS	df	F	P
Operator	131.12	3	13.125	0.000
Error	2,104.73	632		
Total	**4,964.00**	**641**		

ANOVA (MODEL)

Source	Additional SS	df	F	P
Model	40.91	2	6.143	0.002
Error	2,104.73	632		
Total	**4,964.00**	**641**		

Result 3: The relationship between the rate of unscheduled landings and the time since the last block repair and the number of block repairs depends on the aircraft model and operator.

Using the indicator variables, it is possible to write out the fitted equation for each (operator and model) type. The estimates for equation parameters are given in table 7. The equation for operator $O_2(q = 3)$ is slightly different from the equation using only O_2 data, owing to the greater variation in using multiple operators and models. However, it is a good reference for understanding the changes in the equation with the operator/model variations. The biggest operator effect is the difference of $O_3(q = 2,5)$ from the others on the estimate $\hat{\beta}_1$. For aircraft models, the estimate of $\hat{\beta}_4$ is most affected.

TABLE 7 Comparison of Estimates

q	V_1V_2U	β_1^q	β_2^q	β_4^q	β_5^q
1	000	–1.46	2.84	0.18	0.014
2	100	0.99	4.70	0.18	0.014
3	010	–1.46	6.35	0.18	0.014
4	001	–1.46	2.89	0.04	0.018
5	101	0.99	4.70	0.04	0.018

DISCUSSION

The operating condition of aging aircraft has been a hotly discussed topic for more than a decade. The Federal Aviation Administration's position is that the operation of older aircraft is an economic decision and not a safety issue; that is, aircraft can be repaired to a safe operating condition and the cost of those repairs is the issue.

This study considers the trajectory of the health status of an individual aircraft, with an emphasis on episodes where flights are interrupted because of mechanical failures affecting safety. In the context of a model for mechanical failure, two experiments were carried out. In the first experiment, a single model and carrier were analyzed for the potential impact of aircraft age (accumulated use) and repair on schedule reliability. The study assumes that all the selected aircraft are equivalent except for age, and the fleet management practices of the carrier remain consistent over time. In this setting, the variability in the failure rate (unscheduled landings) can be partly attributed to the aging of the aircraft and incomplete repair during preventive maintenance. The second experiment involved multiple operators and aircraft models. With the same failure model, the differential effect of operational practices and aircraft design can be studied.

The following can be concluded from the results of this study.

1. The percentage of variation in unscheduled landings that can be explained by degradation with age and incomplete repair is high.
2. Age (accumulated hours of use) has a statistically significant effect on failures (unscheduled landings), with an increasing failure rate as age increases.
3. The improvement in the operating condition with planned preventive maintenance is not to good-as-new.
4. The relationship between failures and degradation differs from model to model.
5. The relationship between failures and repair differs from operator to operator.

The clear relationship between unscheduled landings and degradation/repair in the regression model has implications for the maintenance policies of operators. The operator has control over the repair intervals—(τ, δ) and the repair effort β—the fraction of good-as-new. The dependence of the regression coefficients on the maintenance parameters (τ, δ, β) is implied in this paper, but that connection can be made more explicit by using the actual derivatives in the series approximations. In that way, changes in the values of the maintenance parameters would translate into changes in the rate of unscheduled landings. Therefore, an operator could explore the outcome (in terms of unscheduled landings) of changes in the repair parameters, for example, the block repair interval.

The purpose of our research was to establish the feasibility of predicting unscheduled landings from data on use and maintenance. An earlier study (Nowlan and Heap 1978) found that 89% of aviation mechanical malfunctions were unpredicted using operating limits or undertaking repeated checks of equipment. The results of this work indicate that important problems in the operation of aircraft can be studied with existing field data. Use of these results in the management of an airline would require additional study, but a step in that direction has been taken here.

REFERENCES

Alfred, L., R. Ellis, and W. Shepard. 1997. The Dynamics of Aircraft Aging, mimeo. Decision Dynamics Inc.

AlgoPlus. 2004. Available at http://www.cislpiloti.org/Technical/GAIN/WGB.

Aven, T. and U. Jensen. 1998. *Stochastic Models in Reliability.* New York, NY: Springer Verlag.

AvSoft. 2004. Available at http://www.avsoft.co.uk/.

Block, H., W. Borges, and T. Savits. 1985. Age-Dependent Minimal Repair. *Journal of Applied Probability* 22:370–385.

Brown, M. and F. Proschan. 1983. Imperfect Repair. *Journal of Applied Probability* 20:851–859.

Friend, C. 1992. *Aircraft Maintenance Management.* Essex, England: Longman Scientific and Technical.

Kanafani, A., T. Keeler, and S. Sathisan. 1993. Airline Safety Posture: Evidence from Service Difficulty Reports. *Journal of Transportation Engineering* 119(4):655–664.

Lincoln, J.W. 2000. Risk Assessments of Aging Aircraft. *Aging Aircraft Fleets: Structural and Other Subsystem Aspects,* RTO Lecture Series 218, RTO-EN-015.

McCullagh, P. and J. Nelder. 1989. *Generalized Linear Models,* 2nd ed. New York, NY: Chapman and Hall.

Nowlan, F.S. and H.F. Heap. 1978. *Reliability Centered Maintenance,* Report No. AD-A066-579. Available from the National Technical Information Service, Springfield, VA.

Phelan, M. 2003. Health Check. *Flight International* 163(4865):26–27.

Sampath, S.G. 2000. Safety and Service Difficulty Reporting, *Aging Aircraft Fleets: Structural and Other Subsystem Aspects,* RTO Lecture Series 218, RTO-EN-015.

Air Traffic at Three Swiss Airports: Application of *Stamp* in Forecasting Future Trends

MIRIAM SCAGLIONE [1,*]
ANDREW MUNGALL [2]

[1] Institute for Economics & Tourism
University of Applied Sciences Valais
TECHNO-Pôle Sierre 3
CH 3960 Sierre
Switzerland

[2] Ecole Hôtelière de Lausanne
Le Chalet-à-Gobet
CH – 1000 Lausanne 25
Switzerland

ABSTRACT

This paper presents forecasting trends for numbers of air passengers and aircraft movements at the three main airports in Switzerland: Zurich, Geneva, and Basel. The case of Swiss airports is particularly interesting, because air traffic was affected in the recent past not only by the September 11, 2001, terrorist attacks but also by the bankruptcy of the national carrier, Swissair, that same year. A structural time series model (STS) is created using *Stamp* software to facilitate forecasting. Results, based on readily available data (i.e., passengers and movements), show that STS models yield good forecasts even in a relatively long run of four years.

INTRODUCTION

Airports are now widely recognized as having a considerable economic and social impact on their surrounding regions. These impacts go far beyond the direct effect of an airport's operation on its neighbors and extend to the wider benefits that access to air transport brings to regional business interests and consumers.

The economic benefits of air transport may be assessed by looking at the full extent of the industry's impact on the overall economy, from the movement

Email addresses:
* Corresponding author—miriam.scaglione@hevs.ch
A. Mungall—Andrew.Mungall@ehl.ch

KEYWORDS: Structural time series, forecast, airports, Switzerland, air traffic.

of passengers and cargo to the economic growth that the industry's presence stimulates in a local area.

In this respect, Switzerland represents an interesting case. First, half of the earnings of the Swiss economy come from abroad. Fast, direct access to the different markets around the world is therefore very important, especially when Switzerland's dependence on exports will increase in the future. In this economy, the industry with the highest share of exports is metallurgy and mechanical engineering. Around 75% of its production is exported. In the Swiss tourism industry, exports represented by the expenses of foreign tourists are also important. Of a total tourism revenue of 22.2 billion Swiss francs (CHF) in 2003 (Swiss Federal Statistical Office 2004), 12.6 billion CHF (60% of the amount) came from foreign tourists.

Second, Swiss air transport has recently been facing rapid change. During the last half of the 1990s, Swiss airports revisited their respective strategies following the decision made by the national carrier, Swissair, in 1996 to concentrate all long-haul flights at Zurich-Unique Airport (ZRH). Also at this time, the Basel/Mulhouse Airport (BSL) made plans to become a European hub or a spoke for Swissair, and, therefore, decided on significant expansion investments. This obliged Geneva International Airport (GVA) to adopt an open-sky policy, where foreign air carriers could benefit from so-called "fifth-freedom" rights, enabling them to provide intercontinental services to and from Geneva. The Swissair bankruptcy in 2001 changed the context, calling the expansion policy of ZRH airport into question.

This study uses the data from the three main Swiss airports—ZRH, GVA, and BSL—to show the overcapacity of two of the three by forecasting these series to 2006. The average share of overall aircraft movements at the three airports was: 46% (±5), 32% (±3), and 22% (±3); for passengers, the share was 59% (±5), 32% (±4), and 9% (±3), respectively.

The structure of the paper is as follows: the next section describes the data used for this analysis, then we provide the forecasts based on three models (one for each airport) to a horizon of 2006. All sections use structural time series (STS) models and the *Stamp* software for STS (Koopman et al. 2000). The last section provides conclusions based on these results.

DATA

For our analysis, we obtained data from the statistical department of each airport studied. Each provided yearly observations from 1949 to 2003, except for GVA, whose data began in 1922. We considered two principal indicators: air movements and passengers.

Traffic Forecasts

We built bivariate models for each airport using the vector of variables

$$y_t = \begin{pmatrix} m_t \\ p_t \end{pmatrix},$$

where m_t denoted the number of movements for a given airport, and p_t denoted the number of passengers. The choice of the bivariate model seems appropriate, because aircraft and passengers belong to the same economic system and this kind of model allows for interaction between the two variables. These models are called Seemingly Unrelated Time Series Equations (SUTSE) and are an extension of univariate forms, with the advantage of allowing for cross-correlation leads between variables. In *Stamp*, SUTSE are particularly appealing because, on the one hand, models with common factors emerge as a special case; on the other, the direct analysis of the unobservable components provides a more efficient forecast and inference (see appendix 1 for details). In this study, the variables are transformed to logarithms so that the model is multiplicative. Such a transformation allows for percentage changes rather than absolute changes in traffic levels and also helps to stabilize the variances of the variables.

Analysis for GVA

For GVA, while annual data are available from 1922, the first major facility was built in 1949. Given that the period 1949 to 1952 was one of transition, we used data from 1953 to 2002. The observation for 2003 was used to evaluate the probability of a structural break using the post-predictive features of *Stamp* (Koopman et al. 2000, pp. 39–40).

Table 1 shows the hyper-parameters of the model (table A1 in the appendix provides an interpretation of these values). Table 2 shows statistics tests normally used to evaluate the goodness-of-fit of the model. In the case of GVA, only the Box-Ljung (test of residual serial correlation) statistic is slightly significant (p-value = 0.08) for passengers. The model is a local linear trend and a common cycle (table 3).

Table 1 also provides the list of intervention variables.[1] There is no intervention for passenger series, but there are two interventions in the aircraft movement component. The first occurs in 1956 and is a positive level shift, probably explained by the increase of movements owing to the start of the jet aircraft era. The outlier intervention (AO) in 1967 is difficult to explain; however, we retained it in the model for consistency.

The analysis of the components obtained in the STS modeling (i.e., trend, slopes, and cycles) highlights some interesting features of the phenomena under study. Figure 1 shows some of the components of the models. First, the slopes are parallel in logarithms, suggesting that the rates of growth, though different, have a parallel evolution. In statistical terms, it means that the system composed of the passengers and movements is co-integrated to an order of (2,2) and there is a combination that is stationary (see Koopman et al. 2000, p. 86; Song and Witt 2000, p. 56). As the data are in logarithms, this component represents the rate of growth and can be read as tracking its acceleration and deceleration.

[1] See appendix 1 for the definition of the intervention variables.

TABLE 1 Hyper-Parameters Given as Standard Deviations and Interventions

	Airports					
	GVA		ZRH		BSL	
Variable	Passengers	Movements	Passengers	Movements	Passengers	Movements
Level	3.95E–02	7.3E–03	nil	nil	3.9E–02	2.0E–03
Slope	1.2E–02	8.6E–03	1.9E–02	6.7E–03	1.8E–02	2.2E–02
Cycle	ρ = 0.83		ρ = 0.88		ρ = 0.98	
	2.1E–02	1.6E–02	1.2E–02	2.4E–02	3.6E–02	3.0E–02
Irregular	0.0E+00	2.5E–02	1.9E–02	6.7E–03	9.8E–03	3.5E–02
Interventions	nil	+LS 1956**	+AO 1957**	–AO 1957*	nil	nil
	nil	+AO 1967**	–LS 2002***	–LS 2002**	nil	nil
Model	LLT+cycle		Smooth trend		LLT+cycle	
Cointegration	C(2,2)+common cycles		nil		C(2,2)	

Key: *** = p-value < 0.01; ** = p-value < 0.05; and * = p-value < 0.1; the sign in the intervention row indicates the sign of its coefficient.

TABLE 2 Statistical Values of the Goodness-of-Fit Test

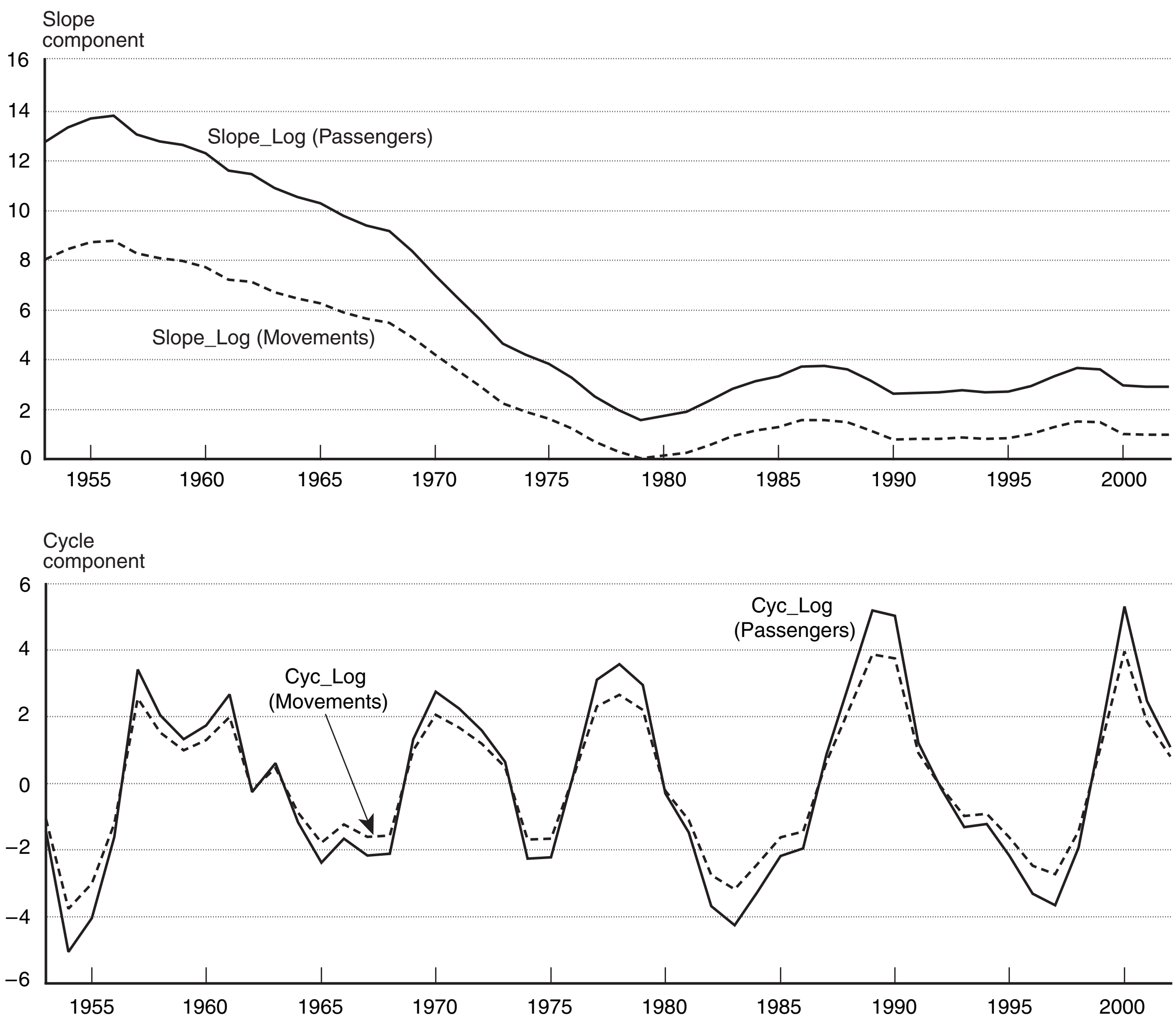
	Airport					
	GVA (1953–2002)		ZRH (1953–2003)		BSL (1956–2002)	
Variable	Passengers	Movements	Passengers	Movements	Passengers	Movements
Standard error	5.17E–02	4.43E–02	4.39E–02	3.75E–02	7.61E–02	8.00E–02
Normality	5.92	0.12	3.92	2.31	2.52	0.61
Heteroskedasticity	0.51	0.37	1.62	0.51	0.44	0.31
Durbin-Watson	1.98	1.92	1.89	1.89	1.51	1.78
Box-Ljung	11.20*	6.45	7.59	12.84*	5.40	7.64
R-squared	0.36	0.64	0.63	0.34	0.51	0.48

Key: * = p-value < 0.1.
Note: For definitions, see the appendix.

TABLE 3 Cycle Parameters and Final Growth Rates for Three Swiss Airports

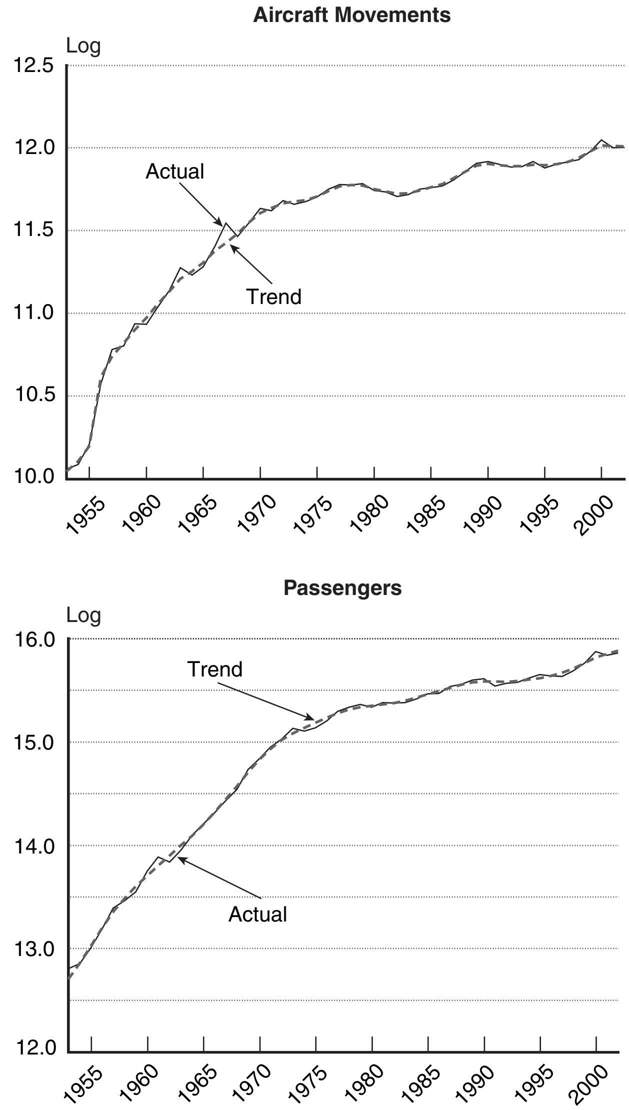
Airport	Variable	Cycles: Amplitude at the end of the series	Cycles: Period (years)	Cycles: Covariance	Slope: Growth rate at the end of the series
GVA (1953–2002)	Passengers	2.2	11.90	1.0	2.9
	Movements	1.6			1.0
ZRH (1953–2003)	Passengers	3.6	8.96	4.2E–01	–0.9
	Movements	7.2			1.1
BSL (1956–2002)	Passengers	18.2	9.11	9.8E–01	2.0
	Movements	11.1			–1.9

FIGURE 1 Slope Components and Cycles for Aircraft Movements and Passengers at GVA: 1953–2002



The growth rates gradually declined from the mid-1950s to the mid-1970s and then stabilized at about 3% for passengers and 1% for movements (see the last column in table 3). This difference in rates clearly reflects the increase in the jet aircraft era. The mid-1950s brought the prospect of commercial jet airliners in the near future, with all it would entail in terms of longer runways and greater

FIGURE 2 Trend Plus Cycle and Actual Values (in Logarithms) for Movements and Passengers at GVA: 1953–2002



terminal capacity. The Swiss and French authorities reached an agreement concerning an exchange of land with France. Provision was also made for a sector of the future terminal to become a "French Airport," linked to Ferney-Voltaire in France by an extra-territorial connection. The agreement was ratified by the Federal Assembly in 1956 and by the French Parliament in 1958.

Other important features summarized in table 3 for GVA are the existence of a cycle of approximately 12 years that is stochastic, but with a correlation of 1; this means that the two cycles (passengers and air movements) move together. The table also shows that the amplitude of the cycle (at the end of the series) is less than 2.5% of the trend for passengers and less than 2% for movements. Figure 2 shows the trend and the actual values in logs with the contribution of the cycles.

Table 4(A) shows the number of passengers forecast and observed for 2003 and 2004. The forecasts underestimate the observed figures by about 3% for 2003 and 6% for 2004. Table 4(B) shows that aircraft movements observed and forecast are very close for both 2003 and 2004, both approximately 1.6 hundreds of thousands.

In spite of the slightly high error for passengers, table 4(A) shows that the numbers observed remain inside the confidence interval of one standard error, and therefore the model does not need to be reviewed. In 2003, GVA outperformed the 2000 record for passengers; this upward tendency was confirmed in 2004. This increased passenger level can be explained by the innovative policy carried out by the GVA authorities and the reinforcement of the presence of low-cost carriers that have a high passenger load rate for their flights. As an example, on April 22, 2004, the GVA Board decided to adopt a loyalty policy for both the old and new companies operating in the airport. Under this policy, GVA would return up to 40% of the airport taxes (excluding the share relating to security) to all companies that signed a commitment to operate from the airport for three to five years. At the same time, GVA decided to segment its terminals. The principal terminal remains a conventional one, and GVA renovated the old terminal and offered to let all companies (even though low-cost companies suggested this measure) use it at a lower tax level than the principal one. Table 4 (B) shows that GVA will need more than two years to return to the maximum level of movements registered in 2000 (1.71 hundreds of thousands).

Figure 3 also shows the forecasts for 2004 to 2006, which indicate a strong upward trend for passengers and a slight upward trend for aircraft movements. Once more, this difference in the speed of evolution between movements and passengers can be explained by the aggressive GVA policy in trying to capture the low-cost market (e.g., *easyJet,* which has an excellent aircraft occupancy rate). Thus, with 25% of the market share of GVA traffic in 2003,

TABLE 4 Forecasts of Numbers of Passengers and Aircraft Movements

A. PASSENGERS (millions)

Airport	Year	2003	2004	2005	2006
GVA (1953–2002)	Forecast	7.86	8.05	8.29	8.19
	Observed	8.09	8.59		
	RMSE	0.43	0.69	1.02	1.33
	Error	–2.84%	–6.29%		
ZRH (1953–2003)	Forecast	in sample	16.41	16.55	16.59
	Observed	17.02	17.25		
	RMSE	in sample	1.62	2.41	3.27
	Error	in sample	–4.87%		
BSL (1956–2002)	Forecast	2.85	2.84	3.07	3.49
	Observed	2.48	2.59		
	RMSE	0.23	0.44	0.70	0.10
	Error	14.92%	9.65%		

B. AIRCRAFT MOVEMENTS (hundreds of thousands)

Airport	Year	2003	2004	2005	2006
GVA (1953–2002)	Forecast	1.64	1.66	1.67	1.69
	Observed	1.63	1.67		
	RMSE	0.07	0.10	0.14	0.17
	Error	0.61%	–0.60%		
ZRH (1953–2003)	Forecast	in sample	2.78	2.95	3.16
	Observed	2.69	2.67		
	RMSE	in sample	0.20	0.29	0.37
	Error	in sample	4.12%		
BSL (1956–2002)	Forecast	1.01	0.97	0.99	1.04
	Observed	1.00	0.78		
	RMSE	0.09	0.15	0.22	0.30
	Error	1.00%	24.36%		

Key: RMSE = root mean squared error.

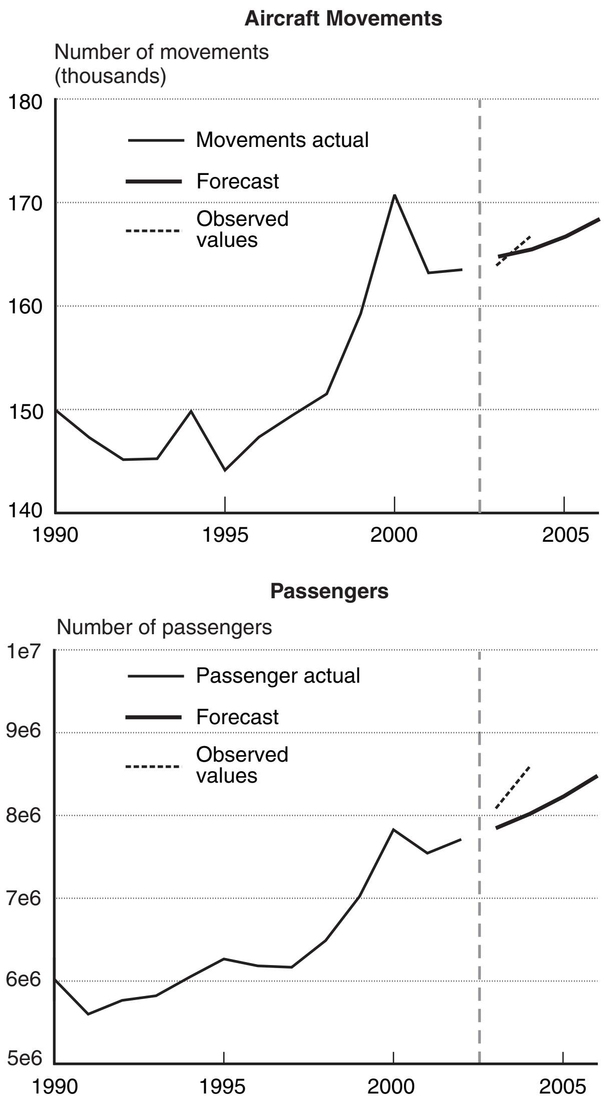
easyJet has become, for the first time, the most important carrier in Geneva. The market share of the national airline, Swiss, amounted only to 21.3%.

Analysis for ZRH Airport

For ZRH airport, we considered data from 1953, which coincided with the formal inauguration of the present location (also called Kloten Airport). During the 1980s and 1990s, the airport experienced rapid expansion. The number of passengers reached 12 million in 1990 and 22 million in 2000. Thus, in 1995 the Zürich electorate accepted a further expansion program for the airport, with a new terminal, a new airside center (linking the different terminals) and underground facilities. Initially, the expectation was to complete the expansion by 2005, with the capacity of the airport increasing from 20 million to 40 million passengers per year. As a consequence of the events of 2001, which form the object of this analysis, this extension phase is being implemented more slowly and Terminal B (capacity of about 5 million passengers) has been closed down.

The model calculated on the sample data from 1953 to 2002 shows that the observed totals for 2003 lay outside the forecast confidence interval of one standard deviation. Therefore, the authors reviewed the model, taking 2003 as the last observation, and a level shift intervention in the model for 2002. Table 1 shows that the intervention has a

FIGURE 3 Forecast Figures for Movements and Passengers at GVA



significant negative coefficient for both series, passengers and movements, indicating that the shift in both trends is decreasing. The model is a smoothed trend with a drift.

Figure 4 shows the slope, namely the rate of growth for the series. For movements, the range was about 5% (from 6% to 1%) and was quite steady. For passengers, the range was about 19% (from 18% to –1%). The figure also shows the external shocks for both series. Indeed, the rate of growth ZRH passengers becomes negative in the same year that Swissair went bankrupt.

Analysis for BSL Airport

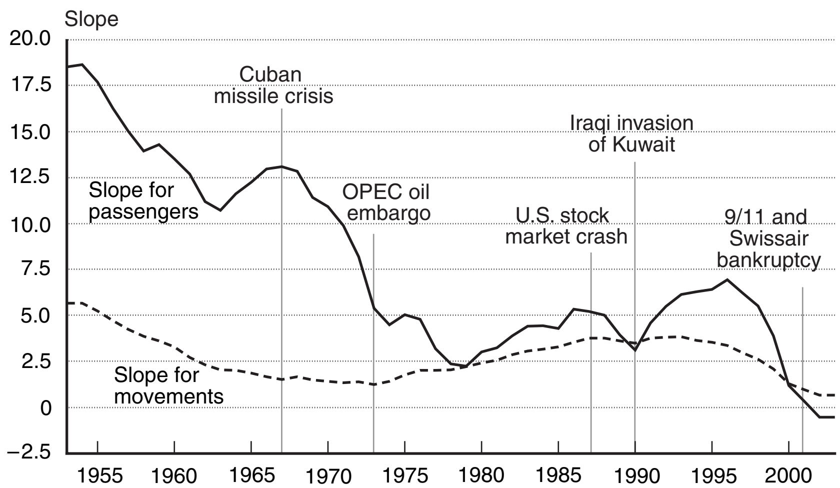
The analysis of BSL airport uses observations from 1956 to 2002. The reason for chosing 1956 is that at that time the facilities development process was quite mature. When the airport reached 3 million passengers per year at the end of 1998, expansion seemed to be essential and urgent. Extension work on the terminal buildings will allow for further expansion in the number of passengers in the future, to a capacity of 5 million per year.

The model is a local linear trend, plus a cycle, having a common slope, which means a cointegration (2,2) for the two series. The estimated growth rate of the fitted stochastic trend is positive for passengers (about 2%) and negative for movements (–2%) at the end of the sample period.

Table 4(A) shows a high error in the overestimation for passengers in 2003, which could be due to the drastic reduction of flights by the national carrier, Swiss (elimination of 11 destinations and 7 transfer flights since March 2003), which strongly affected the number of transit passengers. The good performance in the forecast of movements could be explained, in part, by the increasing number of charters (over 3% against 2002) following the arrival of new carriers (EuroAirport 2004).

The 2004 forecast is better for passengers than for movements, nevertheless both forecast figures remain inside the confidence interval of one standard deviation. On the one hand, the number of passengers on scheduled flights increased by 8% against 2003, given the arrival of *easyJet* offering four new destinations. On the other hand, the decrease in the number of movements is explained by the increasing load rate and the use of aircraft with larger capacity (EuroAirport 2005). The former fact explains the apparent contradiction of the increasing number of passengers despite a decreasing number of movements. Finally, this high error in the forecast suggests that Basel Airport is in a structural change phase (i.e., the percentage of transit passengers in 2004 was 2%, whereas in the past it was approximately 28%). Nevertheless, BSL seems to be growing once again owing to a policy centered more on low-cost carriers and charters and much less on its original vocation of being a Euro-Hub.

FIGURE 4 Slope Components for the Zurich Airport Model with External Shocks Chronology: 1953–2002



CONCLUSION

The forecasts here show that neither ZRH nor BSL seems likely to return to the level of the record year of 2000, either for passengers or for aircraft movements, by 2006. The only airport that was able to beat the record numbers achieved in 2000 was GVA, but only in the case of passengers.

The innovative policy carried out by GVA was a good solution to overcome the 2001 crises of the bankruptcy of Swissair and the U.S. terrorist attacks. This being said, GVA had begun to rethink its strategy earlier than the other two airports, owing to the decision of the national carrier to concentrate all long-haul flights at ZRH in 1996.

The high errors in the forecast figures for BSL are a result of the structural changes taking place at that airport since 2003; namely, an evolution away from being a spoke and toward becoming a city-to-city European airport. Therefore, the analysis of those differences could be a tool for assessing the effectiveness of the measures undertaken by BSL. In fact, table 4(A) shows that a slightly decreasing trend was forecast for passengers between 2003 and 2004, whereas the observed figures show the opposite. This may be due, at least in part, to the success of the new policy adopted by the Board of BSL.

Finally, the use of *Stamp* software on the series of passengers and movements through a SUTSE model appears to be an interesting tool for forecasting air transportation data. On the one hand, the forecasts are good if there are no structural changes (as in the case of BSL); on the other hand, analysis of the components (i.e., trends, slopes, and cycles) gives a good insight into the dynamic of the series. Moreover, the data used (i.e., passengers and movements) are easily available.

ACKNOWLEDGMENTS

The authors wish to thank the following individuals for their invaluable support: Merrick Fall (EHL), Regula Catsantonis (Unique Airport Verkehrsdaten/Statistik), Robert Weber (Statistical Department of Aéroport International de Genève), Valérie Meny (Statistical Department of EuroAirport), and Merk Jürg (Stab/Statistik Bundesamt für Zivilluftfahrt). Thanks also go to Professor Andrew Harvey (Cambridge University) and to the anonymous referees for their helpful comments on the earlier version of this study.

REFERENCES

EuroAirport. 2004. Face à Une Situation Économique Perturbée et à la Chute de Son Trafic, l'EuroAirport s'Adapte aux Nouveaux Besoins du Marché et Affirme sa Volonté de Renouer avec la Croissance en 2004.

_____. 2005. Evolution Positive du Traffic à l'EuroAirport: Innovation et Adaption Réussie aux Nouveaux Besoins du Marché. En 2005: Retour de la Croissance.

Harvey, A.C. 1990. *Forecasting, Structural Time Models, and the Kalman Filter.* Cambridge, UK: Cambridge University Press.

Koopman, S.J., A.C. Harvey, J.A. Doornik, and N. Shepard. 2000. *Stamp: Structural Time Series Analyser, Modeller and Predictor.* London, England: Timberlake Consultants, Ltd.

Song, H. and S.F. Witt. 2000. *Tourism Demand Modeling and Forecasting*. Amsterdam, The Netherlands: Pergamon.

Swiss Federal Statistical Office. 2004. *Annuaire Statistique de la Suisse.* Edited by Neue Zürcher Zeitung. Zürich, Switzerland.

APPENDIX

The structural time series model aims to capture the salient characteristics of stochastic phenomena, usually in the form of trends, seasonal or other irregular components, explanatory variables, and intervention variables. This model can reveal the components of a series that would otherwise be unobserved, greatly contributing to thorough comprehension of the phenomena. We describe here only the elements necessary for this study; for a complete description see Harvey (1990) and Koopman et al. (2000). An STS multivariate model may be specified as:

Observed variables = trend + cycle + intervention + irregular

The algebraic form for the N series is:

$$\boldsymbol{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\psi}_t + \Lambda \mathbf{I}_t + \boldsymbol{\varepsilon}_t$$

$$\boldsymbol{\varepsilon}_t \sim NID(0, \boldsymbol{\Sigma}_\varepsilon^2) \qquad t = 1,\ldots,T \tag{1}$$

Unless otherwise stated, the elements in equation (1) are $(\boldsymbol{N}\, x\, \mathbf{1})$ vectors,

where $\boldsymbol{y}_t$ = the vector of observed variables,

$\boldsymbol{\mu}_t$ = the stochastic trend,

$\boldsymbol{\psi}_t$ = the cycle,

$\boldsymbol{\Lambda}$ = the $\boldsymbol{N}\, x\, \boldsymbol{K}^*$ matrix of coefficients for the interventions, and

$\mathbf{I}_t$ = the $K^* \times \boldsymbol{1}$ vector of interventions.

The stochastic trend is intended to capture the long-trend movements in the series and trends other than linear ones, and is composed of two elements: the level (2) and the slope (3). The trend described below allows the model to handle these.

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}_{t-1} + \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t$$

$$\boldsymbol{\eta}_t \sim NID(0, \boldsymbol{\Sigma}_\eta^2) \qquad t = 1,\ldots, T \tag{2}$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{s}_t$$

$$\boldsymbol{s}_t \sim NID(0, \boldsymbol{\Sigma}_s^2) \qquad t = 1,\ldots, T \tag{3}$$

If the variances of the irregular components $\boldsymbol{\varepsilon}_t$ in (1), the disturbances of the level $\boldsymbol{\eta}_t$ in (2), and at least one of the slope terms ζ_t are simultaneously strictly positive, the model is a local linear trend.

When the level component is fixed and different from zero, and when the two other variances are not zero, the model is called a "smoothed trend with a drift."

For a univariate model, the cycle $\boldsymbol{\psi}_t$ has the following statistical specification:

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_{t-1} \end{bmatrix},$$
$$t = 1,\ldots, T \tag{4}$$

where λ_c is the frequency, in radians, in the range $0 < \lambda_c < 1, \kappa_t$, and κ_{t-1}^* , are two mutually uncorrelated white noise disturbances with zero means and common variance σ_κ^2, and ρ is the damping factor. The period of the cycle is $2\pi/\lambda_c$.

Cycles can also be introduced into a multivariate model. The disturbances may be correlated; the same, incidentally, can occur with any components in multivariate models. Because the cycle in each series is driven by two disturbances, there are two sets of disturbances and *Stamp* assumes that they have the same variance matrix (Koopman et al. 2000, p. 76), that is:

$$E\left(\boldsymbol{\kappa}_t \boldsymbol{\kappa}'_t\right) = E\left(\boldsymbol{\kappa}_t^* \boldsymbol{\kappa}_t^{*\prime}\right) = \boldsymbol{\Sigma}_\kappa$$

$$E\left(\boldsymbol{\kappa}_t \boldsymbol{\kappa}_t^{*\prime}\right) = 0, t = 1,\ldots, T \tag{5}$$

where $\boldsymbol{\Sigma}_\kappa$ is a $N \times N$ variance matrix.

Stamp has pre-programmed the following exogenous intervention variables used in this study:

1. AO: it is an unusually large value of the *irregular* disturbance at a particular time. It can be captured by an *impulse* intervention variable that takes the value of the outliers as one at that particular time, and zero elsewhere. If t_{ao} is the time of the outlier, then the exogenous intervention variable $\mathbf{I}_{t_{ao}}$ has the following form:

$$\mathbf{I}_{t_{ao}} = \begin{cases} 1 \text{ if } t = t_{ao} \\ 0 \text{ if } t \neq t_{ao} \end{cases} \qquad t = 1,\ldots, T$$

2. LS: this kind of intervention handles a *structural break* in which the level of the series shifts up or down. It is modeled by a *step* intervention variable that is zero before the event and one after it. If t_{LS} is the time of the level shift, then the exogenous intervention variable $\mathbf{I}_{t_{LS}}$ has the following form:

$$\mathbf{I}_{t_{LS}} = \begin{cases} 0 \text{ if } t < t_{LS} \\ 1 \text{ if } t \geq t_{LS} \end{cases} \qquad t = 1,\ldots, T$$

Output

Table A-1 illustrates the nature of the outputs used in the main text. The figures are taken from table 1 in the text.

Diagnostics

The diagnosis test statistics for a single series in an STS model are the following (see Koopman 2000, pp. 182–183):

- Normality test: the Doornik-Hansen statistic, which is the Bowman-Shenton statistic with the correction of Doornik and Hansen. Under the null hypothesis that the residuals are normally distributed, the 5% critical value is approximately 6.0.
- Heteroskedasticity test: A two-sided *F*-test that compares the residual sums of squares for the first and last thirds of the residuals series.
- DW: The Durbin-Watson statistic for residual autocorrelation; under the null hypothesis, it is distributed approximately as $N(0,1/T)$, T being the number of observations.
- Box-Ljung *Q*-statistic: A test of residual serial correlation, based on the first P residual autocorrelations and distributed as chi-square, with $P–n+1$ df, when n parameters are estimated.

TABLE A-1 Interpretation of the Hyper-Parameters of the Model for Geneva International Airport

Variables	Passengers	Movements	Comments
Level	3.95E–02	7.3E–03	Variances of level terms
Slope	1.2E–03	8.6E–03	Variances of slope terms
Cycle	ρ = 0.83		Damping factor
	2.1E–02	1.6E–02	Variances of cycles
Irregular	0.0E+00	2.5E–02	Variances of irregular terms
Interventions	nil	+LS 1956**	Interventions: timing and coefficient's sign (+/–)
	nil	+AO 1967**	
Model	LLT+cycle		
Cointegration	C(2,2)+common cycles		

Key: ** = *p*-value < 0.05.

Improved Estimates of Ton-Miles

SCOTT M. DENNIS

Bureau of Transportation Statistics
Research and Innovative Technology Administration
U.S. Department of Transportation
400 7th St. SW, Room 3430
Washington, DC 20590
Email: scott.dennis@dot.gov

ABSTRACT

This paper describes recent improvements in measuring ton-miles for the air, truck, rail, water, and pipeline modes. Each modal estimate contains a discussion of the data sources used and methodology employed, presents a comparison with well-known existing estimates for reference purposes, and discusses the limitations of the data. The resulting estimates provide more comprehensive coverage of transportation activity than do existing estimates, especially with respect to trucking and natural gas pipelines.

INTRODUCTION

The Bureau of Transportation Statistics (BTS) is improving some of its basic estimates of transportation activity. This paper describes proposed ton-mile estimates for the air, truck, rail, water, and pipeline modes. Each modal estimate contains a discussion of the data sources used and methodology employed, presents a comparison with well-known existing estimates for reference purposes, and discusses the limitations of the data. This paper should be viewed as part of a continuing series of steps forward. Additional planned work will allow BTS to further improve its basic estimates of transportation activity.

KEYWORDS: Transportation measurement, ton-miles.

CONCEPT

Ton-miles is the primary physical measure of freight transportation output. A ton-mile is defined as one ton of freight shipped one mile, and therefore reflects both the volume shipped (tons) and the distance shipped (miles). Ton-miles provides the best single measure of the overall demand for freight transportation services, which in turn reflects the overall level of industrial activity in the economy. In addition, a ton-mile estimate is necessary in order to construct other estimates of transportation system performance, such as energy efficiency and accident, injury, and fatality rates.

Domestic ton-mile estimates are usually developed by aggregating data for individual freight transportation modes. Data for air freight, railroad, and water transportation are readily available as a result of government provision of infrastructure or residual economic regulation. Comprehensive pipeline data are difficult to obtain, because a significant percentage of pipeline traffic is "in-house" transportation for companies that produce and refine oil. Ton-mile data for the trucking sector are even more problematic due to the large number of shippers, receivers, and trucking firms, as well as the substantial percentage of in-house trucking traffic. All data sources suffer, at least to some degree, from gaps in the desired scope of coverage.

The Eno Transportation Foundation has published historical ton-mile estimates for many years (Eno 2002, p. 42), but no longer does so. In more recent years, BTS has provided alternative estimates in *National Transportation Statistics* (*NTS*) (USDOT BTS 2003). But due to the problems described above, these well-known sources do not appear to provide complete, reliable estimates of this basic transportation measure. BTS, therefore, undertook a research program to address these shortcomings. This paper presents the results of this research.

DATA SOURCES

BTS developed its improved estimates of domestic ton-miles (traffic within and between the 50 states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands) to maintain compatibility with other U.S. Department of Transportation Strategic Plan data. These annual ton-mile estimates illustrate long-term trends. Comprehensive coverage is achieved by combining reported data from established sources, estimates from surveys, and calculations based on certain assumptions. Table 1 briefly compares the scope of the improved BTS ton-mile estimates with the *NTS* and Eno estimates, while figure 1 presents all three estimates for all modes (also see appendix table A1). The *NTS* and Eno estimates are not available for the most recent years.

Air

Figure 2 shows air freight ton-mile data from the three datasets (also see appendix table A2). The improved BTS data are compiled in *Air Carrier Traffic Statistics Monthly* (USDOT BTS 1990–2003), which presents the results of the T-100 reporting system, supplemented by special tabulations of data on domestic all-cargo operators from the Federal Aviation Administration (FAA).

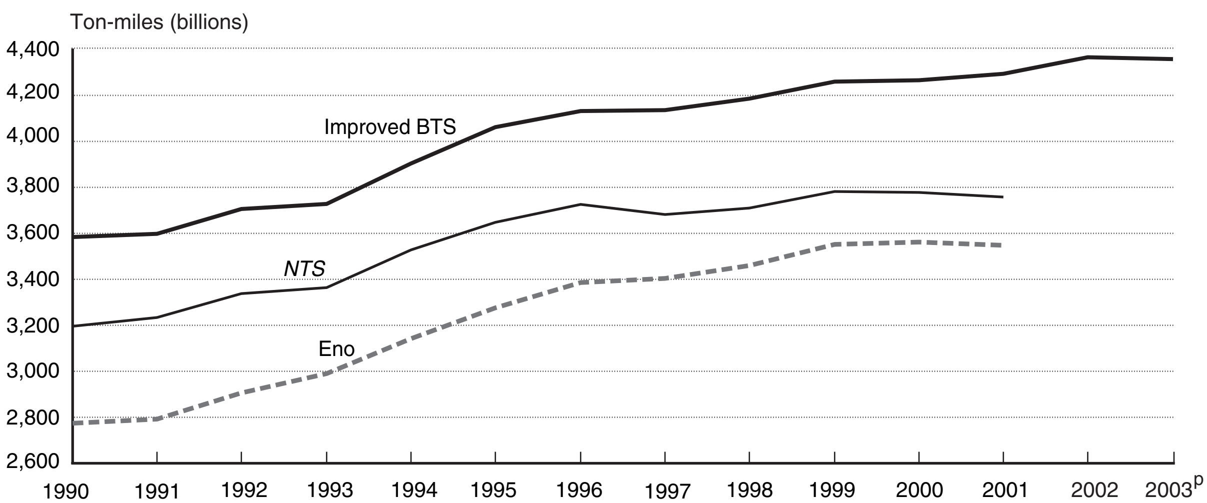
The T-100 data represent the population of all domestic freight traffic for Section 401 air carriers, which operate planes with a passenger seating capacity of more than 60 seats or a maximum payload capacity of more than 18,000 pounds. These data include the vast majority of all domestic air freight traffic. As a result of a BTS rulemaking, data for smaller carriers have been included in this source starting with the fourth quarter of 2003. The inclusion of smaller carriers does not substantially affect the value of the data series. Domestic all-cargo operators (Section 418 carriers) have been gradually integrated into *Air Carrier Traffic Statistics Monthly.* The FAA data captured those carriers who had not yet reported in *Air Carrier Traffic Statistics Monthly,* thus allowing representation of the full population of domestic all-cargo operators.

BTS's proposed estimates of air freight ton-miles are essentially the same as the Eno estimates. Neither estimate includes private carriage of air freight or air freight forwarders who do not use T-100 reporting carriers. These exceptions account for well under 5% of all air freight traffic. The substantial difference between the two data series in 2001 is due apparently to Eno's use of preliminary data.

TABLE 1 Comparison of Annual Data Coverage

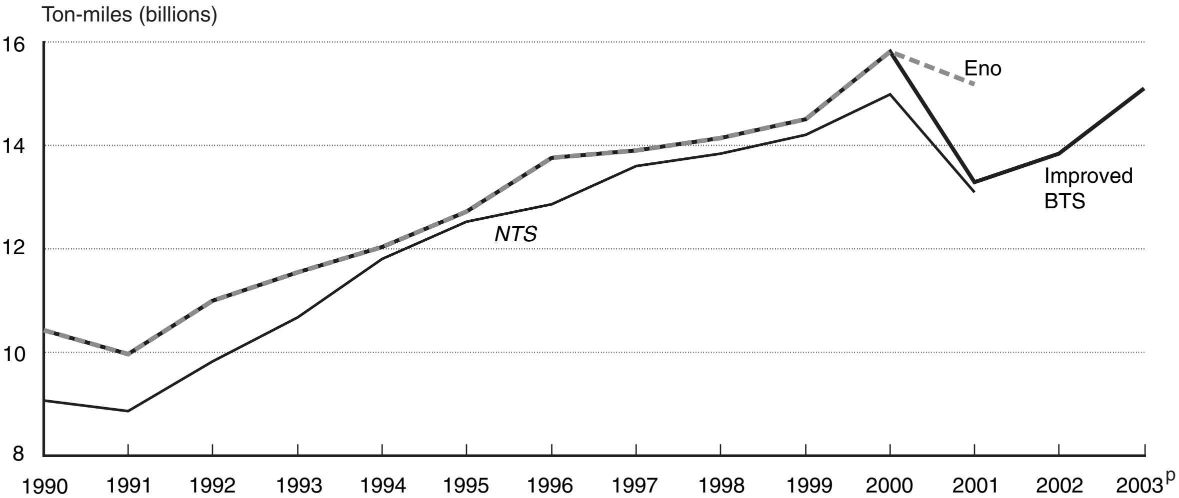
Mode	Improved BTS	*NTS*	Eno
Air	Section 401 carriers Section 418 carriers	Section 401 carriers	Section 401 carriers Section 418 carriers
		Excludes Section 418 carriers	
	Excludes private carriage and some freight forwarders	Excludes private carriage and some freight forwarders	Excludes private carriage and some freight forwarders
Truck	Excludes household, retail, service, government, and certain noncommercial freight shipments	Excludes intracity traffic	Excludes intracity traffic
Railroad	All traffic	Excludes small railroads	All traffic
Water	All domestic traffic	All domestic traffic	Excludes coastal traffic and traffic to and from Alaska, Hawaii, and Puerto Rico
Pipeline	Oil and oil products pipelines Natural gas pipelines	Oil and oil products pipelines	Oil and oil products pipelines
	Excludes chemical and coal slurry pipelines	Excludes natural gas, chemical, and coal slurry pipelines	Excludes natural gas, chemical, and coal slurry pipelines

FIGURE 1 All Ton-Miles



Key: p = preliminary.

FIGURE 2 Air Freight, Express, and Mail Revenue Ton-Miles



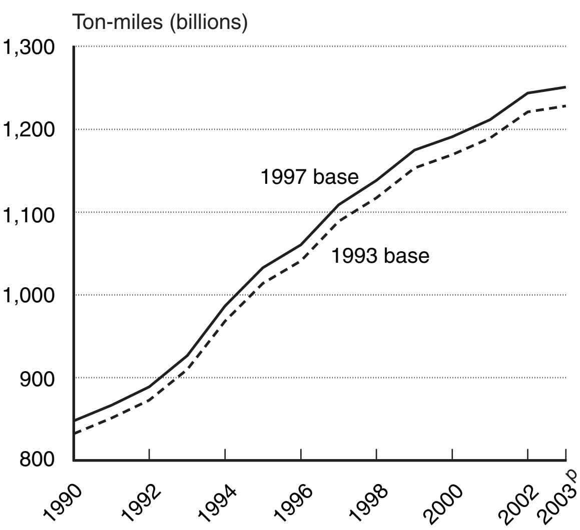
Key: p = preliminary.

Truck

Oak Ridge National Laboratory produced estimates of truck ton-miles based on the 1993 and 1997 Commodity Flow Surveys (CFS) (USDOT and USDOC 1993 and 1997), supplemented with data on farm-based shipments and imports arriving by truck from Canada and Mexico. *Transportation Statistics Annual Report* provides this 1997 estimate of truck ton-miles (USDOT BTS 2000, p. 124). To produce the improved BTS estimate, the 1993 and 1997 estimates were updated and backdated using intercity and intracity vehicle miles-traveled (VMT) for single-unit and combination trucks, as reported in *Highway Statistics* (USDOT FHWA 1990–2003). Figure 3 presents the resulting estimates (also see appendix table A3). The trend in both series is the same, because the same VMT data were used to update each series. After making these adjustments for different time periods and population coverage, the difference between the 1993 and 1997 estimates is less than 2%.[1]

[1] The VMT data include a substantial amount of truck traffic that is outside the scope of the CFS, e.g., shipments by households and retail, service, utility, and government establishments (including the U.S. Postal Service); and certain noncommercial freight shipments, e.g., construction traffic and municipal solid waste. The VMT data can, therefore, be taken to provide a reasonable estimate of the trend in truck ton-miles, but not the level, and should not be used to make inferences about operational parameters such as empty mileage or average load per truck.

FIGURE 3 Development of Truck Ton-Mile Estimates



Key: p = preliminary.

The CFS captures export movements, as well as movements of imports once they reach their first domestic destination, such as a warehouse. In order to provide a more complete estimate of truck traffic, the data in figure 3 were further adjusted to reflect truck ton-miles from maritime movements prior to reaching their first domestic destination. The number of loaded 20-foot equivalent unit (TEU) containers shipped through U.S. ports is reported in *U.S. Waterborne Container Traffic by Port* (U.S. Army Corp of Engineers WCSC 2003). These figures were then divided by 2.4 to convert them to an equivalent number of 48-foot trucks. Estimates of the percentage of import traffic, the truck share of import traffic, miles to the first domestic destination, and tons per truck for East, Gulf, and West Coast ports were obtained through interviews with port personnel in New York, Houston, and Los Angeles, respectively. The resulting estimates added between 7 billion and 12 billion truck ton-miles each year. This represents approximately 1% of all truck ton-miles currently estimated.

Figure 4 shows trucking ton-mile estimates (also see appendix table A4). The improved BTS estimates are based on the Oak Ridge National Laboratory supplement to the 1997 study, which is the more recent of the two studies. The improved BTS estimate is about 10% higher than the *NTS* and Eno estimates, each of which reflects only intercity truck traffic. Therefore, the improved BTS estimate provides a more comprehensive estimate of truck traffic.

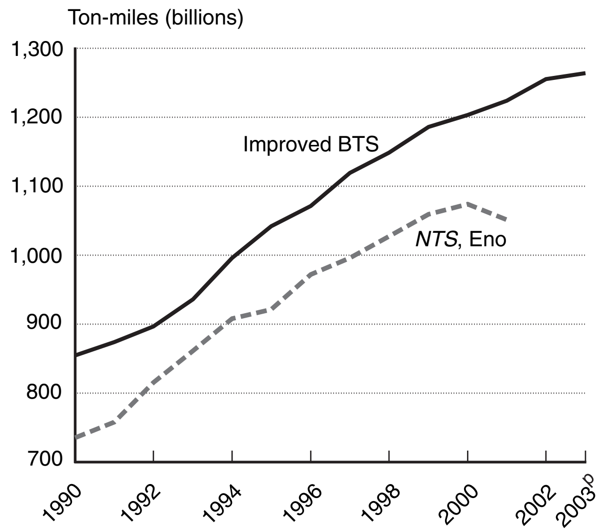
The CFS data used to construct the improved trucking ton-miles estimate exclude shipments by households and retail, service, utility, and government establishments (including the U.S. Postal Service); and certain noncommercial freight shipments, such as construction traffic and municipal solid waste. The existing *NTS* and Eno estimates do not include intracity traffic. Therefore, it appears that a significant percentage of truck VMT and a somewhat smaller percentage of truck ton-miles are not included in any of these estimates. Clearly more work is needed in this area.

Railroad

BTS developed its improved railroad ton-miles estimates using data from the *Carload Waybill Sample* (USDOT STB 1990–2003). The population estimate in this source is based on a 500,000 record sample of all traffic terminating on all railroads in the United States. The sample implicitly includes traffic originating on U.S. railroads and terminating on Mexican railroads, because almost all such traffic is rebilled to U.S. border crossings.[2]

[2] Traffic originating on U.S. railroads and terminating on Mexican railroads is treated for accounting purposes as if it terminated at the U.S. border crossing, and is therefore included in the *Carload Waybill Sample*. This practice is known as rebilling.

FIGURE 4 Truck Ton-Miles

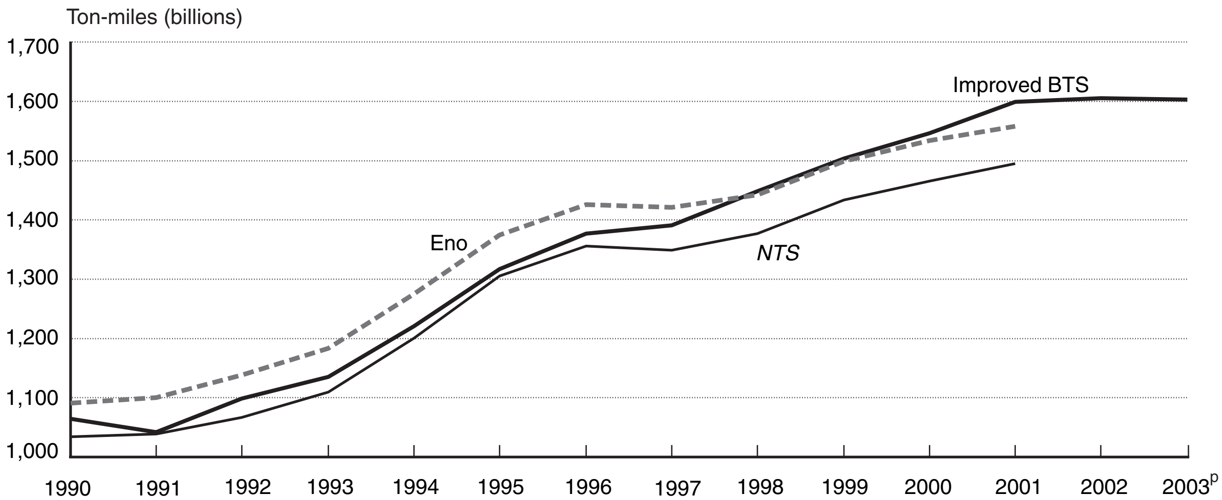


Key: p = preliminary.

Population data on the tonnage of railroad shipments originating in the United States and terminating in Canada come from *Transportation in Canada* (Transport Canada 1990–2002) for years prior to 2003. The average length of haul for U.S. railroad shipments was applied to this tonnage to obtain an estimate of U.S. railroad ton-miles for shipments terminating in Canada. This assumption seems reasonable, because even though much of this traffic originates in states bordering Canada, more distant states such as California, Texas, and Georgia are also among the 10 largest originating states. The *Carload Waybill Sample's* improved coverage of Canadian terminations of rail shipments originating in the United States allows estimates of all railroad ton-miles for 2003 and subsequent years.

Figure 5 illustrates the railroad ton-mile estimates (also see appendix table A5). From 1998 to 2001 (the most recent years for which all three estimates are available), the improved BTS estimates are about 5% greater than the *NTS* estimates, which include only Class I railroads; about 1% greater than the *Waybill* estimates, which do not include Canadian terminations; and almost identical to the Eno estimates, which include both non-Class I railroads and Canadian terminations. The increase in the improved BTS estimates relative to the other estimates is probably due to better coverage of the rapidly growing railroad shipments originating in the United States and terminating in Canada. Further,

FIGURE 5 Railroad Ton-Miles



Key: p = preliminary.

while Eno's ton-mile estimates for non-Class I railroads are based on financial survey data, the improved BTS estimates are based on actual ton-mile data and should be considered more reliable.

Finally, railroad ton-mile data may not include shipments originating in Mexico and terminating in Canada. Based on data from Transport Canada, it appears that these shipments account for less than one tenth of one percent of all U.S. railroad traffic.

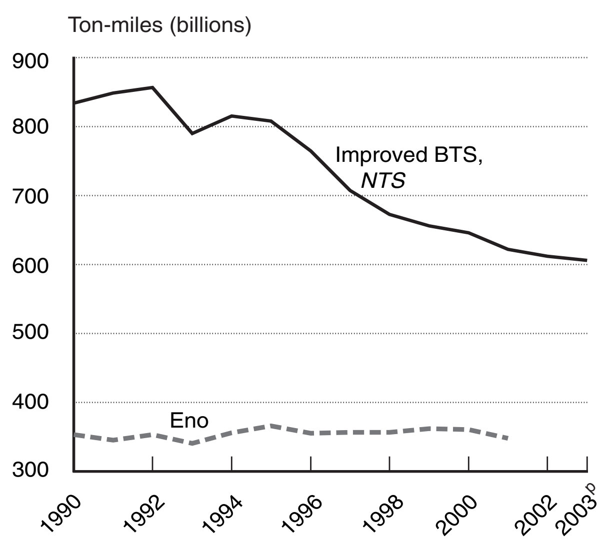
Water

Domestic waterborne ton-mile estimates are presented in figure 6 (also see appendix table A6). The current *NTS* estimates of annual water transportation ton-miles were taken from *Waterborne Commerce of the United States* (U.S. Army Corps of Engineers 2003). Data in this source are developed from lock data and individual trip reports that must be filed with the U.S. Coast Guard. Therefore, this source represents the entire population of all domestic water traffic, including inland waterways, coastwise, Great Lakes, and intraport traffic, along with traffic to and from Alaska, Hawaii, and Puerto Rico. The *NTS* and Eno estimates differ substantially, because *NTS* includes coastwise (domestic ocean) traffic and Eno does not. Thus, the current *NTS* data, which are proposed for use here, are more comprehensive than Eno's estimates.

Pipeline

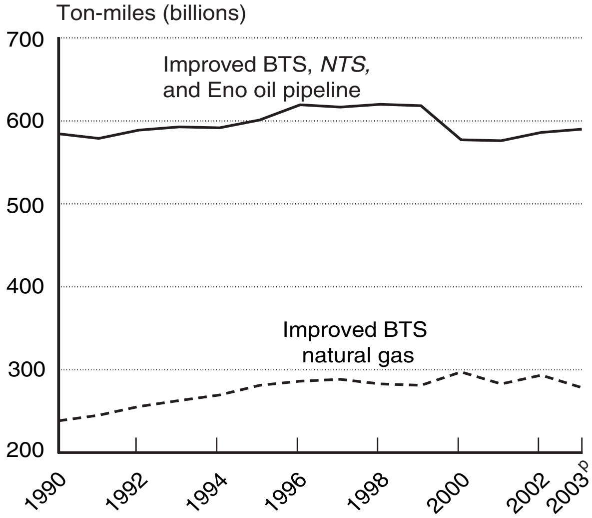
Figure 7 shows pipeline ton-miles (also see appendix tables A7(a) and A7(b)). Annual oil and oil products pipeline ton-miles were obtained from *Shifts in Petroleum Transportation* (Association of Oil Pipelines 2003). These data represent the entire population of crude petroleum and petroleum products carried in domestic transportation by both federally regulated and nonfederally regulated pipelines. Both *NTS* and Eno currently use these data, which we also propose for use here.

FIGURE 6 Waterborne Ton-Miles



Key: p = preliminary.

FIGURE 7 Pipeline Ton-Miles



Key: p = preliminary.

Natural gas pipeline ton-miles are also presented in figure 7. These new estimates are based on natural gas deliveries reported in the *Annual Energy Review* (USDOE 2003a). BTS first converted the gas deliveries, measured in cubic feet, to metric tons and then to tons using a standard conversion factor of 48,700 cubic feet per metric ton as reported in the *International Energy Annual* (USDOE 2001).

There are no data available on the length of haul for natural gas shipments, because natural gas is drawn from a common pipeline rather than shipped to a specific consignee. Origination and termination data indicate that natural gas has a distribution pattern similar to oil and oil products (USDOE 2003b and 2003c). The oil and natural gas pipeline networks are also very similar. Therefore, the length of haul for oil and oil products was applied to the tonnage of natural gas to estimate natural gas ton-miles in transmission lines.[3]

Natural gas ton-miles in distribution lines (i.e., local utilities) were estimated using 5% of transmission length of haul, which is approximately half the diameter of a major metropolitan area. Natural gas ton-miles in gathering lines (i.e., from well to processing plant) were estimated using the same length of haul as in distribution lines. The ton-miles for gathering, transmission, and distribution lines were then summed to provide an estimate of total natural gas ton-miles. Natural gas ton-miles, which have not to our knowledge been previously estimated, represent nearly as much traffic as that carried on the inland waterway system. These new estimates fill a substantial gap in the existing ton-mile data.

The natural gas pipeline data do not include gas used to repressurize gas fields or power the pipeline itself, because these uses do not represent gas carried in revenue transportation. The pipeline data also exclude coal slurry, ammonia, and other types of pipelines. There are only a few such pipelines, which tend to have either short haul or low volume, and appear to account for well under 1% of all pipeline ton-miles. BTS will investigate the recent decline in the oil pipeline ton-mile data and the resulting reduction in the estimate of natural gas ton-miles in order to improve the most recent estimates.

CONCLUSION

The improved ton-mile estimates for the air, truck, rail, water, and pipeline modes described in this paper are both more comprehensive and more reliable than well-known existing estimates. The improvements are most noticeable with respect to trucking and natural gas pipelines. Additional work will allow BTS to further improve these basic estimates of transportation activity.

BTS has already incorporated these improved estimates of domestic freight ton-miles into the *Transportation Statistics Annual Report* (USDOT BTS 2004, p. 213).[4] BTS plans to extend the improved estimates back to 1980 in the fall 2005 update to *National Transportation Statistics*.[5] Future research will be conducted to further extend the improved estimates back to 1960 where data are available.

REFERENCES

Association of Oil Pipelines. 2003. *Shifts in Petroleum Transportation.* Washington, DC. Table 1.

Eno Transportation Foundation. 2002. *Transportation in America.* Washington, DC.

[3] The 2000 to 2003 oil pipeline length of haul data are somewhat suspect; thus, the average length of haul from 1990 to 1999 was used in place of these data.

[4] Estimates presented in this paper reflect subsequent revisions in the source data.

[5] Available on the BTS website at www.bts.dot.gov.

Transport Canada. 1990–2002. *Transportation in Canada*. Ottawa, Ontario, Canada. Addendum, table A6-10.

U.S. Army Corps of Engineers. 2003. *Waterborne Commerce of the United States*. Washington, DC. Part V, Section 1, Table 1-4, Total Waterborne Commerce.

U.S. Army Corps of Engineers, Waterborne Commerce Statistics Center (WCSC). 2003. *U.S. Waterborne Container Traffic by Port*. Washington, DC.

U.S. Department of Energy (USDOE), Energy Information Administration. 2001. *International Energy Annual*. Washington, DC. Table C-1.

_____. 2003a. *Annual Energy Review*. Washington, DC. Table 6.5.

_____. 2003b. *Natural Gas Annual*. Washington, DC. Table 12.

_____. 2003c. *Petroleum Supply Annual, Volume 1*. Washington, DC. Table 33.

U.S. Department of Transportation (USDOT), Bureau of Transportation Statistics (BTS). 2000. *Transportation Statistics Annual Report*. Washington, DC.

_____. 2003. *National Transportation Statistics*. Washington, DC. Table 1-44.

_____. 2004. *Transportation Statistics Annual Report*. Washington, DC.

U.S. Department of Transportation (USDOT), Bureau of Transportation Statistics (BTS), Office of Airline Information. 1990–2003. *Air Carrier Traffic Statistics Monthly*. Washington, DC. Freight, Express, and Mail Revenue Ton-Miles table, p. 2, line 3.

U.S. Department of Transportation (USDOT), Bureau of Transportation Statistics and U.S. Department of Commerce (USDOC), Economics and Statistics Administration, U.S. Census Bureau. 1993. *Commodity Flow Survey*. Washington, DC.

______. 1997. *Commodity Flow Survey*. Washington, DC.

U.S. Department of Transportation (USDOT), Federal Highway Administration (FHWA). 1990–2003. *Highway Statistics*. Washington, DC. Table VM-1.

U.S. Department of Transportation (USDOT), Surface Transportation Board (STB). 1990–2003. *Carload Waybill Sample*. Washington, DC.

APPENDIX

TABLE A1 All Ton-Miles

Improved BTS Ton-Miles
(billions)

Mode	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003[p]
Air	10	10	11	12	12	13	14	14	14	15	16	13	14	15
Truck	854	874	896	936	996	1,042	1,071	1,119	1,149	1,186	1,203	1,224	1,255	1,264
Railroad	1,064	1,042	1,098	1,135	1,221	1,317	1,377	1,391	1,448	1,504	1,546	1,599	1,606	1,604
Water	834	848	857	790	815	808	765	707	673	656	646	622	612	606
Coastwise	479	502	502	448	458	440	408	350	315	293	284	275	264	279
Lakewise	61	55	56	56	58	60	58	62	62	57	58	51	54	48
Internal	292	290	298	284	298	306	297	294	295	305	303	295	293	278
Intraport	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Pipeline	822	824	844	856	860	882	905	904	902	899	874	859	879	868
Oil and oil products	584	579	589	593	591	601	619	617	620	618	577	576	586	590
Natural gas	238	245	255	263	269	281	286	288	283	281	297	283	293	278
TOTAL	**3,584**	**3,597**	**3,706**	**3,727**	**3,904**	**4,061**	**4,131**	**4,136**	**4,186**	**4,259**	**4,285**	**4,317**	**4,366**	**4,357**

Key: p = preliminary.

***NTS* Ton-Miles**
(billions)

Mode	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Air	9	9	10	11	12	13	13	14	14	14	15	13
Intercity truck	735	758	815	861	908	921	972	996	1,027	1,059	1,074	1,051
Class I railroad	1,034	1,039	1,067	1,109	1,201	1,306	1,356	1,349	1,377	1,433	1,466	1,495
Water	834	848	857	790	815	808	765	707	673	656	646	622
Coastwise	479	502	502	448	458	440	408	350	315	293	284	275
Lakewise	61	55	56	56	58	60	58	62	62	57	58	51
Internal	292	290	298	284	298	306	297	294	295	305	303	295
Intraport	1	1	1	1	1	1	1	1	1	1	1	1
Pipeline	584	579	589	593	591	601	619	617	620	618	577	576
Oil and oil products	584	579	589	593	591	601	619	617	620	618	577	576
Natural gas	—	—	—	—	—	—	—	—	—	—	—	—
TOTAL	**3,196**	**3,233**	**3,337**	**3,364**	**3,527**	**3,648**	**3,725**	**3,682**	**3,710**	**3,781**	**3,778**	**3,757**

(*continued on next page*)

TABLE A1 All Ton-Miles *(continued)*

Eno Ton-Miles
(billions)

Mode	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Air	10	10	11	12	12	13	14	14	14	15	16	15
Intercity truck	735	758	815	861	908	921	972	996	1,027	1,059	1,074	1,051
Railroad	1,091	1,100	1,138	1,183	1,275	1,375	1,426	1,421	1,442	1,499	1,534	1,558
Water	353	345	353	340	356	366	355	356	357	362	360	348
Rivers/canals	292	290	298	284	298	306	297	294	295	305	303	297
Great Lakes domestic	61	55	56	56	58	60	58	62	62	57	58	51
Pipeline	584	579	589	593	591	601	619	617	620	618	577	576
Oil and oil products	584	579	589	593	591	601	619	617	620	618	577	576
Natural gas	—	—	—	—	—	—	—	—	—	—	—	—
TOTAL	**2,774**	**2,792**	**2,906**	**2,989**	**3,142**	**3,276**	**3,386**	**3,404**	**3,459**	**3,552**	**3,561**	**3,548**

TABLE A2 Airline Freight, Express, and Mail Revenue Ton-Miles
(billions)

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003[p]
Air (improved BTS)[1,2]	10.420	9.960	10.990	11.540	12.030	12.720	13.760	13.900	14.140	14.500	15.810	13.288	13.837	15.096
Air (*NTS*)[3]	9.064	8.860	9.820	10.675	11.803	12.520	12.861	13.601	13.840	14.202	14.983	13.088		
Air (Eno)[4]	10.420	9.960	10.990	11.540	12.030	12.720	13.760	13.900	14.140	14.500	15.810	15.180		

[1] U.S. Department of Transportation, Bureau of Transportation Statistics, Office of Airline Information, *Air Carrier Traffic Statistics Monthly* (Washington, DC: 1990–2003), Freight, Express, and Mail Revenue Ton-Miles table, p. 2, line 3.

[2] Federal Aviation Administration, supplementary statistics.

[3] U.S. Department of Transportation, Bureau of Transportation Statistics, *National Transportation Statistics* (Washington, DC: 2003), table 1-44.

[4] Eno Transportation Foundation, *Transportation in America* (Washington, DC: 2002).

Key: p = preliminary.

TABLE A3 Development of Truck Ton-Mile Estimates

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003p
Oak Ridge National Lab estimate (billions)[1]				910				1,109						
Single-unit truck VMT (billions)[2]	52	53	54	57	61	63	64	67	68	70	71	72	76	78
Combination truck VMT (billions)[2]	94	97	100	103	109	115	119	125	128	132	135	137	139	138
Total truck VMT (billions)	146	150	153	160	170	178	183	191	196	203	206	209	215	216
Truck VMT index, 1993 base	0.915	0.935	0.959	1.000	1.065	1.114	1.144	1.197	1.228	1.268	1.285	1.307	1.342	1.350
Estimated truck traffic, 1993 base (billion ton-miles)	832	851	873	910	968	1,014	1,041	1,089	1,117	1,153	1,169	1,189	1,221	1,228
Truck VMT index, 1997 base	0.764	0.782	0.802	0.836	0.890	0.931	0.956	1.000	1.026	1.059	1.074	1.092	1.122	1.128
Estimated truck traffic, 1997 base (billion ton-miles)	848	867	889	927	987	1,033	1,060	1,109	1,138	1,175	1,191	1,212	1,244	1,251

[1] U.S. Department of Transportation, Bureau of Transportation Statistics, *Transportation Statistics Annual Report* (Washington, DC: 2000), p. 124.

[2] U.S. Department of Transportation, Federal Highway Administration, *Highway Statistics* (Washington, DC: 1990–2003), table VM-1.

Key: p = preliminary; VMT = vehicle-miles traveled.

TABLE A4 Truck Ton-Miles
(billions)

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003p
Intercity truck (*NTS,* Eno)[1, 2]	735	758	815	861	908	921	972	996	1,027	1,059	1,074	1,051		
All truck, 1997 base (improved BTS)	854	874	896	936	996	1,042	1,071	1,119	1,149	1,186	1,203	1,224	1,255	1,264

[1] U.S. Department of Transportation, Bureau of Transportation Statistics, *National Transportation Statistics* (Washington, DC: 2003), table 1-44.

[2] Eno Transportation Foundation, *Transportation in America* (Washington, DC: 2002), p. 42.

Key: p = preliminary.

TABLE A5 Railroad Ton-Miles
(billions, unless otherwise noted)

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003[p]
Waybill ton-miles, U.S. terminations[1]	1,055	1,033	1,089	1,124	1,208	1,303	1,362	1,374	1,433	1,490	1,530	1,581	1,587	
U.S.-Canada terminations (million metric tons)[2]	11.479	10.398	11.362	12.482	14.502	15.391	16.474	18.403	17.099	15.175	17.624	19.813	19.606	
Conversion to short tons (millions)	12.654	11.462	12.525	13.759	15.986	16.966	18.160	20.286	18.849	16.728	19.427	21.840	21.612	
Average U.S. length of haul (miles)[3]	726	751	763	794	817	843	842	851	835	835	843	859	853	
Ton-miles, U.S.-Canada terminations	9.183	8.611	9.550	10.927	13.057	14.295	15.285	17.261	15.740	13.966	16.383	18.750	18.435	
Waybill ton-miles, all traffic														1,621
Waybill tons, all traffic (millions)														2,078
Waybill tons, traffic within U.S. (millions)														1,965
Mileage in Canada														150
Canadian ton-miles														17
Railroad ton-miles (improved BTS)	1,064	1,042	1,098	1,135	1,221	1,317	1,377	1,391	1,448	1,504	1,546	1,599	1,606	1,604
Class I ton-miles (*NTS*)[4]	1,034	1,039	1,067	1,109	1,201	1,306	1,356	1,349	1,377	1,433	1,466	1,495		
Railroad ton-miles (Eno)[5]	1,091	1,100	1,138	1,183	1,275	1,375	1,426	1,421	1,442	1,499	1,534	1,558		

[1] U.S. Department of Transportation, Surface Transportation Board, *Carload Waybill Sample* (Washington, DC: 1990–2003).
[2] Transport Canada, *Transportation in Canada* (Ottawa, Ontario, Canada: 1990–2002), Addendum, table A6-10.
[3] Association of American Railroads, *Railroad Facts* (Washington, DC: Various years), p. 36.
[4] U.S. Department of Transportation, Bureau of Transportation Statistics, *National Transportation Statistics* (Washington, DC: 2003), table 1-44.
[5] Eno Transportation Foundation, *Transportation in America* (Washington, DC: 2002), p. 42.

Key: p = preliminary.

TABLE A6 Waterborne Ton-Miles
(billions)

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003^{P}
Water (*NTS,* improved BTS)[1]	834	848	857	790	815	808	765	707	673	656	646	622	612	606
Coastwise	479	502	502	448	458	440	408	350	315	293	284	275	264	279
Lakewise	61	55	56	56	58	60	58	62	62	57	58	51	54	48
Internal	292	290	298	284	298	306	297	294	295	305	303	295	293	278
Intraport	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Water (Eno)[2]	353	345	353	340	356	366	355	356	357	362	360	348		
Rivers/canals	292	290	298	284	298	306	297	294	295	305	303	297		
Great Lakes, domestic	61	55	56	56	58	60	58	62	62	57	58	51		

[1] U.S. Army Corps of Engineers, *Waterborne Commerce of the United States* (Washington, DC: 2003), Part V, Section 1, Table 1-4, Total Waterborne Commerce.

[2] Eno Transportation Foundation, *Transportation in America* (Washington, DC: 2002), p. 42.

Key: p = preliminary.

TABLE A7(a) **Oil and Oil Products Pipeline Ton-Miles**
(billions)

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003[p]
Oil pipeline (*NTS,* improved BTS, Eno)[1, 2]	584	579	589	593	591	601	619	617	620	618	577	576	586	590

TABLE A7(b) **Natural Gas Pipeline Ton-Miles**
(billions, unless otherwise noted)

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003[p]
Total consumption, cubic feet (trillion)[3]	19.174	19.562	20.228	20.790	21.247	22.207	22.609	22.737	22.246	22.405	23.333	22.239	23.018	21.894
Lease and plant fuel, cubic feet (trillion)[3]	1.236	1.129	1.171	1.172	1.124	1.220	1.250	1.203	1.173	1.079	1.151	1.119	1.114	1.123
Pipeline fuel, cubic feet (trillion)[3]	0.660	0.601	0.588	0.624	0.685	0.700	0.711	0.751	0.635	0.645	0.642	0.625	0.667	0.635
Gas to consumers, cubic feet (trillion)	17.278	17.832	18.469	18.994	19.438	20.287	20.648	20.783	20.438	20.681	21.540	20.495	21.237	20.136
Gas to consumers, tons[4]	0.391	0.404	0.418	0.430	0.440	0.459	0.467	0.471	0.462	0.468	0.488	0.464	0.481	0.456
Oil pipeline ton-miles[1]	584	579	589	593	591	601	619	617	620	618	577	576	586	590
Oil pipeline tons[2]	1.057	1.048	1.061	1.067	1.064	1.081	1.114	1.108	1.116	1.131	—	—	—	—
Length of haul (miles)	553	552	555	556	556	556	556	556	555	546	554	554	554	554
Gas transmission pipeline ton-miles	216	223	232	239	245	255	260	262	257	256	270	257	267	253
Gas gathering pipeline ton-miles (estimated)[5]	11	11	12	12	12	13	13	13	13	13	14	13	13	13
Gas distribution pipeline ton-miles (estimated)[5]	11	11	12	12	12	13	13	13	13	13	14	13	13	13
Total gas pipeline ton-miles (improved BTS)	238	245	255	263	269	281	286	288	283	281	297	283	293	278

[1] Association of Oil Pipelines, *Shifts in Petroleum Transportation* (Washington, DC: 2003), table 1.
[2] Eno Transportation Foundation, *Transportation in America* (Washington, DC: 2002), p. 42.
[3] U.S. Department of Energy, Energy Information Administration, *Annual Energy Review* (Washington, DC: 2001), table 6.5.
[4] Conversion factor from U.S. Department of Energy, Energy Information Administration, *International Energy Annual* (Washington, DC: 2001), table C-1.
[5] Estimated at 5% of transmission ton-miles.

Key: p = preliminary.

Spatial and Temporal Transferability of Trip Generation Demand Models in Israel

ADRIAN V. COTRUS[1]
JOSEPH N. PRASHKER[2]
YORAM SHIFTAN[2,*]

[1] Department of Civil Engineering
Technion, Israel Institute of Technology
Haifa 32000, Israel

[2] Transportation Research Institute
Technion, Israel Institute of Technology
Haifa 32000, Israel

ABSTRACT

This research investigates the transferability of person-level disaggregate trip generation models (TGMs) in time and space using two model specifications: multinomial linear regression and Tobit. The models are estimated for the Tel Aviv and Haifa metropolitan areas based on data from the 1984 and 1996/97 Israeli National Travel Habits Surveys. The paper emphasizes that Tobit models perform better than regression or discrete choice models in estimating nontravelers. Furthermore, the paper notes that variables and file structures in household surveys need to be consistent. Results of the study show that the estimated regression and Tobit disaggregate person-level TGMs are statistically different in space and in time. In spite of the transferred forecasts, the aggregate forecasts were also similar.

INTRODUCTION

Trip generation models (TGMs) are used as a first step in classical four-step travel demand modeling and, therefore, any over- or underprediction of trip generation rates can cause errors throughout the entire transportation planning process. Inappropriate decisionmaking due to these types of errors can

Email addresses:
* Corresponding author—shiftan@tx.technion.ac.il
A. Cotrus—cotrus@walla.co.il
J. Prashker—prashker@netvision.net.il

KEYWORDS: Trip generation, transferability, multinomial linear regression, Tobit model.

account for premature investments in infrastructure in the case of overprediction and loss of labor hours, pollution, and low levels of service in the case of underprediction.

TGMs are usually estimated based on periodic surveys of the travel habits of individuals or households. They are expensive and difficult to perform and are not conducted often. For our research, we used the last Israeli National Travel Habits Surveys collected in 1984 and 1996/97. Transportation planners use models previously estimated and sometimes in different contexts. The planners can perform forecasts for the same areas and, if justifiable, transfer the models to other areas. Hence, it is important to know whether these models can be transferred in time and in space.

Recently, researchers and planning agencies began to implement tour-based activity modeling systems rather than trip-based modeling systems.[1] The advantage in using an activity modeling approach is the ability to model each individual's tours. However, in this paper, we were only able to investigate the stability of individual predictions of trips. This research presents the characteristics of trip generation in Israel and tries to answer the question of whether linear regression and Tobit TGMs can be transferred in time and space, given the dynamic changes in metropolitan areas and socioeconomic characteristics. The TGMs estimated and analyzed in this research include only vehicular trips.

We estimated the models for the geographically diverse metropolitan areas of Haifa and Tel Aviv in Israel and tested them for transferability in space and in time. The topography of Haifa is hilly, with the core of the city poorly connected to the rest of the metropolitan area. Tel Aviv lies on level topography, with a well connected road network. The metropolitan areas also differ structurally: Tel Aviv is interconnected like a spider web, including several minor cores with high-density population and employment concentrations. On the other hand, Haifa's less connected road network and rolling terrain give it a lower level of accessibility. When comparing the areas by land use, the Tel Aviv metropolitan area consists of neighborhoods that combine residential shopping and personal business areas. Haifa, on the other hand, contains highly separated areas with each area consisting of a uniform land use.

Given the difference in accessibility and land use, the calibrated models were restricted to demographic and socioeconomic variables. The average number of daily trips per person was higher in the Haifa metropolitan area (2.14 in 1984, and 2.03 in 1996/97) than in the Tel Aviv metropolitan area (1.83 in 1984, and 1.91 in 1996/97). The difference in the average trip rates may be explained based on the variation in land uses. The lack of mixed land uses in Haifa may encourage the generation of more trips. Furthermore, Haifa's hilly topography may encourage more vehicular trips than in Tel Aviv, where shorter trips are probably done on foot. The comparison between these metropolitan areas and the calibrated models is possible due to the similarity of the distribution of most demographic and socioeconomic variables. (See table 1 for a partial presentation of the comparison.) A more detailed comparison of Haifa and Tel Aviv characteristics is presented in Cotrus (2001).

LITERATURE REVIEW

Over the last few decades, several papers have discussed the transferability of trip generation models. The debate among researchers, in general, focused not only on the transferability of models in space and time but also on the model specification and level of aggregation. The aggregation levels are usually defined as area (zonal), household, and person. Estimating the models (see, e.g., Ortuzar and Willumsen 1994) at more disaggregate levels improves the transferability of TGM.

Atherton and Ben-Akiva (1976) emphasized that disaggregated models tend to maintain the variance and behavioral context of the response variable and, therefore, are expected to give better estimates when transferred. Downes and Gyenes (1976) pointed out that when the explanatory power of the model is of interest rather than the aggregate forecasts, the disaggregate level should be selected. Wilmot (1995) indicated that disaggregate models are preferred because of their independence from

[1] *Tours* refer to a sequence of trips usually starting and ending at home. *Trips* refer to just the movement between an origin and destination.

TABLE 1 Demographic Characteristics by Metropolitan Area and Year

Household and personal characteristics	Haifa			Tel Aviv		
	1996			1996		
	In terms of 1996*	In terms of 1984*	1984	In terms of 1996*	In terms of 1984*	1984
HOUSEHOLDS						
Total (thousands)	245	160	130	766	678	522
Persons in household (%)						
1	15.8	17.3	20.7	16.7	17.3	18.8
2	25.6	29.9	29.3	24.3	25.4	26.0
3–4	36.0	37.1	32.6	35.5	35.2	34.1
5–6	18.4	14.4	14.8	20.3	19.4	18.2
7+	4.3	1.2	2.6	3.1	2.7	3.0
Car availability						
0	38.4	36.7	51.6	34.3	33.8	51.1
1	42.0	40.7	40.3	40.0	39.9	39.1
2	11.4	12.3	7.4	17.5	17.5	9.4
3+	1.6	1.9	—	2.6	2.8	—
Not known	6.6	8.4	—	5.6	6.0	—
Households with vehicles (%)	55.0	54.9	48.4	60.1	60.2	48.9
TOTAL POPULATION						
Total (thousands)	803	470	380	2,477	2,145	1,632
Age (%)						
0–7	13.5	11.0	14.8	13.7	13.0	14.8
8–17	18.5	15.8	15.7	17.3	16.8	17.6
18–29	18.0	17.7	16.3	18.2	18.3	17.7
30–64	38.2	40.1	39.4	39.5	39.8	37.9
65+	11.8	15.4	13.8	11.4	12.0	12.0
POPULATION AGED 8+						
Total (thousands)	695	419	324	2,138	1,866	1,390
Sex (%)						
Men	48.5	48.1	47.9	48.3	48.2	49.7
Women	51.5	51.9	52.1	51.7	51.8	50.3
Years of schooling (%)						
0–8	31.3	24.2	34.7	26.5	25.8	39.2
9–12	37.9	37.8	40.7	41.0	41.0	40.1
13+	25.1	30.1	24.6	27.4	27.6	20.7
Not known	5.7	7.9	—	5.0	5.5	—
POPULATION AGED 15+						
Total (thousands)	589	367	281	1,833	1,609	1,177
Employed (%)						
Total	47.3	46.8	49.1	51.3	51.2	50.3
Men	57.8	55.1	61	54.1	53.5	61.5
Women	42.2	44.9	39	45.9	46.5	38.5
POPULATION AGED 17+						
Total (thousands)	559	351	270	1,752	1,540	1,127
Those having a driver's license (%)	53.0	54.2	46.6	59.8	60.0	49.8
Men (%)	63.0	59.7	64.8	58.6	57.9	66.0
Women (%)	37.0	40.3	35.2	41.4	42.1	34.0

* In terms of 1984 refers to the geographic definition of the metropolitan area in that year. In term of 1996 refers to the geographic definition of the metropolitan area that year.

zonal definitions. In Supernak et al. (1983) and Supernak (1987), the person level was preferred for TGM because of the identity of the response factor (trip) and the generative (the person). One advantage of disaggregate person-level models is the reduced amount of data required for model estimation. (For more details, see Fleet and Robertson 1968; and Ortuzar and Willumsen 1994.) Other types of model specification techniques include cross-classification, regression, logit-based models, artificial neural networks, fuzzy logic, and simulations.

A number of studies found spatial transferability of models satisfactory (Wilmot 1995; Atherton and Ben-Akiva 1976; Supernak 1982, 1984; Duah and Hall 1997; Walker and Olanipekun 1989; Rose and Koppelman 1984; Caldwell and Demetsky 1980; and Kannel and Heathington 1973). On the other hand, Smith and Cleveland (1976) and Daor (1981) found spatial transferability unsatisfactory. We should emphasize that Smith and Cleveland pointed out that although the explanatory variables are distinctive, their effects vary in space. A number of researchers found the transferability of models in time (i.e., their temporal stability) satisfactory (Downes and Gyenes 1976; Yunker 1976; Walker and Peng 1991; Kannel and Heathington 1973; and Karasmaa and Pursula 1997). Unsatisfactory results, however, were obtained in other studies (Doubleday 1977; Smith and Cleveland 1976; and Copley and Lowe 1981).

While several international studies explored model transferability in time and space, in Israel the transferability of discrete mode choice models has been the main focus (Prashker 1982; Silman 1981). This study deals with the investigation of trip generation characteristics but also provides local estimates of TGMs and their validation for transferability in time and space. The study also explores the implementation of Tobit models in TGM.

We often approach trip generation from an economic viewpoint, where trips are defined as the product and the person/household as the customer. The strongest argument to model trips on a disaggregate level is that any zonal outcome is based on the aggregation of several customers, ignoring the heterogeneity among them. The explanatory variables for the power of consumption of each person/household can be found in several categories including demographic, geographic, and economic.

As discussed above, several approaches exist to model trip generation, including regression-based models such as multiple linear regression (Wilmot 1995) and cross-classification (Walker and Olanipekun 1989); discrete choice models such as probit, logit, and ordered probit (Zhao 2000); simulations such as Smash, Amos, and the Starchild System; fuzzy logic models; and artificial neural networks (Huisken 2000). Clearly, the issue of trip generation can be approached from several directions and tested for transferability in time and space. Therefore, researchers will choose the modeling approach based on the size of the database at hand, the nature and structure of the variables, the aggregation level desired, as well as other considerations. The main problem with using a regression model is the treatment of trip rates as continuous rather than discrete variables. Discrete choice models and spatially ordered response models may better account for the behavioral process of trip generation. However, due to practical reasons, in most models to date, the dependent variable is treated as a continuous variable. For this reason, we perform our analysis on such models.

METHODOLOGY

In this research, we first estimated regression models for each metropolitan area for each year, taking into account the inconsistency of the household surveys (table 2). We then estimated Tobit TGMs based on the same variables and tested whether these models are suitable for trip generation estimation and for transferability. The regression model form is presented in equation 1.

$$\begin{aligned} y_i &= \alpha + \beta_1 \bullet x_{1,i} + \beta_2 \bullet x_{2,i} ... + \beta_k \bullet x_{k,i} + \zeta_i \\ \forall i &= 1,...,n \\ \forall k &= 1, 2,...,k \end{aligned} \quad (1)$$

where

y_i = trip rate generated by individual i,

$x_{k,i}$ = explanatory variable k for individual i,

n = the number of observations,

k = number of explanatory variables, and

ζ_i = error term of the ith observation.

TABLE 2 Differences Among the 1972, 1984, and 1996/97 Traveling Habits Surveys

No.	Topic	1972 survey	1984 survey	1996/97 survey
1	Trip modes surveyed	Motorized trips only, although data on travel on foot were also gathered	Motorized and nonmotorized trips	Only motorized trips
2	Survey population	Permanent residents of Israel	Permanent residents of Israel	Permanent residents of Israel, including Jews domiciled in Judea, Samaria, and Gaza.
		Israeli residents abroad for less than 1 year	Israeli residents abroad for less than 2 months	Israeli residents abroad for less than 1 year
		Potential immigrants: persons who were eligible for immigrant cards and wished to stay in Israel for more than 3 months	Immigrants and potential immigrants who arrived in Israel before June 1983	Immigrants who reached Israel before the survey date
		Tourists, volunteers, and temporary residents in Israel for more than 1 year	Same as in 1972	Tourists, volunteers, and temporary residents in Israel for more than 3 months
		Excluding diplomats and UN personnel	Same as in 1972	Same as in 1972
3	Age	5+	8+	8+
4	Tenants of institutions	Including institutions where 5 or more persons spend the night	Institutions not included	Including students dormitories, immigrant-absorption centers, and sheltered housing
5	Gross sample size	56,000 households	5,000 households	17,700 households
6	Geographic spread and locality type	Countrywide (excluding Judea, Samaria, and Gaza)	Two expanded conurbations and Jerusalem and its surroundings	Countrywide
		Localities with populations of 10,000+	Not including Arab localities in Judea District	Not including Druze localities on the Golan Heights, institutional local cities, and areas outside settled area, including Bedouin tribes in the Negev
		Small localities directly related to the conurbation	Not including kibbutzim and villages	Including kibbutzim and villages
7	Investigation period	November 1972–June 1973; not including summer months and holidays	January 1984–May 1985; 2 month interruption: July–August 1984	March 1996–February 1997; supplementary period: March–August 1997.
8	Number of investigation days	One 24-hour day, starting at 14:00 on day preceding enumerator's visit and ending at time of visit	1.5 24-hour days, starting on day preceding enumerator's visit and ending at time of visit	3–4 investigation days, starting at beginning of day preceding enumerator's visit and ending at end of day after enumerator's visit
9	Investigation days	Sunday–Thursday; the entire period from Thursday evening to Sunday noon was not investigated; holiday eves were not investigated; Sundays were investigated at the previous day level, and Thursday at the current day level	Same as in 1972	All days of the week; holidays were investigated; Thursdays were not investigated at the previous day level; Sundays were not investigated at the next-day level; and Friday–Saturday were not investigated at the current-day level

(continues on next page)

TABLE 2 Differences Among the 1972, 1984, and 1996/97 Traveling Habits Surveys *(continued)*

No.	Topic	1972 survey	1984 survey	1996/97 survey
10	Investigation variables	Demographic and economic variables of household and individual	Same as in 1972	Same as in 1972
		Travel-related variables: origin, destination, hour, etc.	Same as in 1972	Same as in 1972
		Variables related to vehicles available to household: number of vehicles, type of vehicles, etc.	Same as in 1972	Same as in 1972, plus variables related to parking
11	Investigation method	Questionnaire filled in by enumerator	Same as in 1972	Questionnaire filled in by enumerator and diary left with respondents to fill in
12	Geographic regions	Tel Aviv and Haifa conurbations	Tel Aviv and Haifa conurbations and Jerusalem and its surroundings	Division into geographic regions performed only at survey data analysis phase: Tel Aviv metropolitan area, Haifa metropolitan area, planned metropolitan area of Beer Sheva, and Jerusalem and its surroundings
		Tel Aviv conurbation: 27 localities	Tel Aviv conurbation: 41 localities (4 more were added during the survey, for a total of 45)	Tel Aviv metropolitan area: 259 localities
		Haifa conurbation: 8 localities	Haifa conurbation: 14 localities; Jerusalem and its surroundings: 12 localities	Haifa metropolitan area: 101 localities
				Planned metropolitan area of Beer Sheva: 130 localities
13	Trips	Several actions—switching vehicles, stopping and exiting vehicle, switching drivers, and changing travel routes—were defined as having created two separate trips	A trip was defined by its main purpose even if there were stops on the way	Same as in 1972
			Reaching the destination by two means of transport or along two routes was considered one trip	
14	Means of transport	One means of transport per trip	Possibility of more than one means of transport per trip	Same as in 1972

Source: Israeli Ministry of Foreign Affairs, Central Bureau of Statistics, *1996/97 Household Survey: Description of Survey* (English in origin).

Hald (1949) first presented the model that, in its final form, is called the Tobit model (1958). Tobit models differentiate from regression models by the incorporation of truncated or censored dependent variables. Tobit analysis assumes that the dependent variable has a number of its values clustered at a limiting value, usually zero. The Tobit model can be presented as a discrete/continuous model that first makes a discrete choice of passing the threshold and second, if passed, a continuous choice regarding the value above the threshold. This approach is appropriate for trip generation, as an individual must decide whether to make any trips and, if so, how many trips to make.

Tobit analysis uses all observations when estimating the regression line, including those at the limit (no trips) and those above the limit (those who chose to travel). As shown by McDonald and Moffitt (1980), Tobit analysis can be used to determine the changes in the value of the dependent variable if it is above the limit, as well as changes in the probability of being above the limit. Since the surveys include observations at the limit (i.e., persons that are not traveling), it was also interesting to find out how well the Tobit model can predict persons doing no travel at all. The Tobit model form is presented in equation 2:

$$\begin{aligned} y_i &= X_i \bullet \beta + \zeta_i \qquad && if \quad X_i \bullet \beta + \zeta_i > 0 \\ y_i &= 0 && if \quad X_i \bullet \beta + \zeta_i \le 0 \\ \forall i &= 1, 2, 3, \ldots, (N-1), N && \end{aligned} \tag{2}$$

where
N = number of observations,
y_i = trip rate generated by observation i,
X_i = vector of independent variables,
β = vector of coefficients, and
ζ = independently distributed error term $\sim \left(0, \sigma^2\right)$.

Because Tobit models have not been used previously in the context of trip generation, this research investigates their suitability for that purpose. We also compared the predictions obtained using regression models with those produced using the analogous Tobit model. The best specification of the regression model is not necessarily the best specification of the Tobit model. However, in order to allow for basic comparisons of the model parameters, we estimated the regression models first; then, after the determination of the final variables in the model, we estimated Tobit models with the same variables.

All models were estimated at the disaggregate person level. At the person level of modeling, we maintained the heterogeneity among observations and kept a good identity between the consumer of the product (the person) and the outcome (number of daily trips taken by the person). As discussed above, disaggregate models tend to show better transfer results than aggregate models and also incorporate the power to understand and control the production of trips. The models were estimated for a 24-hour period[2] and tested for transferability in space and in time. Figure 1 presents the sequence of the analysis.

Statistical tests were conducted to determine the spatial and temporal stability of the estimated models by assessing the transferability of the coefficients from one area to another, and for each metropolitan area between the two survey years. Transferability was also tested by comparing the overall aggregate prediction obtained by the transferred model with the local model. Furthermore, we analyzed the ability of Tobit models to represent and evaluate nontravelers, that is, people who do not generate trips based on the given survey data.

Data Sources and Descriptions

The Israeli Central Bureau of Statistics (CBS) conducted some limited scope[3] Traveling Habits Surveys in the 1960s. Comprehensive National Traveling Habits Surveys have been conducted by CBS every 12 years since 1972. Because the 1972 survey is not available on magnetic media, it was not possible to do a computer-based statistical analysis. Therefore, we based this research on the 1984 and 1996/97 household surveys.

The main problems we encountered in doing this research were related to the inconsistency in the investigated variables, the structure of the surveys, the definition of variables, the period of investigation, the geographic deployment, and the database structure (see table 2 again). The 1984 and 1996/97 household surveys differ in several ways: the geographic deployment (number and size of jurisdictions in the survey), the size of the survey (number of households), the definitions of the investigation period, and the variables that were excluded from the surveys. For example, income is included in the 1984 survey but is omitted from the 1996/97 survey.

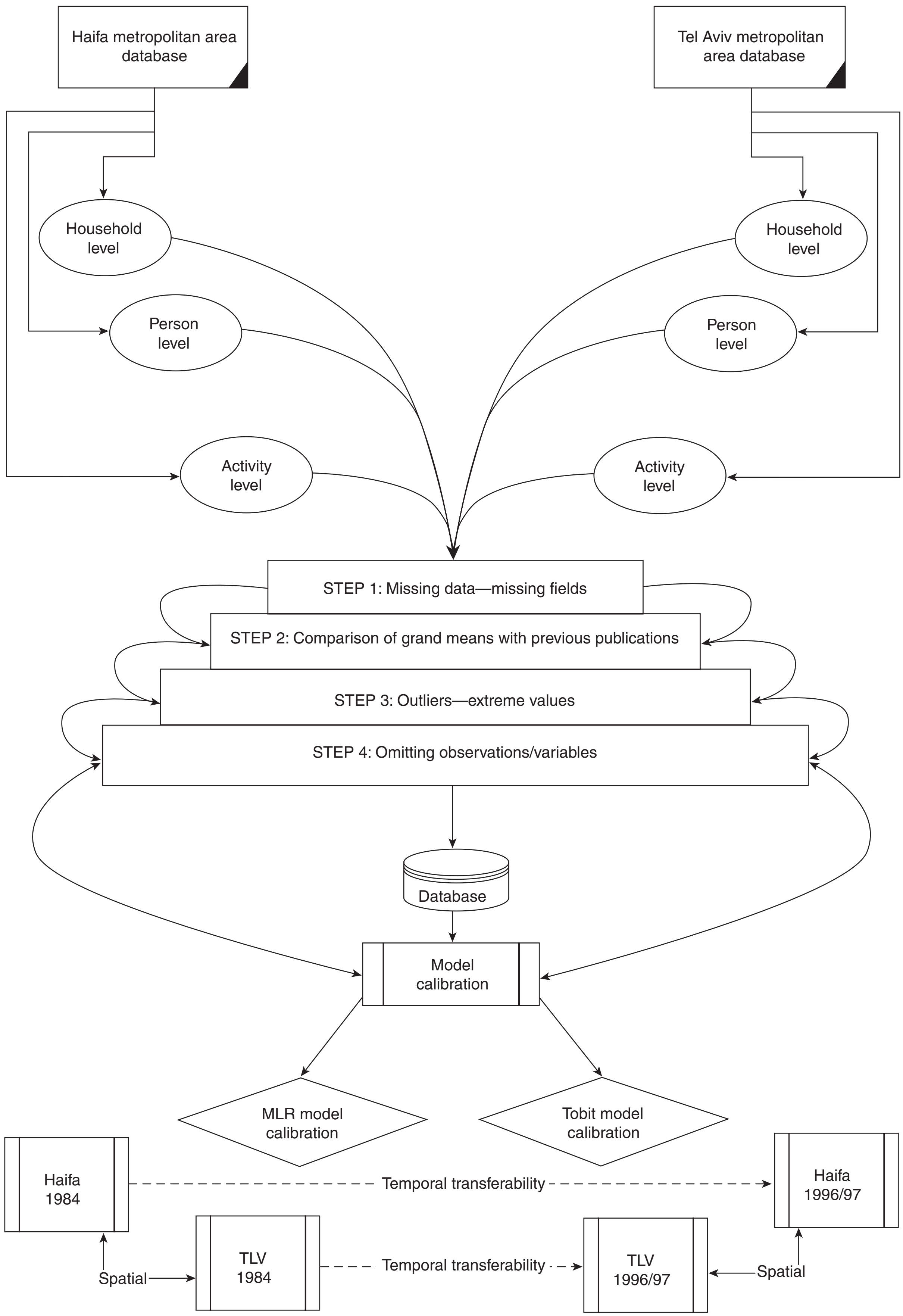
Despite definition and database differences in the two surveys (1984 was an activity survey and 1996/97 was a trip survey), we were able to bring the variables in the models to a common basis. In particular, the 1984 survey included bicycle and walking trips among the means to accomplish the activities, while the later survey excluded them. To resolve this difference, we excluded walking and biking trips from the 1984 database; only motorized trips were considered for each person.

The 1984 survey files included data for 5,420 persons in the Tel Aviv metropolitan area and 4,056 persons in the Haifa metropolitan area. The final files used for model calibration after sieving

[2] The 1984 household survey contained 1.5 days of data for each person; the 1996/97 survey contained 3 to 4 days of data for each person.

[3] These surveys were restricted to work-related activities only.

FIGURE 1 Research Sequence of Stages



incomplete and anomalous observations totaled 4,385 and 3,258 persons, respectively. The 1996/97 files included data for 20,436 persons in the Tel Aviv metropolitan area and 6,417 persons in the Haifa metropolitan area. The final files used for model calibration totaled 15,729 and 5,041 persons, respectively.

RESULTS

The selected trip generation models included six categorical variables: age, car availability, possession of a driver's license, employment, education, and status in the household. Data fell into five age categories: 8–13, 14–17, 18–29, 30–64, and 65 and over. Ortuzar and Willumsen (1994) found that life cycle variables were an important factor for explaining trip generation. Different trip rates can be expected for households and people at various stages of life. Furthermore, age should correlate with employment, having a driver's license, and marital status. Car availability included three categories: 0, 1, and ≥ 2 cars in the household. Clearly, households with more cars available will generate more trips. The driver's license category has only one variable: whether the person has a license (including motorcycle) or not.

The employment variable indicates whether the person was employed or not. Employed persons were expected to generate more trips, because they usually make at least two trips: to and from work.

Household status refers to whether the person defines himself or herself as the head of the household. This variable indicates the responsibility and availability of household resources as an incentive for consumption of trips.

Finally, four education categories were defined based on the number of years of study (0, 1–8, 9–12, and 13 or more). The literature shows good correlation between education and income. In the absence of a pure economic indicator, education is used also as a proxy for income. Respondents with higher education (hence higher income) were expected to generate more trips. All variables were found significant and the coefficients corresponded with our expectations.

Table 3 shows the estimation results for the regression models for 1984 for both metropolitan areas. As can be seen from the table, all coefficients were found to be significant at the 95% level. The number of observations remaining in the estimation process resulted from the limited scope of this survey and the elimination of incomplete observations in the original database.

Estimation results for these models show that all variables affected trip generation as expected. The education coefficients show that people with higher education generated more trips. This can be explained not only by the assumption of the relationship between education and income but also by assuming that a person with higher education is more likely to pursue culture and perhaps leisure activities. Also, as expected, persons with driver's licenses and employed persons tended to generate more trips than the equivalent nonworking and/or nondriving persons. Heads of household tended to generate more trips, as assumed, because of the responsibility and availability of resources. The coefficients of the age categories indicate that persons aged 14 to 17 travel more than people with similar characteristics of other age groups, probably because they are young and active and have less household or work responsibilities.

The overall R^2 of the 1984 models was 0.33 for the Tel Aviv model and 0.34 for the equivalent Haifa model. These R^2 values are modest but not anomalous for trip generation modeling. They indicate that a substantial portion of trip generation can be explained by nonhousehold factors, such as relative location of residence, employment, and other parameters. Statistical Z tests (assuming known variances, normal distributions, and independence of populations) for the transferability of the coefficients (without updating) show that, at a 95% level of confidence, the coefficients differ, except for the coefficient defining the head of household. To verify the results we also conducted Chow tests for the transferability of the models. The calculated statistic was 7.86 in comparison with the critical 1.72 F-value (at a 95% level of confidence), and yield the same conclusion that the 1984 models are not transferable in space.

Table 4 shows the results of transferring the models in space by showing the predictions from the estimated models and the 1984 database, each applied for both metropolitan areas. The table

TABLE 3 Regression Models by Metropolitan Areas: 1984

	Haifa			Tel Aviv		
Variable	**Coefficient**	***t* statistic**	**Standard error**	**Coefficient**	***t* statistic**	**Standard error**
Regression intercept	4.10	30.26	0.13	3.51	32.47	0.11
Age						
8–13	0.60	4.32	0.14	0.35	3.22	0.11
14–17	0.72	4.98	0.14	0.80	6.75	0.12
18–29	0.39	3.19	0.12	0.53	5.33	0.09
30–64	0.24	2.33	0.10	0.42	4.89	0.08
65+	0.00	—	—	0.00	—	—
Number of cars per household						
0	–0.83	–7.74	0.11	–0.80	–10.25	0.08
1	–0.39	–4.22	0.09	–0.46	–6.71	0.07
2+	0.00	—	—	0.00	—	—
Driver's license						
No license	–0.99	–11.23	0.09	–0.67	–9.80	0.07
License	0.00	—	—	0	—	—
Employed						
Not employed	–1.06	–13.17	0.08	–1.00	–15.58	0.06
Employed	0.00	—	—	0.00	—	—
Household status						
Head (0 = no)	–0.24	–3.13	0.07	–0.24	–3.82	0.06
Head (1 = yes)	0.00	—	—	0.00	—	—
Education						
0 years	–0.77	–3.98	0.19	–0.87	–6.50	0.13
1–8 years	–0.69	–6.80	0.10	–0.68	–8.57	0.08
9–12 years	–0.24	–2.95	0.08	–0.36	–5.50	0.07
13+ years	0.00	—	—	0.00	—	—
Overall R^2		0.34			0.33	
Number of observations used		3,243			4,356	

Notes: Chow test for 1984 model transferability:
Ess_1 = error sum of squares of Haifa set
Ess_2 = error sum of squares of Tel Aviv set
Ess_3 = error sum of squares of combined set
K = number of parameters in the model including the constant
N_i = number of observations in model i

$$H_0\colon \beta_i^1 = \beta_i^2 \quad \forall i$$

$$H_1 = else \qquad \alpha = 0.05$$

$$\frac{Ess_3 - (Ess_1 + Ess_2)}{K} \bullet \frac{N_1 + N_2 - 2 \bullet K}{Ess_1 + Ess_2} = \frac{20124 - (9351 + 10505)}{13} \bullet \frac{3243 + 4356 - 2 \bullet 13}{9351 + 10505} = 7.86$$

$$F_{(k, N_1 + N_2 - 2 \bullet K)} = F_{(13, \infty)} = 1.72 \Rightarrow H_1$$

TABLE 4 Aggregate Estimation Results Using Regression Models: 1984

	Haifa data			Tel Aviv data		
	Tel Aviv model	Haifa model	Actual	Haifa model	Tel Aviv model	Actual
Average trips per person per day	1.84	2.18	2.18	2.19	1.84	1.84
Negative estimations	24	0	—	0	32	—
Minimum trip rate	–0.07	0.20	0	0.20	–0.07	0
Maximum trip rate	4.04	4.49	11	4.49	4.04	10
Total	5,973	7,060	7,059	9,525	8,007	8,019
Difference between model estimation and actual values	**–15.37**	**0.00**	**—**	**18.79**	**–0.15**	**—**
Standard deviation	1.10	1.21	2.08	1.18	1.07	1.90

shows that the Haifa model overpredicts the actual trip rate when applied to the Tel Aviv database, in comparison with the Tel Aviv model applied to the Haifa database. This was expected, as the Haifa trip rate is higher than that for Tel Aviv.

Table 5 presents the estimation results for the 1984 Tobit models using the 1984 household survey data. When we transferred the estimated Tobit models in space and used them to predict the average trip rate in the other city, we found that the Haifa Tobit model overestimated the trips in Tel Aviv by 21.9% (table 6). When we used the estimated Tel Aviv Tobit model to predict the average trip rate for Haifa, we found that it underestimated the total trips by 27.5%. However, transferability *t*-tests at a 95% level of confidence showed that most of the coefficients are not significantly different in space, except for the license variable and the 8–13 and 30–64 age category variables. On the other hand, χ^2 tests at the 95% level of confidence strengthen the alternative hypothesis, that the models vary in space $\left(\chi^2_{13,\,0.95} = 22.36 < 91.198\right)$.

An important issue was to find out whether the Tobit model could explain and capture the nontravelers in the population. As can be seen in table 7, the results are not consistent for the two models. The Haifa 1984 model correctly estimated only 13.5% of the observed nontravelers in the Haifa data and 18.5% in the Tel Aviv data. The Tel Aviv model obtained better results estimating correctly 41.9% of the observed nontravelers in the Tel Aviv data and 34.7% in the Haifa data. These results encourage further research.

Table 8 presents the estimation results of the 1996/97 regression models. As can be seen, the 1996/97 coefficients differ substantially from those of 1984, and most of the coefficients are significant at the 95% confidence level. The best model specification for 1984 was found to be also the best specification for the 1996/97 model, indicating that the most important variables affecting trip generation are similar in both models. The main problem raised during the basic comparison was the difference in the geographic scope (definition of the metropolitan survey area) for the two.

The overall R^2 of the 1996/97 models, 0.21 for the Tel Aviv model and 0.23 for the equivalent Haifa model, are even smaller than the values achieved for the 1984 models. But they are still not anomalous in the field of trip generation modeling. Statistical *Z*-tests conducted at the 95% level of confidence show that none of the coefficients are the same for the two metropolitan areas; that is, the coefficients differ in space. To verify the results, we conducted Chow tests for the transferability of the models. The calculated statistic was 6.91 in comparison with the 1.72 tabular *F*-value (at a 95% level of confidence), thus yielding the same conclusion, that the 1996/97 models are not transferable in space.

TABLE 5 Tobit Models by Metropolitan Areas: 1984

	Haifa		Tel Aviv	
Variable	**Coefficient**	***t* statistic**	**Coefficient**	***t* statistic**
Regression	3.91	41.67	3.15	40.23
Age				
8–13	1.25	7.79	0.81	6.01
14–17	1.46	7.97	1.76	10.74
18–29	0.95	6.95	1.30	11.08
30–64	0.63	6.36	0.99	11.43
65+	0.00	—	0.00	—
Number of cars per household				
0	–1.10	–11.30	–1.11	–12.76
1	–0.48	–4.91	–0.57	–6.77
2	0.00	—	—	—
Driver's license				
No license	–1.12	–13.24	–0.89	–11.92
License	0.00	—	0.00	—
Employed				
Not employed	–1.65	–19.80	–1.67	–22.20
Employed	0.00	—	0.00	—
Household status				
Head (0 = no)	–0.36	–4.29	–0.42	–5.85
Head (1 = yes)	0.00	—	0.00	—
Education				
0 years	–1.70	–5.36	–1.70	–7.04
1–8 years	–1.06	–9.82	–1.02	–11.14
9–12 years	–0.32	–3.38	–0.41	–4.79
13+ years	0.00	—	0.00	—
Standard error of *U*	2.40		2.50	
Overall pseudo R^2	0.33		0.32	
Number of observations used	3,243		4,356	
Log likelihood	–5,742.89		–7,173.62	
Log likelihood constant only	–6,453.65		–8,134.61	

TABLE 6 Aggregate Estimation Results Using Linear Tobit Models, by Year

	Haifa data			Tel Aviv data		
1984 data and models	**Tel Aviv model**	**Haifa model**	**Actual**	**Haifa model**	**Tel Aviv model**	**Actual**
Average trips per person per day	1.84	2.18	2.18	2.18	1.84	1.84
Observations with negative estimates	24	0	—	0	0	—
Minimum trip rate	0	0	0	0	0	0
Maximum trip rate	4.43	5.21	11.00	5.21	4.86	10.00
Total	5,115	7,321	7,059	9,772	7,883	8,019
Difference between model estimation and actual values	**–27.54**	**3.71**	**—**	**21.85**	**–1.69**	**—**
Standard deviation	1.30	1.52	2.08	1.50	1.48	1.90

	Haifa data			Tel Aviv data		
1996/97 data and models	**Tel Aviv model**	**Haifa model**	**Actual**	**Haifa model**	**Tel Aviv model**	**Actual**
Average trips per person per day	1.93	2.02	2.07	2.21	2.06	2.07
Observations with negative estimates	0	0	—	0	0	—
Minimum trip rate	0	0	—	0	0	—
Maximum trip rate	4.30	4.57	11.00	5.05	4.41	11.00
Total	9,672	10,184	10,414	34,797	32,363	32,455
Difference between model estimation and actual values	**–7.13**	**2.20**	**—**	**7.21**	**–0.28**	**—**
Standard deviation	1.28	1.42	2.08	1.42	1.29	1.99

TABLE 7 Tobit Model Estimation of Nontravelers: 1984

	Haifa data			Tel Aviv data		
	Tel Aviv model	**Haifa model**	**Observed**	**Haifa model**	**Tel Aviv model**	**Observed**
Estimated nontravelers (persons)	752	277	1,655 (37.99% of total observations)	244	635	1,014 (31.27% of total observations)
Observed nontravelers predicted as nontravelers	575	224	—	188	425	—
Right estimation (%)	34.7	13.5	—	18.5	41.9	—
Common observed and estimated (%)	76.5	80.6	—	77.0	66.9	—

Key: Right estimation = observed nontravelers predicted as nontravelers / observed * 100 ; common = observed nontravelers predicted as nontravelers / estimated * 100.

Table 9 presents the predicted average daily trips using the 1996/97 data for each model and each metropolitan area. As in 1984, the 1996/97 Haifa model overpredicts trips compared with the Tel Aviv model, however, the differences are smaller. One should remember that in regression models, the regression line always passes through the average (center of gravity of the observations). Since the observed average number of trips (in Tel Aviv and Haifa) was equal in the 1996/97 metropolitan files, the estimation difference was expected to be smaller.

However, the similar predicted aggregate trip rates indicate overrepresentation of particular sections of the population. For example, the calculated average car availability per household in metropolitan Tel Aviv was higher in the 1996/97 surveys than in metropolitan Haifa (0.60 > 0.55), but in 1984 the average car availability per household was almost the same (0.489 ≈ 0.484). Statistically, different definitions of the sampling areas could affect the transferability of the estimated models

Table 10 shows the Tobit model results for 1996/97. A comparison with table 8 shows the resemblance in the effect of the explanatory variables and the difference in the magnitude of the coefficients between the Tobit and the regression models. Transferring the models in space and evaluating the estimated average trip rate from the models, for 1996/97, we found that the Haifa Tobit model overestimated the trips in the Tel Aviv file by 7.2% and the Tel Aviv Tobit model underestimated the trips in the Haifa file by 7.1% (table 6). These values are quite similar to the over- and underprediction of the equivalent regression models shown in table 9. Spatial transferability *t*-tests held at the 95% level of confidence show that most of the coefficients are not significantly different between the metropolitan areas, except age categories and the education "non-educated" category. χ^2 tests for the spatial transferability of the models at the same level of confidence reach the same conclusion $\left(\chi^2_{13,\,0.95} = 22.36 << 82.01\right)$.

In table 11, the 1996/97 Tobit model prediction of nontravelers is even worse than in the analogous 1984 models. When trying to represent nontravelers, the Haifa 1996/97 Tobit model captured only 6.8% of the observed nontravelers in the Tel Aviv file and 11.8% in the Haifa file. The Tel Aviv 1996/97 Tobit model captured only 3.8% of the nontravelers in the Haifa file and 9.8% in the Tel Aviv file. A point worth mentioning is the resemblance in the proportion of nontravelers in the two surveys (about 35% of the persons represented in the sample files did not generate trips). Finally, about 70% to 75% of the estimated nontravelers are observed nontravelers.

Table 12 shows the estimation results for temporal transferability of the regression and Tobit models. When we tested for temporal transferability using the 1984 models to predict 1996/97 trip rates, we observed that the 1984 Tel Aviv regression model underestimated the observed total number of trips in Tel Aviv in 1996/97 by 7%. The Haifa 1984 regression model overestimated the observed total number of trips in 1996/97 Haifa data by only 2.8%. Taking into account that the average number of daily trips in the Haifa 1984 survey was 2.17 and in the 1996/97 survey it was 2.07, the difference is not surprising. However, it may also be affected by the different definition of the geographic scope of the two household surveys.

Chow tests of the temporal stability of the 1984 models compared with the 1996/97 show that the statistic for the 1984 Tel Aviv model was 4.53, bigger than the 1.72 tabular *F*-statistic at the 95% level of confidence, meaning that the 1984 coefficients differ from the 1996/97 coefficients. The statistic for the temporal stability of the 1984 Haifa model was 8.39 compared with the 1.72 tabular *F*, reaching the same conclusion. Transferability χ^2 tests of the Tobit models in time show that, at the 95% level of confidence, we can reject the null hypothesis; that is, the models for the two time points are different $\left(\chi^2_{13,\,0.95} = 22.36 < 128.72\right)$.

CONCLUSIONS

In our research, statistical tests indicated that the regression and Tobit models estimated for two metropolitan areas and two time periods differ statistically in time and in space. One exception was the Tobit transferability in space, where the coefficients

TABLE 8 Regression Models by Metropolitan Area: 1996/97

Variable	Haifa Coefficient	Haifa t statistic	Haifa Standard error	Tel Aviv Coefficient	Tel Aviv t statistic	Tel Aviv Standard error
Regression coefficient	3.34	29.67	0.11	3.30	55.10	0.06
Age						
8–13	0.19	1.62	0.12	0.14	2.11	0.07
14–17	0.73	5.70	0.13	0.40	5.55	0.07
18–29	0.45	4.65	0.10	0.36	6.54	0.05
30–64	0.55	6.35	0.09	0.34	7.00	0.05
65+	0.00	—	—	0.00	—	—
Number of cars per household						
0	–0.81	–9.50	0.08	–0.73	–16.76	0.04
1	–0.40	–5.52	0.07	–0.35	–9.96	0.03
2+	0.00	—	—	0.00	—	—
Driver's license						
No license	–0.79	–10.51	0.07	–0.68	–16.59	0.03
License	0.00	—	—	0.00	—	—
Employed						
Not employed	–0.84	–11.93	0.07	–0.78	–20.74	0.04
Employed	0.00	—	—	0.00	—	—
Household status						
Head (0 = no)	–0.33	–5.28	0.06	–0.34	–10.12	0.03
Head (1 = yes)	0.00	—	—	0.00	—	—
Education						
0 years	0.39	3.39	0.11	0.15	2.29	0.07
1–8 years	–0.40	–4.25	0.09	–0.39	–7.16	0.05
9–12 years	–0.06	–0.90	0.07	–0.19	–5.32	0.04
13+ years	0.00	—	—	0.00	—	—
Overall R^2	0.23			0.21		
Number of observations used	5,027			15,689		

Notes: Chow test for 1984 model transferability:
Ess_1 = error sum of squares of Haifa set
Ess_2 = error sum of squares of Tel Aviv set
Ess_3 = error sum of squares of combined set
K = number of parameters in the model including the constant
N_i = number of observations in model i

$$H_0: \beta_i^1 = \beta_i^2$$

$$H_1 = else \qquad \alpha = 0.05$$

$$\frac{Ess_3 - (Ess_1 + Ess_2)}{K} \bullet \frac{N_1 + N_2 - 2 \bullet K}{Ess_1 + Ess_2} = \frac{65202 - (16635 + 48285)}{13} \bullet \frac{5027 + 15689 - 2 \bullet 13}{16635 + 48285} = 6.91$$

$$F_{(k, N_1 + N_2 - 2 \bullet K)} = F_{(13, \infty)} = 1.72 \Rightarrow H_1$$

TABLE 9 Aggregate Estimation Results Using Regression Models: 1996/97

	Haifa data			Tel Aviv data		
	Tel Aviv model	Haifa model	Actual	Haifa model	Tel Aviv model	Actual
Average trips per person per day	1.93	2.07	2.07	2.22	2.07	2.07
Observations with negative estimates	0	0	—	0	0	—
Minimum trip rate	0.37	0.16	—	0.16	0.37	—
Maximum trip rate	3.67	3.89	11.00	4.28	3.80	11.00
Total	9,748	10,450	10,414	34,788	32,341	32,455
Difference between model estimation and actual values (%)	**–6.39**	**0.35**	**—**	**7.19**	**–0.35**	**—**
Standard deviation	0.91	1.00	2.08	1.01	0.92	1.99

from the two models for the same year were not significantly different. The distinction cannot be well explained, but it might be due partially to geographic, demographic, socioeconomic, and spatial structure differences between the two metropolitan areas. The smaller sample size and scope of the 1984 household survey compared with the 1996/97 household survey (as shown in table 2) did not allow us to represent the ethnicity of the survey participants, a variable believed to be related to trip generation. Also, the incorporation in the models of a pure economic variable such as income was not possible, because it was not included in the 1996/97 survey.

We ascribed the temporal instability of the estimated models to changes in the structure and development of the metropolitan areas of Tel Aviv and Haifa, changes in lifestyle and socioeconomic variables that are not all accounted for in the model, as well as the inconsistency of the two surveys. A partial explanation may be that 1984 was an economically unstable year, featuring high inflation rates and uncertainty, while 1996/97 was considered to be economically stable.

An important conclusion based on our results is that in order for trip generation models to be transferable they need to account for variables not included in the current models: income, land use and spatial structure, the economy, the transportation system and accessibility, and more detailed socioeconomic and life style variables. If we could estimate a perfect disaggregate model accounting for all factors that affect trip generation and with appropriate segmentation, it would likely be transferable. With this data lacking, models are not transferable, because unobserved variables affect coefficients of observed variables with which they are correlated.

Another conclusion is that household surveys conducted on a regular basis will be more useful if the design stays constant. Differences in the structure, variables, range, investigation period, definition of the variables, and database structure affect the transferability of the estimated models.

We also would emphasize the need for further research on the implementation of Tobit models in the context of trip generation. Tobit models tend to represent the mechanism of trip generation more realistically, capturing and estimating (partially) nontravelers. As a combination of regression and discrete choice models, the Tobit model may be more suitable for implementation in TGM than discrete choice or regression models, particularly because Tobit is better formulated to differentiate nontravelers from travelers. The underestimation of nontravelers may be partly due to the fact that we did not necessarily estimate the best Tobit model.

For the linear regression models, almost all variables were significant at the 95% confidence level, but the coefficients were shown to vary in time and space. For the Tobit model, while almost all variables were significant at the 95% confidence level, the coefficients of the models of the two metropolitan areas were statistically similar but they differed in time for each city.

TABLE 10 Tobit Models by Metropolitan Areas: 1996/97

	Haifa		Tel Aviv	
Variable	**Coefficient**	***t* statistic**	**Coefficient**	***t* statistic**
Regression intercept	3.00	36.84	3.04	68.47
Age				
8–13	0.62	3.69	0.46	5.03
14–17	1.53	8.93	0.85	9.14
18–29	1.02	8.83	0.76	12.14
30–64	1.07	12.10	0.68	14.20
65+	0.00	—	0.00	—
Number of cars per household				
0	–1.20	–12.68	–1.07	–20.42
1	–0.52	–6.09	–0.47	–10.40
2+	0.00	—	0.00	—
Driver's license				
No license	–1.08	–12.89	–0.97	–21.38
License	0.00	—	0.00	—
Employed				
Not employed	–1.36	–17.25	–1.27	–29.96
Employed	0.00	—	0.00	—
Household status				
Head (0 = no)	–0.56	–7.42	–0.48	–11.66
Head (1 = yes)	0.00	—	0.00	—
Education				
0 years	0.50	4.21	0.21	3.14
1–8 years	–0.77	–7.27	–0.60	–10.06
9–12 years	–0.12	–1.34	–0.19	–4.21
13+ years	0.00	—	0.00	—
Standard error of *U*	2.85		2.67	
Overall pseudo R^{2*}	0.23		0.21	
Number of observations used	5,026		15,689	
Log likelihood	–9,060.98		–28,376.08	
Log likelihood constant only	–9,806.57		–30,515.70	

* Pseudo R^2 is a measure of goodness of fit similar to R^2 in ordinary least squares. The residuals are calculated based on the maximum likelihood estimators. For details of its calculation, see the *EasyReg* manual, available at http://econ.la.psu.edu/~hbierens/EasyRegTours/ TOBIT_Tourfiles/TOBIT.PDF.

TABLE 11 Tobit Estimation of Nontravelers by Metropolitan Areas: 1996/97

	Haifa data			Tel Aviv data		
	Haifa model	**Tel Aviv model**	**Observed**	**Tel Aviv model**	**Haifa model**	**Observed**
Nontravelers (persons)	293	90	1,769 (35.19%)	722	505	5,167 (32.93%)
Observed nontravelers predicted as nontravelers	209	67	—	504	353	—
Right estimation (%)	11.81	3.78	—	9.75	6.83	—
Common observed and estimated (%)	71.33	74.44	—	69.80	69.90	—

Key: Right estimation = observed nontravelers predicted as nontravelers / observed * 100;
common = observed nontravelers predicted as nontravelers / estimated * 100.

TABLE 12 Temporal Transferability Results of the 1984 Models in the 1996/97 Files

	Haifa data			Tel Aviv data		
TEL AVIV REGRESSION	**1984 Tel Aviv model**	**1996/97 Tel Aviv model**	**Actual**	**1984 Tel Aviv model**	**1996/97 Tel Aviv model**	**Actual**
Average trips per person per day	1.76	1.94	2.07	1.92	2.06	2.07
Negative estimations	545	—	—	1,464	—	—
Minimum trip rate	–0.40	0.37	0	–0.40	0.16	0
Maximum trip rate	4.04	3.67	11	4.04	3.80	11
Total	8,850	9,748	10,414	30,190	32,341	32,455
Difference between model estimation and actual values (%)	**–15.01**	**–6.4**	**—**	**–6.98**	**–0.35**	**—**
Standard deviation	1.27	0.91	2.07	1.28	0.92	1.99
HAIFA REGRESSION	**1984 Tel Aviv model**	**1996/97 Tel Aviv model**	**Actual**	**1984 Tel Aviv model**	**1996/97 Tel Aviv model**	**Actual**
Average trips per person per day	21.30	2.08	2.07	2.31	22.20	2.07
Negative estimations	0	0	—	—	0	—
Minimum trip rate	0.11	0.16	0	0.11	0.16	0
Maximum trip rate	4.49	3.89	11	4.49	4.28	11
Total	10,710	10,450	10,414	36,287	34,788	32,455
Difference between model estimation and actual values (%)	**2.84**	**0.35**	**—**	**11.80**	**7.19**	**—**
Standard deviation	1.34	1.00	2.07	1.35	1.01	1.99

(*continues on next page*)

TABLE 12 Temporal Transferability Results of the 1984 Models in the 1996/97 Files *(continued)*

	Haifa data			Tel Aviv data		
TEL AVIV TOBIT	**1984 Tel Aviv model**	**1996/97 Tel Aviv model**	**Actual**	**1984 Tel Aviv model**	**1996/97 Tel Aviv model**	**Actual**
Average trips per person per day	1.83	1.92	2.07	2.06	1.98	2.07
Minimum trip rate	0	0	0	0	0	0
Maximum trip rate	4.86	4.30	11	4.41	4.84	11
Total	9,203	9,671	10,414	32,363	31,155	32,455
Difference between model estimation and actual values (%)	**–11.60**	**–7.12**	—	**0.28**	**–4.00**	—
Standard deviation	1.58	1.28	2.07	1.30	1.61	1.99
HAIFA TOBIT	**1984 Tel Aviv model**	**1996/97 Tel Aviv model**	**Actual**	**1984 Tel Aviv model**	**1996/97 Tel Aviv model**	**Actual**
Average trips per person per day	2.39	2.03	2.07	2.22	2.32	2.07
Minimum trip rate	0	0	0	0	0	0
Maximum trip rate	5.20	4.57	11	5.05	5.20	11
Total	12,036	10,184	10,414	34,796	36,381	32,455
Difference between model estimation and actual values (%)	**15.58**	**2.20**	—	**7.22**	**12.09**	—
Standard deviation	1.28	1.42	2.07	1.43	1.75	1.99

The nature of the local household surveys raises a need to validate the results of this study in future research. In particular, further research can identify what makes two study areas "similar enough" to justify transferring a model from one to the other. We also suggest further research incorporating Tobit models in TGM and for investigating the characteristics of nontravelers.

REFERENCES

Atherton, T. and M. Ben-Akiva 1976. Transferability and Updating of Disaggregate Travel Demand Models. *Transportation Research Record* 610:12–18.

Caldwell III, L.C. and M.J. Demetsky. 1980. Transferability of Trip Generation Models. *Transportation Research Record* 751:56–62.

Copley, G. and S.R. Lowe. 1981. The Temporal Stability of Trip Rates: Some Findings and Implications. *Transportation Analysis and Models: Proceedings of Seminar N*. London, England: PTRC Education and Research Services, Ltd.

Cotrus, A. 2001. Analysis of Trip Generation Characteristics in Israel for the Years 1984, 1996/7 and Spatial and Temporal Transferability of Trip Generation Models, Ph.D. thesis, Technion, Israel Institute of Technology, Haifa, Israel.

Daor, E. 1981. The Transferability of Independent Variables in Trip Generation Models. *Transportation Analysis and Models: Proceedings of Seminar N*. London, England: PTRC Education and Research Services, Ltd.

Doubleday, C. 1977. Some Studies of the Temporal Stability of Person Trip Generation Models. *Transportation Research* 11:255–263.

Downes, J.D. and L. Gyenes. 1976. Temporal Stability and Forecasting Ability of Trip Generation Models in Reading, TRRL Report 726. Crowthorne, Berkshire: Transport and Road Research Laboratory.

Duah, K.A. and F.L. Hall. 1997. Spatial Transferability of an Ordered Response Model of Trip Generation. *Transportation Research A* 31:389–402.

Fleet, C. and S. Robertson. 1968. Trip Generation in the Transportation Planning Process. *Highway Research Board Record* 240:11–31.

Hald, A. 1949. Maximum Likelihood Estimation of Parameters of a Normal Distribution Which is Truncated at a Known Point. *Skandinavisk Aktuarietidskrift* 32:119–134.

Huisken, G. 2000. Neural Networks and Fuzzy Logic to Improve Trip Generation Modeling, paper presented at the 79th Annual Meetings of the Transportation Research Board, Jan. 9–13, 2000, Washington, DC.

Kannel, E.J. and K.W. Heathington. 1973. Temporal Stability of Trip Generation Relations. *Highway Research Record* 472:17–27.

Karasmaa, N. and M. Pursula. 1997. Empirical Studies of Transferability of Helsinki Metropolitan Area Travel Forecasting Models. *Transportation Research Record* 1607:38–44.

McDonald, J.F. and R.A. Moffitt. 1980. The Uses of Tobit Analysis. *The Review of Economics and Statistics* 62(2):318–321.

Ortuzar, J. and L.G. Willumsen. 1994. Trip Generation Modeling. *Modeling Transportation*, 2nd ed. New York, NY: Wiley.

Prashker, J. 1982. Transferability of Disaggregate Modal-Split Models—Investigation of Models in the Metropolitan Area of Tel Aviv, manuscript, p. 91. Israel Institute of Transportation.

Rose, G. and F.S. Koppelman. 1984. Transferability of Disaggregate Trip Generation Models. *Proceedings of the 9th International Symposium on Transportation and Traffic Theory.* Utrecht, Netherlands: VNU Press.

Silman, L.A. 1981. The Time Stability of a Modal-Split Model for Tel Aviv. *Environment and Planning A* 13:751–762.

Smith, R.L. and D.E. Cleveland. 1976. Time Stability Analysis of Trip Generation and Predistribution Modal—Choice Models. *Transportation Research Record* 569:76–86.

Supernak, J. 1982. Transportation Modeling: Lessons from the Past and Tasks for the Future. *Transportation Analysis and Models, Proceedings of Seminar Q*. London, England: PTRC Education and Research Services, Ltd.

____. 1984. Disaggregate Models of Mode Choices: An Assessment of Performance and Suggestions for Improvement. *Proceedings of the 9th International Symposium on Transportation and Traffic Theory.* Utrecht, Netherlands: VNU Press.

____. 1987. A Method for Estimating Long-Term Changes in Time-of-Day Travel Demand. *Transportation Research Record* 1138:18–26.

Supernak, J., A. Talvitie, and A. DeJohn. 1983. Person-Category Trip Generation Model. *Transportation Research Record* 944:74–83.

Tobin, J. 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica* 26:24–36.

Walker, W.T. and O.A. Olanipekun. 1989. Interregional Stability of Household Trip Generation Rates from the 1986 New Jersey Home Interview Survey. *Transportation Research Record* 1220:47–57.

Walker, W.T. and H. Peng. 1991. Long Range Temporal Stability of Trip Generation Rates Based on Selected Cross-Classification Models in the Delaware Valley Region. *Transportation Research Record* 1305:61–71.

Wilmot, C.G. 1995. Evidence of Transferability of Trip Generation Models. *Journal of Transportation Engineering* 9:405–410.

Yunker, K.R. 1976. Tests of the Temporal Stability of Travel Simulation Models in Southeastern Wisconsin. *Transportation Research Record* 610:1–5.

Zhao, H. 2000. Comparison of Two Alternatives for Trip Generation, paper presented at the 79th Annual Meetings of the Transportation Research Board, Jan. 9–13, 2000, Washington, DC.

Governors Highway Safety Associations and Transportation Planning: Exploratory Factor Analysis and Structural Equation Modeling

SUDESHNA MITRA[1,*]
SIMON WASHINGTON[1]
ERIC DUMBAUGH[2]
MICHAEL D. MEYER[2]

[1] Department of Civil and Environmental Engineering
Arizona State University
PO Box 875306
Tempe, AZ 85287-5306

[2] School of Civil and Environmental Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0355

ABSTRACT

The Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 mandated the consideration of safety in the regional transportation planning process. As part of National Cooperative Highway Research Program Project 8-44, "Incorporating Safety into the Transportation Planning Process," we conducted a telephone survey to assess safety-related activities and expertise at Governors Highway Safety Associations (GHSAs), and GHSA relationships with metropolitan planning organizations (MPOs) and state departments of transportation (DOTs). The survey results were combined with statewide crash data to enable exploratory modeling of the relationship between GHSA policies and programs and statewide safety. The modeling objective was to illuminate current hurdles to ISTEA implementation, so that appropriate institutional, analytical, and personnel improvements can be made. The study revealed that coordination of transportation safety across DOTs, MPOs, GHSAs, and departments of public safety is generally beneficial to the implementation of safety. In addition, better coordination is characterized by more positive and constructive attitudes toward incorporating safety into planning.

KEYWORDS: Transportation planning, transportation safety, structural equation modeling.

Email addresses:
* Corresponding author—mitra_sudeshna@yahoo.com
S. Washington—Simon.Washington@asu.edu
E. Dumbaugh—EDumbaugh@aol.com
M.D. Meyer—mmeyer@ce.gatech.edu

INTRODUCTION

The Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 is, in many ways, a benchmark of federal transportation legislation. Along with the subsequent 1998 Transportation Efficiency Act for the 21st Century (TEA-21), not only did it define the post-interstate transportation program, it also broadened the types of issues that were to be considered as part of the transportation planning process.

By mandating the consideration of a broader range of issues to address in planning, the projects and strategies surviving the planning and programming processes should relate to those issues. There are challenges in meeting this mandate, however, due in part to "institutional inertia" in many state departments of transportation (DOTs) and metropolitan planning organizations (MPOs) to continue the programming emphasis on capital-intensive projects. ISTEA reinforced the change in focus away from capital-intensive projects with the requirement for six management systems, one of which targeted safety. By introducing a process to identify system deficiencies, analyze and evaluate prospective improvement strategies, and monitor implemented projects and strategies, it is possible to determine whether anticipated effects occurred.

Major stakeholders in the transportation and safety fields are varied and have no tradition of interacting within the context of the planning process. Our interest here is whether MPOs and DOTs interact with their respective Governors Highway Safety Association (GHSA) representatives, who are often the focal point for state initiatives dealing with issues such as drunk driving, seat belt use, and teenage driving. Furthermore, it is important to know whether GHSA and MPO/DOT coordination makes a difference in terms of statewide safety.

The inspiration for GHSA dates back to the Highway Safety Act of 1966, which established state offices of highway safety. In an effort to share information among state safety offices, the National Conference of Governors' Highway Representatives was created. GHSA grew out of this and, in 1974, it incorporated. GHSA includes highway safety program managers from all 50 states, the District of Columbia, Puerto Rico, the Northern Marianas, all U.S. territories, and the Indian Nation. The member agencies are tasked to develop, implement, and oversee highway safety programs using behavioral strategies such as training and educating motorists, pedestrians, bicyclists, and school children on safe behavior, and by addressing impaired driving, speeding, aggressive driving, and safety restraint use. Given the mission of the GHSAs, their impact on statewide safety is vital, and their cooperation and coordination with MPOs and DOTs may play a pivotal role in the ultimate success of incorporating safety into the transportation planning process.

This paper presents the results of a telephone survey designed and administered to state GHSA offices to capture the characteristics, attitudes, and activities of these agencies. (The survey was part of National Cooperative Highway Research Program (NCHRP) Project 8-44, "Incorporating Safety into Long-Range Transportation Planning.") Specific objectives of the survey include:

1. understanding if and how the agencies' mission statements, goals, and/or objectives address various safety issues;
2. characterizing the nature of the implemented programs;
3. determining whether an agency considered integrating the state safety program with specific transportation-related activities; and
4. the extent of GHSA participation in regional transportation planning and interaction with MPOs and DOTs.

Two research questions of particular interest arose from the survey.

1. Are the depth and breadth of programs commensurate with statewide safety? In other words, does the safety level within a state drive the adoption of programs? Will a state with a poorer safety record have broader and more extensive safety programs, for example, and does that indicate that GHSA activities and funding are out of step with safety or are lagging?
2. Do GHSA perceptions of the benefits of transportation planning influence programming efforts and/or statewide safety levels? In other words, does cooperation between GHSAs, DOTs, and MPOs lead to more

extensive statewide safety programming and/or improved safety?

We used latent variable models to look for answers to these research questions in a quantitatively rigorous way. We chose this type of model because of the type of data in the study—many variables are not directly observable (latent), and thus their proxies (variables we can measure directly) suffer from measurement errors.

While latent variables and measurement errors of their proxies are widespread in social science research, they are relatively uncommon in transportation research. Latent variables refer to unobservable or unmeasured variables, such as intelligence, education, social and political classes, and attitudes. Often proxies can be used to indirectly measure latent variables, such as IQ score and grade point average, as measures of the latent variable intelligence. Certain effects of these latent variables on measurable variables are observable, along with some random or systematic errors, collectively called measurement errors. Everitt (1984) pointed out that it was indeed one of the major achievements in the behavioral sciences to develop methods that assess and explain the structure in a set of correlated, observed variables, in terms of a small number of latent variables.

In this study, the variables indicating attitudes of GHSA personnel, their planning and programming efforts, and coordination with MPOs and DOTs are not directly measurable. While the survey responses aim to measure these underlying latent variables, some of their dimensions may remain unexplored. This paper presents various latent variable analysis techniques used to examine and extract relationships in the data, including factor analysis (exploratory) and structural equation modeling.

THE SURVEY

During fall 2002 and spring 2003, the research team conducted telephone surveys of GHSA personnel in the 50 states and the District of Columbia to capture their planning attitudes, types of programs, goal-setting criteria, coordination efforts, and perceived influence on transportation planning. Respondents were asked a series of formal survey questions aimed at understanding the relationships between planning efforts and safety issues, as well as several open-ended questions intended to capture the unique viewpoints, activities, and perspectives of each of the individual agencies. The survey instrument was pre-tested in two states, and the final survey instrument was revised based on the pretest results. The survey instrument appears in the appendix of this paper.

While individuals designated as the Governor's Representative for Highway Safety in a specific state were initially targeted as the appropriate agency respondents, conversations with actual governors' representatives soon revealed that the individuals actually managing the development and implementation of GHSA programs were often not the designated representatives themselves, who were typically high-level personnel in other state agencies. Instead, in many cases, individuals hired for the express task of managing these programs were interviewed.

The respondent recruitment effort consisted of multiple attempts to contact each of the respondents via telephone, followed by an email contact to encourage each individual's participation. Ultimately, telephone surveys, averaging 20 minutes in length, were completed for 43 of the 51 potential respondents. Two of the states completed the survey electronically, bringing the total number of states surveyed to 45. Despite an exhaustive effort to reduce survey nonresponse, responses for six states were not obtained. However, among the 45 completed surveys, item nonresponse was not a problem.

The survey was designed to capture six major characteristics of a GHSA.

- the types of planning-related activities undertaken;
- attitudes toward planning as reflected by GHSA efforts to include specific safety issues in transportation planning activities, as well as how much GHSA participated in the regional transportation planning process;
- whether the GHSA office is affiliated with another state agency;
- the extent of coordination with other agencies;
- the planning time horizons of the agency; and
- the number of agency staff.

SUPPLEMENTAL STATE-LEVEL DATA

In addition to the survey responses, state-level safety data were obtained. These data from the National Highway Traffic Safety Administration (NHTSA) included total fatalities, alcohol-related fatalities, and pedestrian and bicycle-related fatalities for calendar year 2001 (USDOT 2001). To compensate for exposure to risk, crash rates were taken into account rather than the total crash statistics. In other words, a state with a large population and relatively higher vehicle-miles traveled (VMT) can be expected to experience a greater number of crashes than a smaller state. Hence, to minimize bias imposed by population size, fatality rates per 100 million VMT were considered for total as well as alcohol-related crashes. We used fatality rates per 100,000 population for pedestrian and bicycle crashes. The crash rates are used as observed endogenous variables and served as proxies for the latent variable *RISK* (motor vehicle safety-related risk) across states. Better metrics for pedestrian and bicycle exposure are theoretically possible but are not generally available.

In addition to these data, this research utilized various other sources of information, including enacted legislation in the states covering seat belt laws, laws related to impaired driving, helmet laws, child restraint laws, and so forth (IIHS 2004). Despite a priori expectations, these variables were not found to be statistically significant in the modeling efforts. Table 1 presents the descriptive statistics of various observed variables employed in the final model.

TABLE 1 Summary of Survey Response

Variable name	Question number	Variable type	Scale range	Average	Minimum	Maximum	Variance
PEDS2	Q-2a	Yes/No	0–1	0.67	0	1	0.227
BIKE2	Q-2b	Yes/No	0–1	0.60	0	1	0.245
DRIVED2	Q-2c	Yes/No	0–1	0.27	0	1	0.200
SCHOOLED2	Q-2d	Yes/No	0–1	0.47	0	1	0.254
ENFORCE2	Q-2e	Yes/No	0–1	0.84	0	1	0.134
COOPDOT2	Q-2f	Yes/No	0–1	0.67	0	1	0.227
COCPLOC2	Q-2g	Yes/No	0–1	0.71	0	1	0.210
COOPPLAN2	Q-2h	Yes/No	0–1	0.40	0	1	0.245
SAFDES2	Q-2i	Yes/No	0–1	0.27	0	1	0.200
SAFOPS2	Q-2j	Yes/No	0–1	0.22	0	1	0.177
DATA2	Q-2k	Yes/No	0–1	0.80	0	1	0.164
STAFF3	Q-3	Ratio scale	NA	8.36	0	19	22.916
ENG4	Q-4a	Yes/No	0–1	0.18	0	1	0.149
PLAN4	Q-4b	Yes/No	0–1	0.33	0	1	0.227
OPS4	Q-4c	Yes/No	0–1	0.31	0	1	0.219
ENF4	Q-4d	Yes/No	0–1	0.76	0	1	0.188
EDU4	Q-4e	Yes/No	0–1	0.76	0	1	0.188
MKTG4	Q-4f	Yes/No	0–1	0.67	0	1	0.227
PERFMS5	Q-5	Ordinal	1–5	4.84	4	5	0.134
SHAREPM6	Q-6	Yes/No	0–1	1.02	0	9	3.249
PERFTGT7	Q-7a	Yes/No	0–1	0.98	0	1	0.022
TGTYEAR7	Q-7b	Ratio scale	NA	5.02	1	30	20.113

TABLE 1 Summary of Survey Response *(continued)*

Variable name	Question number	Variable type	Scale range	Average	Minimum	Maximum	Variance
AGYHZN8	Q-8	Ratio scale	NA	5.38	1	30	26.331
PEDEDU9	Q-9a	Yes/No	0–1	0.64	0	1	0.234
PEDCRWK9	Q-9b	Yes/No	0–1	0.40	0	1	0.245
PEDSCHL9	Q-9c	Yes/No	0–1	0.62	0	1	0.240
BIKEDU10	Q-10a	Yes/No	0–1	0.69	0	1	0.219
BKHELM10	Q-10b	Yes/No	0–1	0.53	0	1	0.254
BKLITE10	Q-10c	Yes/No	0–1	0.29	0	1	0.210
BKBRK10	Q-10d	Yes/No	0–1	0.09	0	1	0.082
DOT12	Q-12a	Ordinal	0–4	3.96	3	4	0.043
SPOLCE12	Q-12b	Ordinal	0–4	3.82	0	4	0.695
LPOLCE12	Q-12c	Ordinal	0–4	3.96	2	4	0.088
PLAN12	Q-12d	Ordinal	0–4	1.80	0	4	2.300
LEGS12	Q-12e	Ordinal	0–4	2.67	0	4	2.363
GOV12	Q-12f	Ordinal	0–4	2.84	0	4	1.907
HWY12	Q-12g	Ordinal	0–4	0.62	0	4	1.695
ENGNR12	Q-12h	Ordinal	0–4	1.24	0	4	2.098
SCHOOL12	Q-12i	Ordinal	0–4	2.87	0	4	2.345
LOCAL12	Q-12j	Ordinal	0–4	3.22	0	4	1.722
SDWLK13	Q-13a	Yes/No	0–1	0.33	0	1	0.227
CRSWLK13	Q-13b	Yes/No	0–1	0.42	0	1	0.249
BIKE13	Q-13c	Yes/No	0–1	0.51	0	1	0.255
SPEED13	Q-13d	Yes/No	0–1	0.58	0	1	0.249
TURN13	Q-13e	Yes/No	0–1	0.47	0	1	0.254
RDSIDE13	Q-13f	Yes/No	0–1	0.53	0	1	0.254
MONITR13	Q-13g	Yes/No	0–1	0.58	0	1	0.249
MPO14	Q14a	Yes/No	0–1	0.38	0	1	0.240
RURAL15	Q15a	Yes/No	0–1	0.24	0	1	0.188
STP16	Q16a	Yes/No	0–1	0.67	0	1	0.186
INFLU17	Q17	Ordinal	0–2	0.69	0	1	0.674
TOLFTVMT	Total fatality rate per 100 million VMT	Ratio scale	NA	1.59	0.96	2.32	0.128
ALCFTVMT	Any alcohol-related fatality rate per 100 million VMT	Ratio scale	NA	0.66	0.29	1.27	0.042
PEDFTPOP	Pedestrian fatality rate per 100,000 persons	Ratio scale	NA	1.48	0.47	2.98	0.331
BIKFTPOP	Bike-related fatality rate per 100,000 persons	Ratio scale	NA	0.22	0	0.77	0.025

Key: VMT = vehicle-miles traveled.

METHODOLOGY

An exploratory factor analysis followed by structured equation modeling (SEM) was used to model structured relationships between latent variables. Latent variable models have been widely applied in various fields, but rarely in transportation; for example, in sociology by Amato and Alan (1995), in psychology by Östberg and Hagekull (2000) and Rubio et al. (2001), and in construction management by Molenaar et al. (2000).

While few applications exist in transportation, Ben-Akiva et al. (1991), working in the area of pavement management, introduced the concept of latent performance in terms of several observable performance indicators and used measurement as well as structural models to model relationships. The results of this model were then used for decisions on, for example, optimal maintenance and inspection policies, expected number of inspections for the optimum policies, and the minimum expected cost of inspecting and maintaining a facility over various planning time horizons. Another important study in transportation by Golob and Regan (2000) used confirmatory factor analysis to look at the interrelationship among the latent policy evaluations through several exogenous variables defining differences in freight operations.

In statistical modeling, applying knowledge of the underlying data-generating process is a critical step when developing a "starter specification." However, in the absence of well developed theories, it is often difficult for an analyst to specify a priori which observed variables affect which latent variable. In this context, Loehlin (2004) discussed exploratory factor analysis (EFA) as a method to discover and define latent variables as well as a measurement model that can provide the basis for a causal analysis of relationships among the latent variables.

Methodological Details

As described in Washington et al. (2003), EFA is not a statistical model and there is no distinction between dependent and independent variables in this analysis. For the EFA is to be useful, there are $K < n$ factors or principal components, with the first factor given as

$$Z_1 = a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1q}x_q + \ldots\ldots + a_{1(p-1)}y_{(p-1)} + a_{1p}y_p \quad (1)$$

which maximizes the variability across individuals, subject to the constraint

$$a_{11}^2 + a_{12}^2 + a_{13}^2 + \ldots\ldots + a_{1q}^2 + \ldots\ldots + a_{1(p-1)}^2 + a_{1p}^2 = 1 \quad (2)$$

where observed variables are denoted by $(p \times 1)$ column vector y, and $(q \times 1)$ column vector x, and influence the latent endogenous and exogenous variables, respectively. Thus, $VAR[Z_1]$ is maximized given the constraints in equation (2), with the constraints imposed to ensure determinacy. A second factor, Z_2, is then sought to maximize the variability across individuals, subject to the constraints

$$a_{21}^2 + a_{22}^2 + a_{23}^2 + \ldots\ldots + a_{2(p-1)}^2 + a_{2p}^2 = 1$$

and $COR[Z_1,Z_2] = 0$, and so on, such that $COR[Z_1, Z_2 \ldots\ldots Z_K] = 0$ for up to K factors. In this paper, the various survey responses are the observed variables, and they are used to identify the underlying latent variables. Manly (1986), Johnson and Wichern (2002), and Washington et al. (2003) provide additional details of EFA.

After identifying useful factors from EFA, SEMs are developed. A SEM is defined with two components, a measurement model and a structural model. SEMs are a natural extension of factor analysis and are used to identify structural relationships between latent as well as observed variables. The measurement model portion of a SEM correlates the observed variables with latent dependent, as well as independent, variables. Observed variables influencing latent endogenous and exogenous variables are denoted by a $(p \times 1)$ column vector of y and $(q \times 1)$ column vector of x, such that

$$y = \Lambda_y \eta + \varepsilon \quad (3)$$

$$x = \Lambda_x \xi + \delta \quad (4)$$

where

$\Lambda_y(p \times m)$ and $\Lambda_x(q \times n)$ are the coefficient matrices that show the relation of y to η and x to ξ, respectively, and $\varepsilon(p \times 1)$, and $\delta(q \times 1)$ are the errors of measurement for y and x, respectively (Washington et al. 2003). For example, if *statewide safety* is a latent dependent variable of interest, it is denoted as η and all the observed variables, such as alcohol-related fatalities, pedestrian and bicycle-related fatalities, etc., would constitute the y vector.

The structural component of a SEM is given as

$$\eta = B\eta + \Gamma\xi + s \tag{5}$$

where

η is an $(m \times 1)$ vector of latent endogenous random variables,

ξ is a vector of $(n \times 1)$ latent exogenous random variables,

B is an $(m \times m)$ coefficient matrix reflecting the influence of the latent endogenous variables on each other,

Γ is an $(m \times n)$ coefficient matrix for the effects of ξ on η, and

s is the vector of regression errors for which $[E(s) = 0]$ and is uncorrelated with ξ. In addition, the error terms of the measurement models are assumed to be uncorrelated with ξ and s.

From the previous simultaneous equation (5), and treating all the observed variables as dependent variables in the model, the covariance matrix is given as

$$\Sigma(\theta) = G(I-\beta)^{-1}\gamma\varphi\gamma(I-\beta)^{(-1)'}G' \tag{6}$$

where G is the selection matrix containing either zero or one to select the observed variables from all the dependent variables in η. Once the SEM model is identified (statistically), the parameters are estimated using a discrepancy function based on the hypothesized model $\Sigma = \Sigma(\theta)$, where Σ is estimated by the sample covariance matrix S. The role of this discrepancy function is to minimize the difference between the sample variance-covariance matrix and the model-implied variance-covariance matrix, and is given as

$$F = F(S, \Sigma(\hat{\theta})) \tag{7}$$

It is important to note that the observed variables in this study are mainly categorical in nature and are not approximated well by normal distributions. According to Bollen (1989), the weighted least squares (WLS) estimator has the desirable property of making minimal assumptions about the distribution of observed variables, unlike maximum likelihood estimators (MLEs), which presuppose the underlying data to be approximately normally distributed. Hence, we used the WLS estimator instead of MLE for this research. The fitting function for WLS is

$$F_{WLS} = [s - \sigma(\theta)]' W^{-1}[s - \sigma(\theta)] \tag{8}$$

where

s is a $1/2(p + q)(p + q + 1)$ vector containing the polychoric and polyserial correlation coefficients for all pairs of latent endogenous and observed exogenous variables,

$\sigma(\theta)$ is the corresponding same-dimension vector for the implicated covariance matrix, and

W is a consistent estimator of the asymptotic covariance matrix of s.

Goodness-of-Fit Measures

To evaluate the overall goodness of fit of the estimated models, χ^2 fit, the root mean square error of approximation (RMSEA), normed fit index (NFI), and Tucker-Lewis index (TLI) were calculated and used to guide final model selection. In this context, it is important to mention that goodness of fit in SEM is an unsettled topic for which many researchers have presented a variety of viewpoints and recommendations. While a detailed explanation of them is beyond the scope of this paper, a brief description about each measure of fit used in this paper is provided.

As described by Washington et al. (2003), a useful feature of discrepancy functions is that they can be used to test the null hypothesis $H_0 : \Sigma(\theta) = \Sigma$, and $(n - 1)$ times the discrepancy function evaluated at $\hat{\theta}$ is approximately χ^2 distributed. The degrees of freedom are $1/2(p + q)(p + q + 1) - t$, where p and q are as described previously, and t is the number of free parameters in θ. This χ^2 divided by model degrees of freedom has been suggested as a useful goodness-of-fit measure. However, in this context, the logic of significance testing is different from significance of coefficient testing in a regression equation (Bollen 1989). In the classical application, we hope to reject the null hypothesis, whereas in the SEM (for the χ^2 test) the null hypothesis assumes that the implied model is equal to the true model and we do not wish to reject it. As a result, a large χ^2 and small p value suggests model lack of fit.

Thus, a relatively large p value and small χ^2 is sought and corresponds to a good fit of the model-implied variance-covariance with the observed one.

Another class of goodness-of-fit measures available in SEM is based on the population discrepancy function as opposed to the sample discrepancy function, such as RMSEA. The RMSEA is obtained by taking the square root of the population-based discrepancy function divided by its degrees of freedom. Practical experience indicates that a value of RMSEA of about 0.05 or less indicates a good fit of the model.

The remaining two goodness-of-fit measures are based on comparisons with a baseline model. The NFI proposed by Bentler and Bonett (1980) indicates the level of improvement in the overall fit of the present model compared with the baseline model and is given as

$$NFI = 1 - \frac{F}{F_b} \tag{9}$$

where F and F_b are the discrepancy functions of the fitted and baseline models. The TLI is similar to this concept, but in addition to the discrepancy functions, the degrees of freedom associated with the fitted model (df), as well as the baseline model (df_b), are considered in calculating the index, thus a penalty can be imposed for larger models similar to an adjusted R^2 in regression. The index is given as

$$TLI = \frac{F_b/df_b - F/df}{F_b/df_b - 1}$$

More information on SEM estimation, goodness of fit, variable selection, specification, and interpretation can be found in Bollen (1989), Arminger et al. (1995), Hoyle (1995), Schumacker and Lomax (1995), Kline (1998), and Washington et al. (2003).

SEM MODEL STARTER SPECIFICATION

The hypotheses mentioned previously, and described in greater detail here, helped to provide the research team with an initial SEM specification.

Hypothesis 1. The breadth and depth of statewide programs should in theory influence statewide safety, albeit with a time lag. It is hypothesized that states with relatively poor safety records will have broad and intensive safety programs, in response to a needed safety improvement. This finding would reflect an appropriate allocation of federal funds for improving safety across states. Because changes in statewide safety typically are not immediate, and because program benefits tend to lag program investments, it is assumed that depth and breadth of safety programming will be negatively associated with statewide safety. This anticipated relationship is aggregate in nature and exceptions may occur.

Hypothesis 2. GHSAs that perceive benefits from participating in the transportation planning process will be more likely to adopt a coordinated approach to safety and will, consequently, be more likely to identify new opportunities for addressing safety, thereby yielding increased safety performance. Understanding how these agencies develop their perceptions of planning activities is difficult to assess. For instance, an unfavorable attitude could represent an institutional unwillingness to coordinate with other agencies or may be the result of a previously unsuccessful attempt at coordinating with other agencies. Alternatively, a positive perception of planning may be the result of successful experiences in previous coordination attempts or may simply represent an appreciation for the potential benefits of a coordinated approach. It could also represent a latent agency willingness to adopt new and innovative approaches to transportation safety, even if the agency has never previously attempted to coordinate their efforts with planning entities. It is hypothesized that in aggregate, agencies with a positive perception of planning are expected to be willing to introduce a broad range of programs addressing safety, as well as yield better than average safety records.

Although these hypotheses represent a priori beliefs about the relationships between latent variables and guided the specification of the structural relationships in the SEM models, the latent factors were identified through an EFA. The survey response—observed endogenous variables (endogenous because they are influenced by the underlying latent variables)—form the x and y vectors in equation (1) and may be related to one or more latent variables. The factor loadings give an indication of how many distinctly different "dimensions" exist in the data.

The exploratory factor analysis output (table 2) clearly shows significant loadings for some variables

TABLE 2 Results of the Exploratory Factor Analysis

Observed variables	Factor 1	Factor 2	Factor 3
PEDEDU9	0.584	–0.197	0.348
PEDS2	0.893	0.052	0.101
SDWLK13	0.518	0.034	0.735
CRSWLK13	0.408	0.035	0.643
MPO14	–0.037	–0.052	0.883
BIKE2	0.708	0.073	0.155
BIKEDU10	0.783	–0.086	0.332
TOLFTVMT	–0.081	0.864	–0.037
ALCFTVMT	–0.062	0.868	0.035
PEDFTPOP	0.048	0.525	–0.052
BIKFTPOP	0.102	0.491	0.162

on a specific factor compared with others. For example, the variable PEDS2 loads significantly on the first factor but not on the second and third factors. These differences in factor loadings help the analyst to determine which observed variables are influenced by a common underlying factor, how many latent variables to consider in a SEM, and how to specify the measurement portion of a SEM (e.g., which observed endogenous variables measure the latent variables). The EFA in this study produced strong evidence in support of three latent variables, which are described as:

1. *Breadth of the programs* (*PROGRAM* in figure 1). This latent variable reflects the extent of programs reported by GHSAs around the United States. The questions that loaded heavily (were influenced by breadth of programs) on this factor included Q2a–Q2e, Q9a–Q9c, and Q10a–Q10d, survey responses that describe agency goals for pedestrian and bike safety programs, education, and enforcement.
2. *Attitude of the agencies toward planning* (*ATTITUDE_PLN* in figure 1). This latent variable influences collectively the responses to questions Q13a–Q13g, Q14a, Q15a, and Q16a, which represent attitudinal and perception-related questions.
3. *Risk*. This latent variable is strongly associated with variables such as total and alcohol-related crash rates and pedestrian and bicycle-related crash rates. Thus, the variable reflects the amount of motor vehicle-related risk across the states. It is worthwhile to note that the observed crash rates measure the degree that something is not safe, and thus the latent variable is labeled *RISK*.

RESULTS

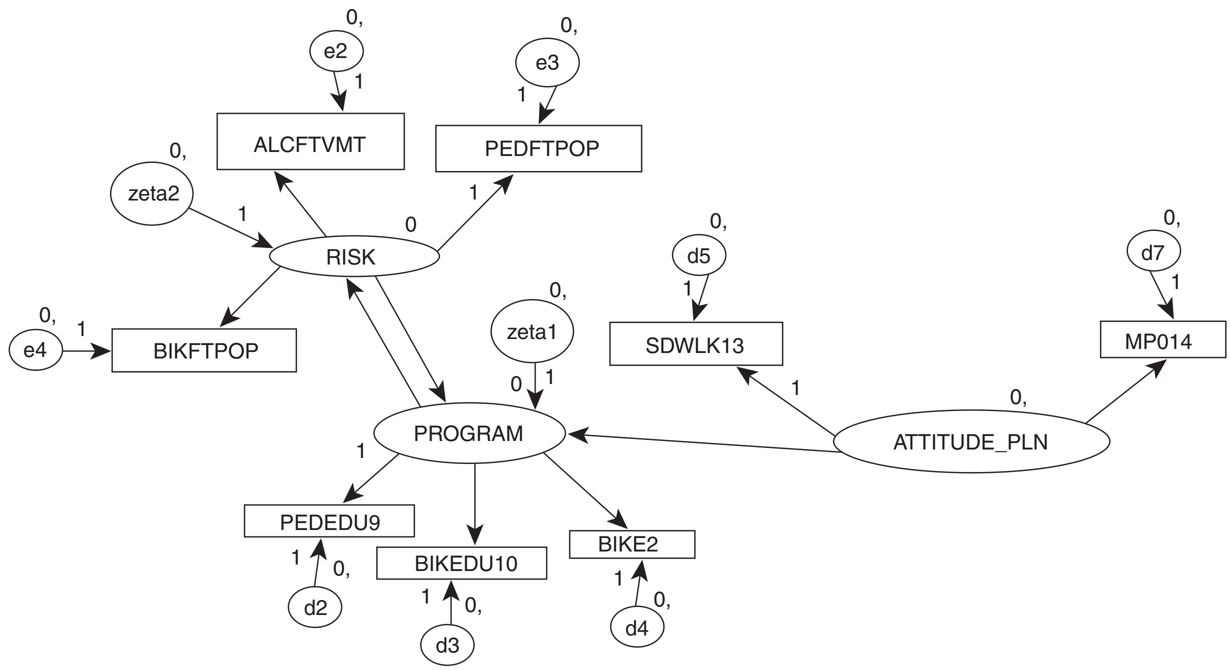
Using *Mplus* software, we obtained the results of the SEM estimated on the survey and statewide crash data. The relationships among the observed and latent variables are shown in figure 1. For ease of understanding, the observed endogenous variables are shown as rectangles, while both latent endogenous and exogenous variables appear as ellipses. The circles represent the unobserved measurement error terms. Arrows in the diagram suggest the direction of influence, thereby identifying the endogenous and exogenous variables. For example, the base of an arrow is attached to an exogenous variable while the variable to which it points is endogenous. A similar concept is also valid for the error terms. A 0 next to a variable indicates that its mean is set to 0 and the 1 near the arrow means the regression weight is given as 1. There must also be at least two observed endogenous variables pointing to a latent variable for identifiability of the model (a necessary condition for ensuring that sufficient information exists to estimate the model parameters).

Overall Model Goodness-of-Fit

As discussed previously, the hypothesized relationships between variables were used to identify structural relationships, while survey questions and state-level crash data were used to formulate the measurement portion of the model. The latent variables in the starter specification model were loaded with all the significant variables obtained from the exploratory factor analysis based on the factor loadings. A variable with a factor loading of greater than 0.5 was considered significant and included in the model. However, including all of the significant observed variables in the measurement model resulted in non-identifiability of the model (too many model parameters relative to the data to estimate).

As a result, an iterative process was adopted where the “least” significant variables were removed and an improved SEM model was obtained based on

FIGURE 1 Final Model Specifications



the model fit and convergence criteria. Note that changes during this process were made only to the measurement model and not to the structural model relating latent variables. For example, *TOLFTVMT,* which presents the total number of crashes per VMT, was dropped from the final model, because a better fitting model was achieved using three other accident-related variables explaining the latent variable *RISK*. Deleting *TOLFTVMT* from *RISK* was not detrimental to model fit, as it explains the total crash including pedestrian, bike, and alcohol-related accidents, and these were already taken into consideration by the other three observed variables. The problem of restricting the measurement models to a few select variables was dependent on sample size and model complexity and is addressed in the section on further research.

Table 3 shows the final SEM specification results. The interpretation of the results in the table is straightforward and consistent with the arrows-between-variables explanation of figure 1. The first row in table 3 shows that the latent variable *PROGRAM* acts as an exogenous variable on the latent variable *RISK* with a coefficient estimate of 0.026, standard error of 0.021, and t value of 1.226. The second part of table 3 presents the estimated intercepts of the observed endogenous variables.

Table 4 shows various goodness-of-fit statistics used during model selection. The chi-square value for the final model is 15.455 with 17 degrees of freedom (p value of 0.5627), which indicates the model fit cannot be rejected at $p = 0.05$. Because the chi-square goodness-of-fit test is used to test for differences between the implied model variance-covariance matrix and the observed one, a model that will not reject the null hypothesis is a desired outcome. The RMSEA for the final model was 0.003, which clearly indicates a close-fitting model. Also, the calculated NFI and TLI are close to 1, indicating considerable improvement over the baseline model.

General Findings

The model results suggest a two-way relationship between risk and the program efforts by GHSAs, or that these two aspects are mutually endogenous. Risk affects programming, and programming also affects risk. States with high safety risk actively implemented a wide variety of safety programs resulting in a breadth of the programs being positively associated with safety risk. This finding agreed with

TABLE 3 Weighted Least Squares Estimate of the Model

SEM regression weights	Regression weights	Standard error	Critical ratio
RISK <--------------- PROGRAM	0.026	0.021	1.226
PROGRAM <---------- RISK	7.081	6.278	1.128
PROGRAM <---- ATTITUDE_PLN	0.563	0.093	6.043
SDWLK13 <--------- ATTITUDE_PLN	1.926	0.806	2.391
MPO14 <----------- ATTITUDE_PLN	1.000	0.000	—
PEDEDU9 <-------------- PROGRAM	1.000	0.000	—
BIKE2 <------------- PROGRAM	1.169	0.102	11.463
BIKEDU10 <---------------- PROGRAM	1.163	0.119	9.769
ALCFTVMT <-------------- RISK	1.000	0.000	0.000
PEDFTPOP <-------------- RISK	7.168	3.753	1.910
BIKFTPOP <--------------- RISK	2.048	1.074	1.910
Intercepts	**Estimate**		
PEDEDU9	–0.370		
SDWLK13	0.431		
MPO14	0.311		
BIKE2	–0.253		
BIKEDU10	–0.493		
ALCFTVMT	0.663		
PEDFTPOP	1.480		
BIKFTPOP	0.219		

Key: SEM = structural equation model.

TABLE 4 Overall Goodness-of-Fit Measures for the Model

Description	Final model	(*p* value)
Chi-square	15.455	(0.5627)
RMSEA	0.003	
Normal fit index	0.999	
Tucker-Lewis index	0.999	

Key: RMSEA = root mean square error of approximation.

expectations that relatively higher safety risk drives the allocation of federal funding and thereby supports a broad range of safety programs targeted toward safety improvements. Also, the effect of risk on programming is significantly larger than the effect of programming on risk. This suggests there is a considerable lag between safety investments and risk reductions, or the combinations of safety programs do not bring about risk reductions proportionate to the effect of risk on programming. The squared multiple correlations (R^2 statistics) for the latent variables *RISK* and *PROGRAM* are 0.686 and 0.431, respectively, indicating that the model explains 68% of the variance in *PROGRAM* and 43% of the variance in *RISK*.

The latent variable *ATTITUDE_PLN*, which captures GHSAs' attitudes toward integrating state safety programs with transportation-related activities, as well as their participation in regional transportation planning, was positively associated with *PROGRAM*. The model also suggests that the attitude of GHSAs directly affects the latent variable *PROGRAM* and indirectly affects the exogenous variable *RISK* through *PROGRAM*. These findings imply that a positive attitude toward safety planning within GHSAs results in a broader and more extensive implementation of safety programs by GHSAs. This result again confirmed a priori expectations that agencies active in safety planning would be likely to implement a broad range of programs to improve statewide transportation safety.

CONCLUSIONS

ISTEA and the TEA-21 legislation raised the visibility of safety conscious planning in the United States. While safety-related planning has historically been reactive, new initiatives and investments are intended to change the status quo and encourage a new approach for considering safety at the transportation planning level. In particular, the NCHRP 8-44 project is oriented toward contributing and understanding tools for proactive planning among DOTs, MPOs, and GHSAs. This paper presents an exploratory analysis aimed at better understanding the relationships between GHSA-implemented safety programs and the actual safety scenario, as well as the effect of coordination among the various agencies on statewide safety.

The study revealed that coordination of transportation safety across DOTs, MPOs, GHSAs, and departments of public safety appears to be generally beneficial to safety, particularly in the long term. In addition, better safety planning coordination is characterized by positive attitudes toward incorporating safety into planning and also implementing a wide range of programs to improve safety. Furthermore, mechanisms for improving cooperation, coordination, and collaboration among the agencies also appear to be worthwhile investments.

FUTURE WORK

The results presented in this paper are exploratory due to a lack of a well-articulated theory regarding the subject matter. Additional information and some controlled data-collection would be required to draw more definitive conclusions. For example, panel data over a period of 5 to 10 years would be required to examine the lag between safety investments and risk, as well as attitudes and programs implemented over time. Additional data from all 50 states would sufficiently increase the number of observations necessary to improve the WLS estimates obtained in this analysis, allowing for more "complex" models. The WLS estimator requires at least $1/2(p + q)(p + q + 1)$ observations where $(p + q)$ are the number of observed dependent variables. Hence, the explanatory power of the model greatly depends on the number of observations.

This study was restricted to three main latent variables and a few carefully selected observed endogenous variables in the measurement model, due partly to data limitations. A larger study would enable the research team to focus on perhaps other relevant variables critical to statewide safety performance, such as the types of programs implemented. As a result, the present model remains speculative and further data are needed to validate it. Furthermore, the results are time dependent, due to the observed response from the 2002 to 2003 time periods. As mentioned previously, there is every possibility of a lagged effect of safety improvement programs on state safety performance, which is not properly captured in this modeling effort. Regardless, some initial insights into relationships among agencies were found and are encouraging for the successful implementation of ISTEA and TEA21 legislation targeted towards national safety improvements.

ACKNOWLEDGMENTS

The authors would like to acknowledge the National Cooperative Highway Research Program (NCHRP 8-44) for providing the funding for this research. This paper represents interim preliminary findings that have not been reviewed nor are endorsed by the NCHRP, its staff, or the panel overseeing this research effort.

REFERENCES

Amato, P.R. and B. Alan. 1995. Changes in Gender Role Attitudes and Perceived Marital Quality. *American Sociological Review* 60(1):58–66.

Arminger, G., C. Clogg, and M. Sobel. 1995. *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York, NY: Plenum Press.

Ben-Akiva, M., F. Humplick, S. Madanat, and R. Ramaswamy. 1991. Latent Performance Approach to Infrastructure Management. *Transportation Research Record* 1311:188–195.

Bentler, P.M. and D.G. Bonett. 1980. Significance Tests and Goodness of Fit in the Analysis of Covariance Structures. *Psychological Bulletin* 88:588–606.

Bollen, K.A. 1989. *Structural Equations with Latent Variables*. New York, NY: Wiley.

Everitt, B.S. 1984. *An Introduction to Latent Variables*. London, England: Chapman & Hall.

Golob, T.F. and A. Regan. 2000. Freight Industry Attitude Towards Policies to Reduce Congestion. *Transportation Research E* 36(1):55–77.

Hoyle, R., ed. 1995. *Structural Equation Modeling: Concepts, Issues, and Applications*. Thousand Oaks, CA: Sage Publications, Inc.

Insurance Institute for Highway Safety (IIHS). 2004. Available at http://www.hwysafety.org/safety_facts/safety.htm.

Johnson, R. and D. Wichern. 2002. *Applied Multivariate Statistical Analysis*. Upper Saddle River, NJ: Prentice Hall.

Kline, R. 1998. *Principles and Practice of Structural Equation Modeling*. New York, NY: Guilford Press.

Loehlin, J.C. 2004. *Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Manly, B. 1986. *Multivariate Statistical Methods: A Primer.* New York, NY: Chapman & Hall.

Molenaar, K., S. Washington, and J. Dickmann. 2000. Structural Equation Model of Construction Contract Dispute Potential. *Journal of Construction Engineering and Management,* July/August:1–10.

Östberg, M. and B. Hagekull. 2000. A Structural Modeling Approach to the Understanding of Parenting Stress. *Journal of Clinical Child Psychology* 29(4):615–625.

Rubio, D.M., M.B. Weger, and S.S. Tebb. 2001. Using Structural Equation Modeling to Test for Multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal* 8(4):613–626.

Schumacker, R.E. and R.G. Lomax. 1995. *A Beginner's Guide to Structural Equation Modeling*. Mahway, NJ: Lawrence Erlbaum Associates.

U.S. Department of Transportation (USDOT), National Highway Traffic Safety Administration. 2001. *Traffic Safety Facts 2001*. Washington, DC.

Washington, S., M. Karlaftis, and F. Mannering. 2003. *Statistical and Econometric Methods for Transportation Data Analysis*. Boca Raton, FL: Chapman & Hall/CRC.

APPENDIX

Survey Questionnaire

Survey Question 1: Is your agency located within, or directly affiliated with, another state agency, such as the Department of Transportation?

1a. affil1—Affiliated with another state agency?
- 1 = Yes
- 0 = No

1b. agency1—Name of affiliated agency.

1c. agcycod—Coding of agency affiliation.
- 0 = No affiliation
- 1 = State DOT
- 2 = State police
- 3 = Department of Public Safety
- 4 = Department of Motor Vehicles
- 5 = Other affiliation

(*survey questionnaire continues on next page*)

Survey Question 2: Do your agency's mission statement, goals, or objectives explicitly address any of the following issues?

2a. peds2—Pedestrian safety.
2b. bikes2—Bicycle safety.
2c. drived2—Driver education.
2d. schooled2—Safety education in school.
2e. enforce2—Traffic law enforcement.
2f. coopdot2—Cooperation with the state DOT.
2g. cooploc2—Cooperation with local officials.
2h. coopplan2—Interaction with regional or local transportation planners.
2i. safdes2—Incorporating safety into the design of transportation facilities.
2j. safops2—Incorporating safety into transportation facility operation.
2k. data2—Collecting safety-related data.
1 = Yes
0 = No

Survey Question 3: How many professional staff members (i.e., those focused on highway safety, not clerical or support staff) does your agency have?

Staff3—Number of agency employees.

Survey Question 4: Do any members of your staff have expertise in the following areas?

4a. Eng4—Transportation engineering.
4b. plan4—Transportation planning.
4c. ops4—Traffic operations.
4d. enf4—Law enforcement.
4e. edu4—Education.
4f. mktg4—Marketing/media relations.
1 = Yes
0 = No

Survey Question 5: Federal regulations require Governors Highway Safety agencies to develop annual plans, as well as performance measures for evaluating program effectiveness. How important are these performance measures in influencing the types of projects and programs implemented by your organization?

perfms5—Importance of annual performance measures on agency projects.
5 = Very important
4 = Somewhat important
3 = No opinion
2 = Not very important
1 = Not at all important

Survey Question 6: Are your agency's performance measures shared by other agencies responsible for the transportation system (e.g., by the state DOT or by regional transportation planning agencies in your state)?

6a. Sharepm6—Other agencies sharing performance measures.
1 = Yes
0 = No
6b. agency6—Names of agencies sharing performance measures, if any.

Survey Question 7: Does your agency develop longer term performance targets beyond the federally required 1-year targets?

7a. perftgt7—Performance targets beyond federal requirements.

1 = Yes

0 = No

7b. tgtyear7—Furthest target year in future.

Survey Question 8: What is the planning time horizon for your agency?

Agyhzn8—Agency planning horizon.

Survey Question 9: Which of the following pedestrian-related safety programs does your agency implement?

9a. Pededu9—Education on safe street crossing.

9b. pedcrwk9—Crosswalk enforcement.

9c. pedschl9—Safe routes to schools program.

1 = Yes

0 = No

9d. other9—Other pedestrian programs, if any.

Survey Question 10: Which of the following bicycle-related safety programs does your agency implement?

10a. Bikedu10—Bicycle education campaigns.

10b. bkhelm10—Bicycle helmet programs.

10c. bklite10—Lights on bicycles at night.

10d. bkbrk10—Bicycle brake requirements.

1 = Yes

0 = No

10e. other10—Other bicycle programs, if any.

Survey Question 11: Has your agency undertaken any innovative safety programs using federal flexible funds, such as Section 407 funds that provide flexible incentive grants for programs aimed at increasing highway safety?

11a. Innov11—Innovative programs using flexible funds.

1 = Yes

0 = No

11b. prgrm111—Name of innovative program 1, if any.

11c. effct111—Effectiveness of program.

5 = Very effective

4 = Somewhat effective

3 = No opinion

2 = Somewhat effective

1 = Not effective

11d. prgrm211—Name of innovative program 2, if any.

11e. effct211—Effectiveness of program 2.

5 = Very effective

4 = Somewhat effective

3 = No opinion

2 = Somewhat effective

1 = Not effective

(*survey questionnaire continues on next page*)

Survey Question 11 *(continued)*: Has your agency undertaken any innovative safety programs using federal flexible funds, such as Section 407 funds that provide flexible incentive grants for programs aimed at increasing highway safety?

11f. prgrm311—Name of innovative program 3, if any

11g. effct311—Effectiveness of program 3.

5 = Very effective
4 = Somewhat effective
3 = No opinion
2 = Somewhat effective
1 = Not effective

11h. prgrm411—Name of innovative program 4, if any.

11i. effct411—Effectiveness of program 4.

5 = Very effective
4 = Somewhat effective
3 = No opinion
2 = Somewhat effective
1 = Not effective

11j. prgrm511—Name of innovative program 5, if any.

11k. effct511—Effectiveness of program 5.

5 = Very effective
4 = Somewhat effective
3 = No opinion
2 = Somewhat effective
1 = Not effective

Survey Question 12: This study seeks to understand how often your agency interacts with other individuals, agencies, or groups that may have an influence on highway safety issues. For each of the following, please indicate whether your agency interacts with them monthly, once every three months, once every six months, once per year, or not at all.

12a. Dot12—Frequency of interaction with the state DOT.

12b. spolce12—Frequency of interaction with the state police.

12c. lpolce12—Frequency of interaction with the local police.

12d. plan12—Frequency of interaction with MPOs.

12e. legs12—Frequency of interaction with state legislators.

12f. gov12—Frequency of interaction with the governor's staff.

12g. hwy12—Frequency of interaction with highway contractors.

12h. engnr12—Frequency of interaction with engineering consultants.

12i. school12—Frequency of interaction with school officials.

12j. local12—Frequency of interaction with local officials.

4 = At least monthly
3 = Once every three months
2 = Once every six months
1 = Once per year
0 = Never

Survey Question 13: Has your agency *considered* integrating state safety programs with any of the following transportation-related activities:

13a. Sdwlk13—Considered sidewalk provisions.

13b. crswlk13—Considered crosswalk signals.

13c. bike13—Considered bike signals.

13d. speed13—Considered design strategies to reduce speeding.

13e. turn13—Considered design strategies to prevent turning movements.

13f. rdside13—Considered eliminating roadside hazards.

13g. monitr13—Considered using monitoring systems.

1 = Yes

0 = No

Survey Question 14: Every urbanized area in a state must have a comprehensive regional transportation planning process. For such areas in your state, has your agency participated in the regional transportation planning process during the last 5 years?

14a. mpo14—Participated in regional transportation planning during the last 5 years.

1 = Yes

0 = No

If mpo14 = 0

14b. blrp14—Benefited from participating in MPO long-range planning process (if no 14mpo).

1 = Yes

0 = No

14c. bgopm14—Benefited from developing MPO goals, objectives, and performance measures (if no 14mpo).

1 = Yes

0 = No

14d. bpep14—Benefited from participating in MPO project evaluation and programming (if no 14mpo).

1 = Yes

0 = No

Survey Question 15: Many states have regional planning agencies that represent rural, non-urbanized portions of a state. Has your agency participated in the transportation planning process for rural areas during the last 5 years?

15a. Rural15—Participated in rural planning process during the last 5 years.

1 = Yes

0 = No

If rural15 = 0

15b. blrp15—Benefited from participating in rural long-range planning process (if no 15rural).

15c. bgopm15—Benefited from developing rural goals, objectives, and performance measures (if no 15rural).

15d. bpep15—Benefited from participating in rural project evaluation and programming (if no 15rural).

1 = Yes

0 = No

(*survey questionnaire continues on next page*)

Survey Question 16: Every state has a statewide transportation planning process that, at a minimum, is responsible for producing a state transportation plan. Has your agency participated in the statewide transportation planning process during the last 5 years?

16a. 16stp—Participated in state transportation planning process during the last 5 years.

1 = Yes

0 = No

If stp16 = 0

16b. blrp16—Benefited from participating in state long-range planning process (if no 16stp).

16c. bgopm16—Benefited from developing state goals, objectives and performance measures (if no 16stp).

16d. bpep16—Benefited from participating in state project evaluation and programming (if no 16stp).

1 = Yes

0 = No

Survey Question 17: To what extent does the transportation planning process in your state influence the programs or initiatives undertaken by your agency?

Influ17—Influence of transportation planning process on highway safety programs.

2 = Strongly influences

1 = Moderately influences

0 = Does not influence

9 = Don't know

Vehicle Breakdown Duration Modeling

WENQUN WANG[1,*]
HAIBO CHEN[2]
MARGARET C. BELL[2]

[1] Atkins Transport Systems
Woodcote Grove Ashley Road
Epsom, England KT18 5BW

[2] Institute for Transport Studies
University of Leeds
Leeds, England LS2 9JT

ABSTRACT

This paper analyzes the characteristics of vehicle breakdown duration and the relationship between the duration and vehicle type, time, location, and reporting mechanisms. Two models, one based on fuzzy logic (FL) and the other on artificial neural networks (ANN), were developed to predict the vehicle breakdown duration. One advantage of these methods is that few inputs are needed in the modeling. Moreover, the distribution of the duration does not affect the results of the prediction. Predictions were compared with the actual breakdown durations demonstrating that the ANN model performs better than the FL model. In addition, the paper advocates for a standard way to collect data to improve the accuracy of duration prediction.

INTRODUCTION

A traffic incident is a nonrecurrent event. It is not a planned closure of a road nor a special event; therefore, there is no advanced notice. Examples include vehicle breakdowns, accidents, natural disasters, and those caused by humans. An accident is a specific type of incident that normally involves human injury or casualty.

Email addresses:
* Corresponding author—wenqun.wang@atkinsglobal.com
H. Chen—hchen@its.leeds.ac.uk
M.C. Bell—mbell@its.leeds.ac.uk

KEYWORDS: Traffic incident management, vehicle breakdown duration, fuzzy logic, neural networks.

Incidents have become one of the main causes of traffic congestion. Lindley (1987) showed that between 50% and 75% of the total traffic congestion on urban motorways in the United States is incident-induced. Moreover, there is a symbiotic relationship between incidents and congestion. As incidents cause more congestion, more congestion brings with it more incidents. Traffic incidents have other impacts: the risk of secondary crashes for other road users and those dealing with the incident; and possible reductions in air quality due to increased fuel consumption caused by the congestion.

In recent years, investment in developing systems to manage incidents has increased. The Federal Highway Administration defines incident management as the systematic, planned, and coordinated use of human, institutional, mechanical, and technical resources to reduce the duration and the impact of incidents, and improve the safety of motorists, crash victims, and incident responders (USDOT 2000). Therefore, incident duration prediction becomes an important tool for incident management. Reliable duration prediction can help traffic managers apply appropriate management strategies, and it can also be used to evaluate the efficiency of the management strategies that are implemented. Furthermore, duration prediction can provide accurate and essential information to road users.

Vehicle breakdown is one type of incident that often occurs on motorways and represents more than 80% of all types of incidents. In this paper, we analyze the characteristics of vehicle breakdowns and develop vehicle breakdown duration models based on fuzzy logic (FL) and artificial neural networks (ANN). We use incident data collected from the M4 motorway in the United Kingdom to validate our models.

LITERATURE REVIEW

Incident duration is the time period between the occurrence and clearance of an incident. During this period, the following activities occur: incident detection, verification, response, clearance, and recovery. Components of incident management include traffic management and traffic information. To accomplish this, information is exchanged between the different parties involved, including the police and the breakdown recovery service.

Golob et al. (1987) analyzed data from over 9,000 accidents involving large trucks and combination vehicles collected over a two-year period on freeways in the greater Log Angeles area. They found that accident duration fitted a log-normal distribution. The factors used in their accident duration model were collision type, accident severity, and lane closures. Their data were shown to be more statistically significantly similar to the log-normal than the log-uniform distribution. However, the sample size of each group was small (between 21 and 57).

Giuliano (1989) extended the research of Golob et al. by applying a log-normal distribution in the incident duration analysis of 512 incidents in Los Angeles. The author found that the factors affecting incident duration were incident type, lanes closed, time of day, accident type, and whether or not a truck was involved. The variance within each category was large making it difficult to forecast the incident duration.

Jones et al. (1991) made further improvements by imposing a conditional probability; that is, given that the incident has lasted *X* minutes, it will end in the *Y*th minute. The authors analyzed 2,156 incidents in the metropolitan Seattle area and found that the duration of incidents conformed to a log-logistic instead of log-normal distribution (they applied a hazard duration model to estimate the incident duration). However, some factors used in their model, such as the age of the driver, were found to be impractical, because this information was often not available when the incident occurred. They stated that more appropriate and accurate data are very important in incident duration analysis.

Nam and Mannering (2000) further developed the hazard duration model in an analysis of incident duration. They analyzed 681 incidents in Washington state, collected over two years. They continued to use the log-logistic model of Jones et al. (1991) but removed the impractical variables and applied hazard-based functions to estimate the incident duration. This study provided evidence that hazard-based approaches are suited to incident analysis for the individual stage of the incident, including detection time, response time, and clearance time.

However, one drawback, highlighted by the authors, is that they could not draw definitive conclusions concerning the actual duration of the incident because data were insufficient.

Sethi et al. (1994) developed a decision tree to predict incident duration. They based their research on the statistical analysis of 801 incidents from the Northwest Central Dispatch. This prediction method was very easy and practical to use; however, all the unknown incident durations were set to 23 minutes, and this oversimplification of the model was detrimental to the accuracy of the predictions.

Other papers that present complementary statistical analyses of incident duration include Wang (1991), Sullivan (1997), Cohen and Nouveliere (1997), Garib et al. (1997), Smith and Smith (2000), and Fu and Hellinga (2002).

FL has been used in the transportation field since the theory was first developed by Zadeh (1965). The method offers much potential in the traffic and transport field, because many problems and parameters are characterized by linguistic variables. Moreover, many problems in this field are ill defined, ambiguous, and vague. Such situations are difficult to model using traditional methods. A review by Teodorovic (1999) of state-of-the-art FL systems for transport engineering clearly showed the potential for the application of FL.

Choi (1996) was the first researcher to use an FL system to predict incident duration. He used incident data on vehicle problems, types of assistance, and the location of disabled vehicles to demonstrate the suitability of FL for solving problems characterized by elements of uncertainty and ambiguity. Moreover, the FL system was shown to perform well with fewer variables compared with the statistical models.

Kim and Choi (2001) updated the model and improved the performance by refining the fuzzy sets. However, the authors did not categorize the type of incidents, and this may have a significant effect on incident duration. Another shortcoming of this work is the limited incident data available to validate the model.

Wang et al. (2002) used FL to model vehicle breakdown duration by analyzing the characteristics of the breakdown by vehicle type, time of day, and location. Over 200 incident records from the M4 Motorway in the United Kingdom were used to demonstrate the credibility of the FL approach for estimating incident duration.

A number of studies have reported the increasing popularity of the application of the ANN theory to transportation. A review by Dougherty (1995) reported its wide application in a number of areas (e.g., traffic control, vehicle detection, driver behavior analysis, traffic pattern analysis, traffic forecasting, and parameter estimation). More recent applications include incident detection analysis by Teng and Qi (2003) and Yuan and Cheu (2003). The theory of ANN is presented later in this paper.

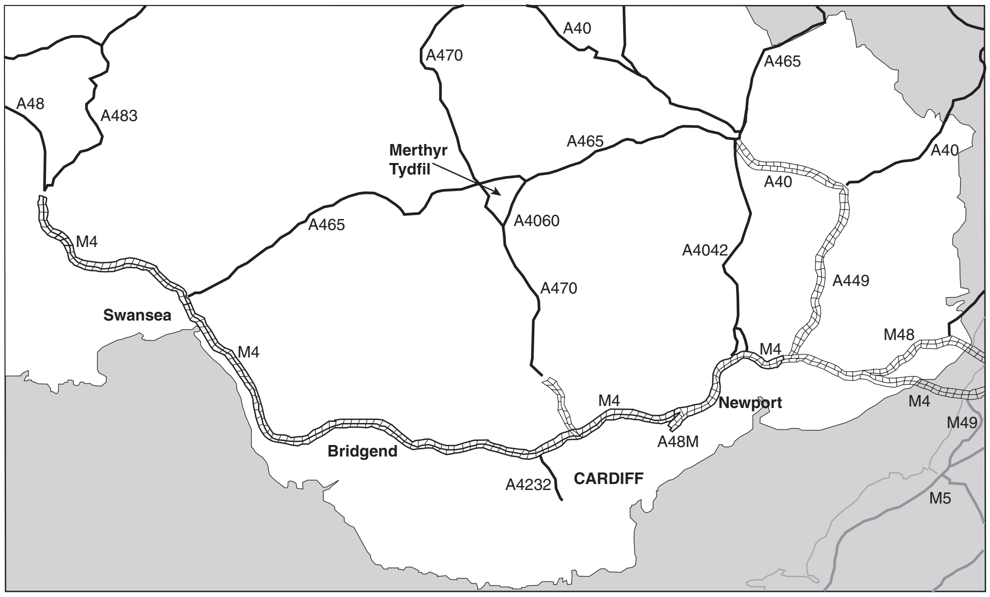
In summary, incident duration research has been developed gradually over the last decade. Various methods have been applied, including statistical analysis and fuzzy logic. However, comparing previous research results is difficult for a number of reasons: different variables have been used by the researchers; the data were collected from different areas in the world; and each dataset had its own characteristics. This review has provided us with the foundation on which we developed an alternative approach to model traffic incident duration using ANN. The results are presented here and are compared with those of an FL model, building on the earlier work of Wang et al. (2002).

DATA DESCRIPTION

For this research, the incident duration data were collected from one of the busiest roads in the United Kingdom, the M4 between Junction 22 and Junction 49. The average traffic flow on this section of the M4 was 65,000 vehicles per day, with a maximum flow of 102,000 vehicles per day in 2001 (Department for Transport 2002).

The MANTAIN CYMRU Traffic Management and Information Centre (TMIC), developed by a public/private partnership led by the National Assembly of Wales, provides a cost efficient method of improving traffic management. TMIC's responsibility includes 129 kilometers of motorway and parts of other trunk roads, as illustrated by figure 1. TMIC collects information using several media including: a closed circuit television system, traffic sensors, roadside meteorological systems, probe vehicles, police traffic reports, and other sources.

FIGURE 1 Map of the M4 Motorway



Source: Available at www.traffic-wales.com.

The Road Network Master Database (RNMD) stores all the information, which can be processed, transferred, and published to a third party as well as the public (James and Wainwright 2002).

We obtained 1,080 incidents records from RNMD for May 2000 to April 2001. The incidents were divided into three types: crashes, vehicle breakdowns, and other incidents. The majority of incidents were vehicle breakdowns, 64% of all the traffic incidents on the motorways. Crashes and other incidents made up the remainder, 20% and 16% of all incidents, respectively.

This paper reports the results of 695 vehicle breakdowns. Many of the records were incomplete; that is, the end time of the incidents was often not recorded. An in-depth look at the data gave us 213 complete incident records, which we present in this paper.

Figure 2 shows the distribution of the incident duration. A Kologorov-Smirnov test shows that it conforms to a Weibull distribution (sig. = 0.432), instead of a log-normal distribution (sig. = 0.043), which is consistent with the research of Nam and Mannering (2000).

Figure 3 demonstrates that incident duration displays a relationship to the time of day and shows peaks during the morning and evening rush hour. The figure also shows that vehicle breakdown duration tends to be longer at night. These characteristics are consistent with the higher traffic flow that causes congestion during the day and the poorer quality of recovery service during the late evening and overnight when traffic flows are substantially lower.

FIGURE 2 Distribution of Vehicle Breakdown Duration

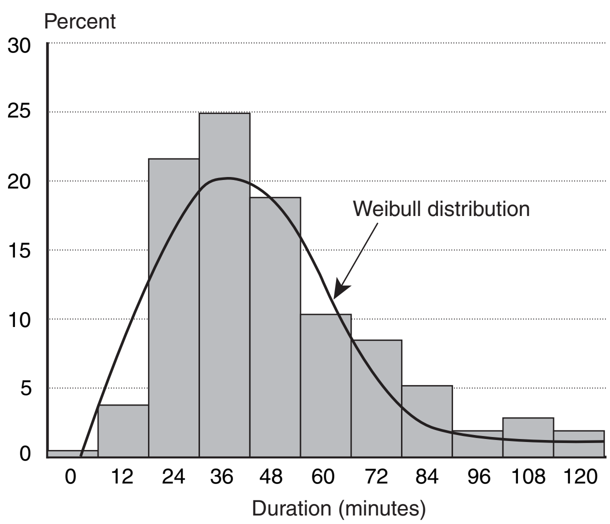


Figure 4 compares the arithmetic and geometric means of the vehicle breakdown data according to vehicle type. As expected, the geometric mean is consistently smaller than the arithmetic mean for all vehicles, because most incidents are of short duration. So the distribution is skewed to the right. The

FIGURE 3 Vehicle Breakdown Duration and Time of Day

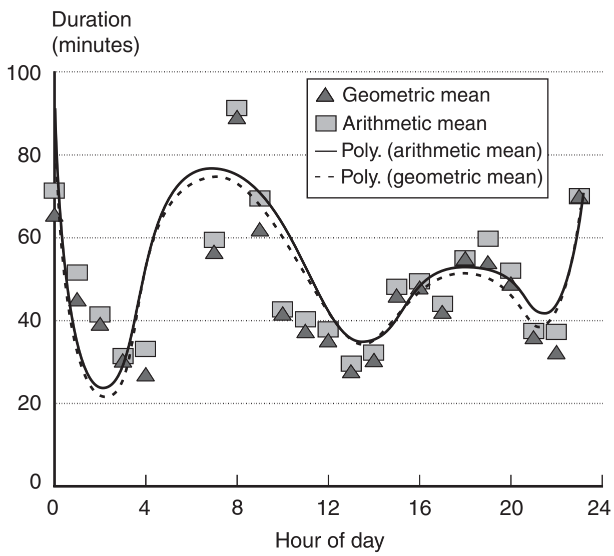
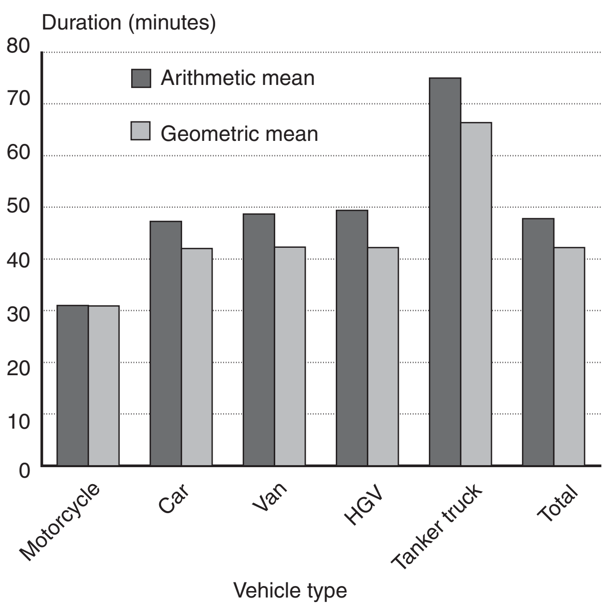


FIGURE 4 Vehicle Breakdown Duration and Vehicle Type



Key: HGV = heavy goods vehicle (over 3,500 kg or 7,716 lbs design, gross vehicle weight).

duration of a tanker breakdown is the greatest, which is not surprising. The latter interpretation, however, should be viewed with caution due to the small sample size for this type of incident.

Based on the available data and discussions with the operators in the traffic control center, the potential variables to be considered in the vehicle breakdown duration model were vehicle type, location,

TABLE 1 Kruskal-Wallis Test of Vehicle Breakdown Duration

Variable	Chi-square	df	Significance level
Report mechanism	16.7	1	0.44*10E–4
Vehicle type	17.1	3	0.67*10E–3
Location	8.5	2	0.014
Time of day	12.8	5	0.025

time of day, and report mechanism. We investigated the difference between the incident duration categories using the Kruskal-Wallis test, the results of which are shown in table 1. This test indicates that the overall differences between the categories are statistically significant, in particular the vehicle types and report mechanisms categories had a significance level of less than 0.001.

VEHICLE BREAKDOWN DURATION MODELING

In this section, we present two vehicle breakdown duration models. The first is based on FL, while the second uses the ANN approach.

The Model Based on FL

This research used the Mamdani-type FL system. Mamdani (1974) proposed this method in an attempt to control a steam engine and boiler combination by synthesizing a set of linguistic control rules obtained from experienced human operators. The Mamdani-type inference expects the output membership functions to be fuzzy sets. After the aggregation process, the fuzzy set for the output variable needs defuzzification. The inputs of the incident duration were vehicle type, location, time of day, and report mechanism. The dependent variable was the vehicle breakdown duration. These are detailed in table 2.

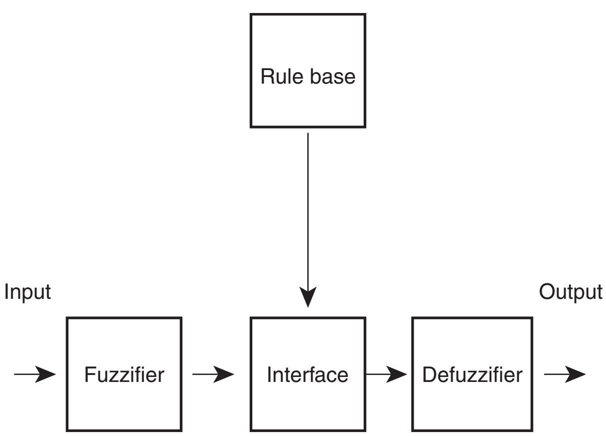
Figure 5 illustrates the structure of the FL system. The system comprises four elements: the fuzzifier that maps the crisp value into a fuzzy set; the rule base that saves the fuzzy rules; the interface that generates the fuzzy output from the input based on the fuzzy rules; and finally the defuzzifier that transfers

TABLE 2 Variables in the Fuzzy Logic Model

	Variable	Fuzzy set
Input	Vehicle type	Small
		Medium
		Big
		Very big
	Location	At node
		Close to node
		Far from node
	Time	Night
		Early morning
		Morning
		Afternoon
		Afternoon peak time
		Evening
	Report mechanism	ETS used
		No ETS used
Output	Duration	Very short
		Short
		Medium
		Long
		Very long

Key: ETS = emergency telephone service.

FIGURE 5 Structure of the FL System



the fuzzy output into a crisp value. The detailed explanation of the FL theory can be found in Pedrycz and Gomide (1998).

This research was based on the fuzzy sets of each variable, the characteristics of the data presented above, and the fuzzy rules derived from an understanding of the experience of the operators at the MANTAIN CYMRU TMIC gained in the interview surveys. One example of the 112 fuzzy rules used in this work is the following:

> If the vehicle is CAR, and location is AT THE JUNCTION, and the time is MORNING, and the report mechanism is ETS (emergency telephone service),
>
> then the vehicle breakdown duration is SHORT.

There are many defuzzification methods, including the *mean of the range of maximal values* and the *center of the area* that returns the center of gravity of the area under the curve. The latter is the most popular method used in defuzzification and the one adopted in this study.

Matlab was used to generate the model and simulate the results (Biran and Breiner 1999). Figure 6 shows the model surface depending on the vehicle type and time of day and clearly illustrates the nonlinear relationship between the inputs and outputs.

Figure 7 displays the value predicted using FL compared with the observed value. The model shows promise as an estimator of breakdown duration and the pattern of results is consistent with the research by Cohen and Nouveliere (1997).

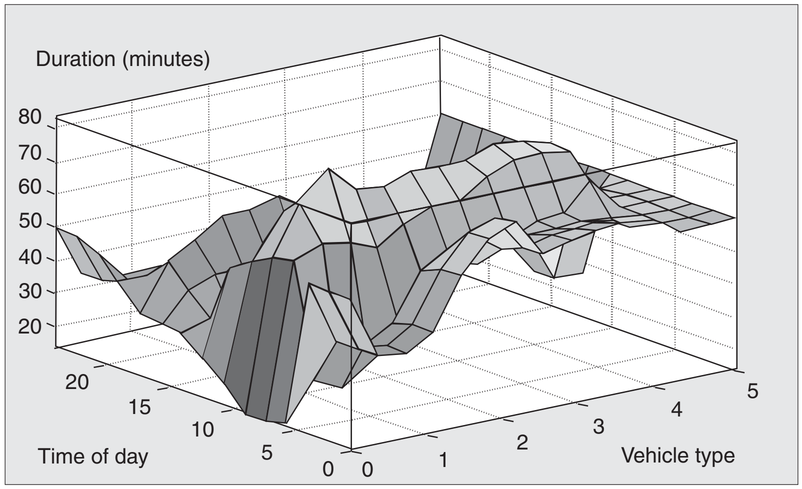
The Model Based on the ANN System

An ANN is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use (Aleksander and Morton 1990). The knowledge is acquired through a learning process and is stored as synaptic weights. The structure of the ANN is described later in this section.

The advantages of the ANN are as follows. First, the ANN is nonlinear, thus it can be applied to model a nonlinear physical mechanism easily. Second, the learning process enables the ANN to be modified, in accordance with an appropriate statistical criterion, to minimize the difference between the desired response and the actual response of the network driven by the input. This makes the ANN a suitable candidate to model incident duration.

In this research, a multilayer perceptron network was used, in which $IW\{n,1\}$ is the input weight matrix; $LW\{n,1\}$ is the layer weight matrix; and $b\{n\}$ are bias vectors, where n is the layer number. The

FIGURE 6 Surface of the FL System



choice of the neurons in the ANN is based on the number of inputs, outputs, and the sample size. The neuron number is determined following the guide by NeuralWare (1993). In this research, to maintain simplicity and avoid redundant architecture the ANN model has 17 neurons in one hidden layer (figure 8).

The output of each layer, which is the input for the next layer, is calculated using the following formula:

$$y_j = f\left(\sum_{i=0}^{N-1} w_{ij}\, x_i - \theta_j\right) \qquad 0 \le j \le N_1 - 1$$

where
y_j = the jth output,
θ_j = the bias in the nodes,
w_{ij} = the weight,
N = the number of inputs, and
N_1 = the neuron number.

In this research, the transfer function of the first hidden layer was the sigmoid function:

$$f(x) = \frac{1}{1 + e^{-x}}$$

In the output layer, a linear function is used as the transfer function to generate the desired output.

The inputs to the ANN system were vehicle type, location, time of day, and report mechanism. The output was vehicle breakdown duration. The ANN model was trained with the back-propagation training algorithm (Lau 1992), which is a generalization of the least mean squares algorithm. It uses a gradient search technique to minimize the mean square difference between the desired and the actual outputs.

FIGURE 7 Result of the FL Model

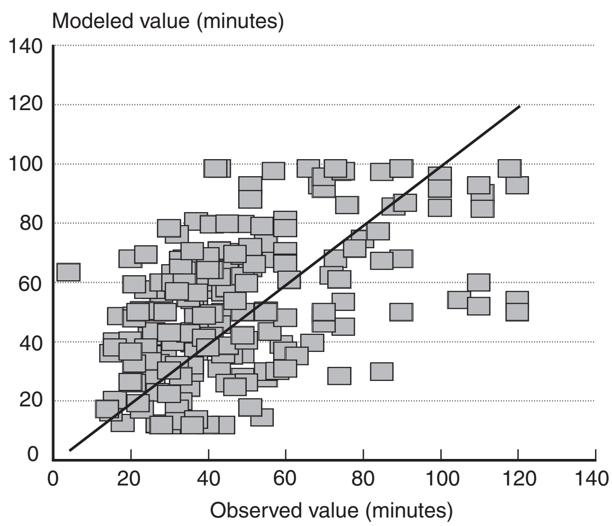


FIGURE 8 Structure of the ANN System

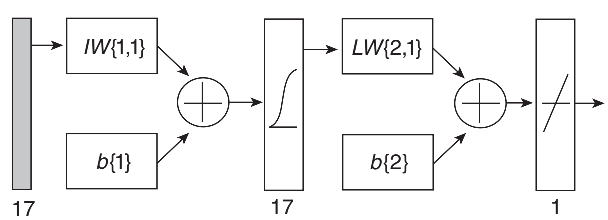
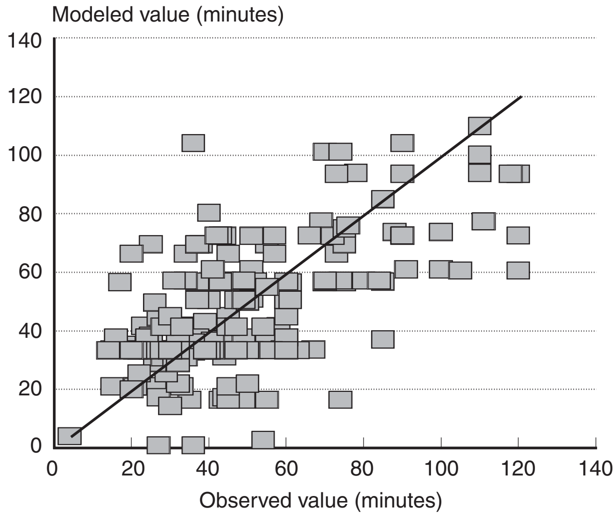


FIGURE 9 Result of the ANN Model



Of the 213 vehicle breakdown incidents, 113 incidents were used to train the model and 50 were used in validation during the training. The remaining 50 incidents were used to test the performance of the model. Figure 9 compares the ANN prediction with the observed value with encouraging results. It demonstrates that the performance was better than that of the FL model, because the predicted value was closer to the observed value. However, it also shows that the ANN model systematically generated the same durations when the observed values were different. In general, the ANN in this case failed to predict the larger values and outliers. One reason for this was that the number of explanatory variables was insufficient. Therefore, the ANN model could not be trained to perform well. This problem could be solved by including additional variables and is a subject of future research.

In order to estimate the influence of the input variables on the output of the model, we conducted a sensitivity test. This was achieved by excluding one input variable at a time and quantifying the deterioration of the performance of the prediction caused by the missing variable. The performance measure used was defined as the percentage change in the root mean square error (*RMSE*). The *RMSE* gives a measure of the difference between the observed and modeled value. It is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(f_n - v_n)^2}$$

where

f_n = the modeled value,

v_n = the observed value, and

N = the number of observations.

The percentage change of the error $P\%$ is given by

$$P\% = \frac{RMSE_{n-1} - RMSE_n}{RMSE_n} \times 100\%$$

where

$RMSE_n$ = the *RMSE* of the model with all n inputs.

The sensitivity test showed that all four variables influenced the performance of the ANN vehicle breakdown duration model, as the error consistently increased when each input was removed from the model. In particular, the report mechanism was found to have the greatest effect, because the error increased by 23% when it was removed from the model. The location had the least effect, with a 12% increase (table 3).

TABLE 3 Sensitivity Test of the ANN Input Variables

Model	*RMSE*	Error increase
ANN model with all four inputs	19.5	
ANN model without report mechanism	24.1	23%
ANN model without vehicle type	23.8	22%
ANN model without time of day	22.5	15%
ANN model without location	22.0	12%

Key: *RMSE* = root mean square error.

COMPARISON OF THE RESULTS OF FL AND ANN

We conducted statistical tests to compare the performance of these two models. In this paper, the R^2 test and the *RMSE* were applied. These methods are commonly used to evaluate the relative performance of traffic models (Clark et al. 2002).

The coefficient of variation R^2 is shown in table 4. We tested two ANN models, one with 17 neurons in the hidden layer and one with 10 neurons in the hidden layer, and found that the number of neurons affects the performance of the model. The table shows that the ANN model with 17 neurons

TABLE 4 Comparison of the FL and ANN Models

Model	R^2	Adjusted R^2	Sig.	*RMSE*
FL model	0.265	0.262	0.000	24.0
ANN model 1*	0.414	0.411	0.000	19.5
ANN model 2**	0.215	0.211	0.000	24.1

* ANN model 1 has 17 neurons in the hidden layer.
** ANN model 2 has 10 neurons in the hidden layer.
Key: *RMSE* = root mean square error.

performed best, while the performance of the FL model fell in the middle of the two ANN models. At the time that an incident occurs, the operator in the control center estimates the anticipated duration of the resulting congestion, based on engineering judgment and experience. The *RMSE* of this estimation is 42 minutes. It shows that both the ANN models and the FL model gave better estimates than the operators judgment.

Both ANN and FL methods show promise in predicting the incident duration. However, given that the R^2 value is not very high, and the *RMSE* value is large, the performance needs to be improved. This can be addressed by including more variables in the model. However, this requires more data to be collected and the cooperation of the operators and those responsible for motorway incident management. Future work will be concentrated in these areas. Despite the fact that the significance levels of the three models are low, the modeled values are consistently better than the estimated values by the operators. Therefore, these results are of interest to the motorway incident management team.

CONCLUSIONS

This paper analyzed the characteristics of vehicle breakdown duration and the main factors that may affect the duration. Two models, one based on FL and the other on ANN, were developed and their performances compared.

The research demonstrated that FL and ANN can provide reasonable estimates for the breakdown duration with few variables. They consistently outperform the existing method based solely on the engineering judgment of the operators. Also, for the specific data used in this research, the ANN model performed better than the FL model according to the characteristics of statistical parameters. However, both models had difficulties in predicting the outliers. As further data characterizing the outliers become available, the relative performance of ANN and FL may change.

Finally, the research highlights the need to collect information required for incident management in a standard way to improve the accuracy of prediction, enhance the management of incidents, and enable the authorities to share the data. Current research, using a specially designed electronic database tool, will improve the quantity and quality of the data records and thus begin to explain more of the variation in the data. In the future, the combined FL-ANN approach can be used to analyze incident duration, because this method can combine the experiences of the experts and the statistical characteristics of ANN.

ACKNOWLEDGMENT

The authors would like to thank W.S. Atkins for supplying the traffic data used in this study.

REFERENCES

Aleksander, I. and H. Morton. 1990. *An Introduction to Neural Computing*. London, England: Chapman & Hall.

Biran, A. and M. Breiner. 1999. *Matlab 5 for Engineers*. Reading, England: Addison-Wesley.

Choi, H.-K. 1996. Predicting Freeway Traffic Incident Duration in an Expert System Context Using Fuzzy Logic, Ph.D. dissertation. University of Southern California.

Clark, S.D., S.M. Grant-Muller, and H. Chen. 2002. Using Nonparametric Tests to Evaluate Traffic Forecasting Performance. *Journal of Transportation and Statistics* 5(1):47–56.

Cohen, S. and C. Nouveliere. 1997. Modelling Incident Duration on an Urban Expressway. *IFAC/IFIP/FORS Symposium*. Edited by M. Papageorgiou and A. Pouliezos. Chania, Greece.

Department for Transport. 2002. *Transport Statistics Bulletin: Road Traffic Statistics: 2001*. Compiled by A. Silvester and A. Smith. London, England.

Dougherty, M. 1995. A Review of Neural Networks Applied to Transport. *Transport Research C* 3:247–260.

Fu, L. and B. Hellinga. 2002. Real-Time, Adaptive Prediction of Incident Delay for Advanced Traffic Management Systems. *Proceedings of the Annual Conference of the Canadian Institute of Transportation Engineers,* May 12–15, 2002, Waterloo, Ontario, Canada.

Garib, A., A.E. Radwan, and H. Al-Deek. 1997. Estimating Magnitude and Duration of Incident Delays. *ASCE Journal of Transportation Engineering* 123:459–466.

Giuliano, G. 1989. Incident Characteristics, Frequency, and Duration on a High Volume Urban Freeway. *Transportation Research A* 23:387–396.

Golob, T.F., W.W. Recker, and J.D. Leonard. 1987. An Analysis of the Severity and Incident Duration of Truck-Involved Freeway Accidents. *Accident Analysis and Prevention* 19:375–395.

James, J.L. and A.K. Wainwright. 2002. Proactive Traffic Management in Wales. *11th IEE International Conference on Road Transport Information and Control*. London, England: IEE Press.

Jones, B., L. Janssen, and F. Mannering. 1991. Analysis of the Frequency and Duration of Freeway Accidents in Seattle. *Accident Analysis and Prevention* 32:239–255.

Lau, C. 1992. *Neural Networks—Theoretical Foundations and Analysis*. New York, NY: IEEE Press.

Lindley, J.A. 1987. Urban Freeway Congestion: Quantification of the Problem and Effectiveness of Potential Solutions. *ITE Journal* 57:27–32.

Kim, H.J. and H.-K. Choi. 2001. A Comparative Analysis of Incident Service Time on Urban Freeways. *Journal of the International Association of Traffic and Safety Sciences* 25:62–72.

Mamdani, E.H. 1974. Application of Fuzzy Algorithms for Control of Simple Dynamic Plant. *Proceedings of IEEE, Control and Science* 121(12):1585–1588.

Nam, D. and F. Mannering. 2000. An Exploratory Hazard-Based Analysis of Highway Incident Duration. *Transportation Research A* 34:85–102.

NeuralWare, Inc. 1993. *Neural Computing*. Carnegie, PA.

Pedrycz, W. and F. Gomide. 1998. *An Introduction to Fuzzy Sets: Analysis and Design*. Cambridge, MA: MIT Press.

Sethi, V., F.S. Koppelman, C.P. Flannery, N. Bhomderi, and L. Schofer. 1994. Duration and Travel Time Impacts of Incidents—ADVANCE Project Technical Report. Northwestern University, Evanston, IL. Quoted in K. Ozbay, and P. Kachroo. 1999. *Incident Management in Intelligent Transportation Systems*. Boston, MA: Artech House.

Smith, B.L. and K.W. Smith. 2000. *An Investigation into Incident Duration Forecasting for Fleetforward*. Charlottesville, VA: University of Virginia.

Sullivan, E.C. 1997. New Model for Predicting Freeway Incidents and Incident Delays. *ASCE Journal of Transportation Engineering* 123:267–275.

Teng, H. and Y. Qi. 2003. Detection-Delay-Based Freeway Incident Detection Algorithms. *Transport Research C* 11:265–287.

Teodorovic, D. 1999. Fuzzy Logic Systems for Transportation Engineering: The State of the Art. *Transportation Research A* 33:337–364.

U.S. Department of Transportation (USDOT), Federal Highway Administration. 2000. *Traffic Incident Management Handbook*. Washington, DC.

Wang, M. 1991. Modelling Freeway Incident Clearance Time, M.S. thesis. Northwestern University, Evanston, IL. Quoted in K. Ozbay, and P. Kachroo. 1999. *Incident Management in Intelligent Transportation Systems*. Boston, MA: Artech House.

Wang W., H. Chen, and M.C. Bell. 2002. A Study of Characteristics of Motorway Vehicle Breakdown Duration. *11th IEE International Conference on Road Transport Information and Control*. London, England: IEE Press.

Yuan, F. and R.L. Cheu. 2003. Incident Detection Using Support Vector Machines. *Transport Research C* 11:309–328.

Zadeh, L. 1965. Fuzzy Sets. *Information and Control* 8:338–353.

Using Generalized Estimating Equations to Account for Correlation in Route Choice Models

MOHAMED ABDEL- ATY

Department of Civil and Environmental Engineering
University of Central Florida
Orlando, FL 32816-2450
Email: mabdel@mail.ucf.edu

ABSTRACT

This paper presents the use of binary and multinomial generalized estimating equation techniques (BGEE and MGEE) for modeling route choice. The modeling results showed significant effects on route choice for travel time, traffic information, weather, number of roadway links, and driver age and education level, among other factors. Each model was developed with and without a covariance structure of the correlated choices. The effect of correlation was found to be statistically significant in both models, which highlights the importance of accounting for correlation in route choice models that may lead to vastly different travel forecasts and policy decisions.

INTRODUCTION

How and when travelers make decisions about what route they will take to their destination is an area of great interest to researchers and decision-makers alike. In this paper, binary and multinomial generalized estimating equation techniques (BGEE and MGEE) are used to model route choice.

Binary and multinomial route choice models may have two different kinds of correlation. First, repeated observations may be correlated. This is

KEYWORDS: Route choice, repeated observations, overlapping routes, BGEE, MGEE, logit.

usually the case for studies that use surveys/simulations where each respondent/subject provides repeated responses. Second, the overlapping distance between alternative routes may be correlated in multinomial route choice models. In a multinomial route choice model, the case is further complicated when the data structure includes both types of correlation.

In the 1980s, most discrete choice models were calibrated using binary logit (BL) and multinomial logit (MNL) models (Yai et al. 1997). BL and MNL models characterize the choice of dichotomous or polytomous alternatives, made by a decisionmaker (in our study, the driver) as a function of attributes associated with each alternative as well as the characteristics of the individual making the choice.

An advantage of both BL and MNL models is their analytical tractability and ease of estimation. However, a major restriction of MNL models is the Independence from Irrelevant Alternatives (IIA) property, which arises because all observations are assumed to have the same error distribution in the utility term based on a Gumbel distribution (IIA arises because the assumption of being Independent and Identically Distributed is made for the Gumbel). Therefore, BL and MNL models assume independence between observations, which is not true if each subject/driver has more than one observation. Also, MNL models assume independence between alternatives, which is not true when routes overlap.

A major statistical problem with cluster-correlated data, for which BL/MNL models do not account, arises from intracluster correlation or the potential for cluster mates to respond similarly. This phenomenon is often referred to as overdispersion or extra variation in an estimated statistic beyond what would be expected under independence. Analyses that assume independence of the observations will generally underestimate the true variance and lead to test statistics with inflated Type I errors (Louviere and Woodworth 1983).

Gopinath (1995) demonstrated that different model forecasts result when the heterogeneity of travelers is not considered. Delvert (1997) argued that models of travel behavior in response to Advanced Traveler Information Systems must address heterogeneity in behavior. When we cannot consider the observations to be random draws from a large population, it is often reasonable to think of the unobserved effects as parameters to estimate, in which case we use fixed-effects methods. Even if we decide to treat the unobserved effects as random variables, we must also decide whether the unobserved effects are uncorrelated with the explanatory variables, which is the case in many situations. To draw accurate conclusions from correlated data, an appropriate model of within-cluster correlation must be used. If correlation is ignored by using a model that is too simple, the model would underestimate the standard errors of modeling effects (Stokes et al. 2000).

This paper reviews the existing methodologies for route choice modeling that account for one or both types of correlation mentioned above. The advantages and drawbacks of each methodology are stated. The main objective of this paper is to suggest a methodology (used in other fields) that accounts for correlation in binary and multinomial route choice modeling. BGEE and MGEE techniques are introduced with a binomial logit link function for BGEE and polytomous logistic link function for MGEE. The advantage of these techniques is that they account for correlation using a simple logistic link function instead of the probit function, which needs tremendous computational effort and cannot be used for relatively high numbers of alternatives or with large networks in multinomial models.

METHODOLOGIES THAT ACCOUNT FOR CORRELATION

Repeated Observations

Statisticians and transportation researchers have developed several methodological techniques to account for correlation between repeated observations made by the same traveler in binary and multinomial route choice models. Louviere and Woodworth (1983) and Mannering (1987) corrected the standard errors produced in a repeated responses regression model by multiplying the standard errors by the square root of the number of repeated observations. Kitamura and Bunch (1990) used a dynamic ordered-response probit model of car ownership with error components. Mannering et al. (1994) used an ordered logit probability model and a duration model with a heterogeneity correlation term. Morikawa (1994) used logit models with error components to treat serial correlation. Abdel-Aty et al. (1997) and Jou (2001) addressed this issue using individual-specific random error components in binary models with a normal mixing distribution. The standard deviation of the error components were found significant in both studies, which clearly showed the need for some formal statistical corrections to account for the unobserved heterogeneity. Jou and Mahmassani (1998) used a general probit model form for the dynamic switching model, allowing the introduction of state dependence and serial correlation in the model specification.

At the multinomial level, Mahmassani and Liu (1999) used a multinomial probit model framework to capture the serial correlation arising from repeated decisions made by the same respondent. Garrido and Mahmassani (2000) used a multinomial probit model with spatial and temporally correlated error structure.

Overlapping Alternatives

The correlation between alternative routes due primarily to overlapping distances has attracted many researchers to overcome the limitations of MNL models. The nested logit (NL) model (proposed by Ben-Akiva 1973) is an extension of the MNL model designed to capture correlation among alternatives. It is based on the partitioning of the choice set into different nests. The NL model partitions some or all nests into subnests, which can in turn be divided into subnests. This model is valid at every layer of the nesting, and the whole model is generated recursively. The structure is usually represented as a tree.

Clearly, the number of potential structures reflecting the correlation among alternatives can be very large. No technique has been proposed thus far to identify the most appropriate correlation structure directly from the data (apart from using a heteroskedastic extreme value choice model as a search engine for specification of NL structures). The NL model is designed to capture choice problems where alternatives within each nest are correlated. No correlation across nests can be captured by the NL model. When alternatives cannot be partitioned into well separated nests to reflect their correlation, the NL model is not appropriate.

Cascetta et al. (1996) introduced the C-logit model as a MNL model that captures the correlation among alternatives in a deterministic way. The authors use a term called "commonality factor," which they add to the deterministic part of the utility function to capture the degree of similarity between the alternative and all other alternatives in the choice set. The lack of theory or guidance on which form of commonality factor should be used is a drawback of the C-logit method.

McFadden (1978) presented the cross-nested logit (CNL) model as a direct extension of the NL model, where each alternative may belong to more than one nest. Similar to the NL model, the choice set is partitioned into nests. Moreover, for each alternative i and each nest m, parameters α_{im}, representing the degree of membership or the inclusive weight of alternative i in nest m, have to be defined. A CNL model is not appropriate for high numbers of alternatives.

Vovsha and Bekhor (1998) proposed and used a link-nested logit model as an application of the CNL model. The largest network they used contained 1 origin-destination pair, 8 nodes, 11 links, and 5 routes. Papola (2000) estimated a CNL model for intercity route choice with a limited number of alternative routes. Swait (2001) proposed the choice set generation logit model, in which choice sets form the nests of a CNL structure. The author acknowledged the computational difficulties of estimating this model when the choice set is large. It was concluded that, for a realistic size network and a realistic number of links per path, the CNL model and its applications become quite complex and therefore computationally onerous.

NL, C-logit, and CNL models are all extensions of the MNL models that use a logit utility function. An alternative technique is the multinomial probit (MP) model, which is derived from the assumption that the error terms of the utility functions are normally distributed. It uses a probit link function instead of a logit function. The MP model captures explicitly the correlation among all alternatives. Therefore, an arbitrary covariance structure can be specified. Mostly, this covariance structure was proportional to overlap length. Routes were also assumed to have heteroskedastic error terms where variance was proportional to route length or impedance. Yai et al. (1997) introduced a function that represents an overlapping relation between pairs of alternatives. The difficulty in implementing the probit model is that no closed form exists for the Gaussian cumulative distribution function, so numerical techniques must be used. Estimating an MP model is difficult even for a relatively low number of alternatives. Moreover, the number of unknown parameters in the variance-covariance matrix grows with the square of the number of alternatives (McFadden 1989).

Ben-Akiva and Bolduc (1996) introduced a multinomial probit model with a logit kernel (or hybrid logit) model, which combines the advantages of logit and probit models. It is based on a utility function that has two error matrices. The elements of the first matrix are normally distributed and capture correlation between alternatives. The elements of the second matrix are independent and identically distributed. These combined models have the same computational difficulties as pure MP. In general, any application of hybrid logit or probit to large-scale route choice is questionable in terms of the computational effort needed for estimating the parameter coefficients and their marginal effects, especially for large networks.

Based on the above review, a clear need exists for a methodology that accounts for the two kinds of correlation in binary and multinomial route choice models with a computationally easy and statistically efficient technique, both for small and large networks. This paper applies BGEE and MGEE with logit functions (binary and polytomous) to account for correlation between repeated observations in binary models and correlation between repeated observations and overlapping routes in multinomial models.

Applications

Route Choice and Switching

Pre-trip and en-route route switching is a direct response to Advanced Traveler Information Systems (ATIS). Network conditions, travel time, travel time variability, delays associated with congestion and incidents, and traveler attributes are significant determinants of route choice (Spyridakis et al. 1991; Adler et al. 1993; Mannering et al. 1994; Abdel-Aty et al. 1995a, b, 1997). Some studies proved that providing information induces greater switching in route choice behavior (Mahmassani 1990; Conquest et al. 1993; Abdel-Aty et al. 1994b). For example, Conquest et al. (1993) reported that 75% of commuters change either departure time or route in response to information. Liu and Mahmassani (1998) concluded that travelers were more likely to change their route when their current choice would cause them to arrive late. They also concluded that drivers exhibited some inertia in route choice, requiring travel time savings of at least one minute on the alternative route.

Benefits of ATIS

Many studies have examined the potential benefits of providing pre-trip and en route real-time information to travelers. Much research focuses on the effects of ATIS on all types of travel decisions. A number of studies show that ATIS results in reduced travel time, congestion delays, and incident clearance time (Wunderlich 1996; Abdel-Aty et al. 1997; Sengupta and Hongola 1998). Empirical evidence supports the hypothesis that travelers alter their behavior in response to ATIS (Bonsall and Parry 1991; Zhao et al. 1996; Mahmassani and Hu 1997). Reiss et al. (1991) reported travel time savings ranging from 3% to 30% and reduction in incident and congestion delays of up to 80% for impacted vehicles.

Drivers' Familiarity with the Network and Diversion

Polydoropoulou et al. (1996) and Khattak et al. (1996) concluded that drivers exhibit some inertia and tend to follow the same route, especially for home-to-work trips. Polydoropoulou et al. found that drivers are more likely to divert to another route when they learn of a delay before a trip. Drivers are less likely to divert during bad weather, as alternative routes may be equally slow. Prescriptive information greatly increases travelers' diversion probabilities, although similar diversion rates are attainable by providing real-time quantitative or predictive information about travel times on usual and alternative routes. The authors suggest that drivers would prefer to receive travel time information and make their own decisions. Abdel-Aty et al. (1994a) showed that ATIS has great potential to influence commuters' route choice even when advising a route different from the usual one.

Studies also indicate that traffic information should be provided along with alternative route information. Streff and Wallace (1993) reported differences in information requirements between commuting, noncommuting trips, and trips in an unfamiliar area. Khattak et al. (1996) found that travelers who were unfamiliar with alternative routes or modes were particularly unwilling to divert. This confirms the work of Kim and Vandebona (2002), which concluded that drivers who were familiar with an area had a high propensity to change their preselected routes. Further, accurate quantitative information might be able to overcome behavioral inertia if commuters are willing to follow advice from a prescriptive ATIS (Khattak et al. 1996; Lotan 1997). Adler and McNally (1994) found that travelers who were familiar with the network were less likely to consult information. Bonsall and Parry (1991) found that user acceptance declined with decreasing quality of advice in an unfamiliar network, and in a familiar network, drivers were less likely to accept advice from the system. However, Allen et al. (1991) found that familiarity does not affect route choice behavior.

GENERALIZED ESTIMATING EQUATIONS

The generalized estimating equations (GEE) technique analyzes discrete and correlated data with reasonable statistical efficiency. Liang and Zeger (1986) introduced GEE for binary models (BGEE) as an extension of generalized linear models (GLM). Lipsitz et al. (1994) extended the BGEE methodology to model correlation between repeated multinomial categorical responses (MGEE).

The GEE methodology models a known function of the marginal expectation of the dependent variable as a linear function of the explanatory variables. With GEE, the analyst describes the random component of the model for each marginal response with a common link and variance function, similar to what happens with a GLM model. However, unlike GLM, the GEE technique accounts for the covariance structure of the repeated measures. This covariance structure across repeated observations is managed as a nuisance parameter. The GEE methodology provides consistent estimators of the regression coefficient and their variances under weak assumptions about the actual correlation among a subject's choices.

In the following section, we provide a brief explanation of the BGEE models. The MGEE methodology is included in the appendix at the end of this paper.

Binary Generalized Estimating Equations

Suppose a number of n_i choices are made by subject i, where the total number of subjects is K, and y_{ij} denotes the jth response from subject i. There are $\sum_{i=1}^{K} n_i$ total choices (measurements). Let the vector of choices made by the ith subject be

$$Y_i = \left(y_{i1},\ldots,y_{in_i}\right)'$$

and let V_i be an estimate of the covariance matrix of y_i. Let the vector of explanatory variables for the jth choice on the ith subject be $X_{ij1} = (x_{ij1},\ldots,x_{ijp})'$.

The GEEs for estimating the $(1 \times p)$ vector of regression parameters β is an extension of the independence estimating equation to correlated data and is given by

$$\sum_{i=1}^{K} \frac{\partial \mu_i'}{\partial \beta} V_i^{-1}(Y_i - \mu_i(\beta)) = 0 \qquad (1)$$

where p is the number of regression parameters,

Since $g(\mu_{ij}) = x_{ij}, \beta$, the $p \times n_i$ matrix of partial derivatives of the mean with respect to the regression parameters for the ith subject is given by

$$\frac{\partial \mu_i'}{\partial \beta} = \begin{bmatrix} \dfrac{x_{i11}}{g'(\mu_{i1})} & \cdots\cdots & \dfrac{x_{in_i1}}{g'\left(\mu_{in_i}\right)} \\ \vdots & & \vdots \\ \dfrac{x_{i1p}}{g'(\mu_{i1})} & \cdots\cdots & \dfrac{x_{in_ip}}{g'\left(\mu_{in_i}\right)} \end{bmatrix} \qquad (2)$$

where

g is the logit link function $g(\mu) = \log(p(1-p))$, which is the inverse of the cumulative logistic distribution function, which is:

$$F(x) = \frac{1}{1+e^{-x}} \tag{3}$$

Working Correlation Matrix in BGEE

Let $R_i(\alpha)$ be an $n_i \times n_i$ "working" correlation matrix that is fully specified by the vector of parameters α (the correlation between any two choices). The (j, k) element of $R_i(\alpha)$ is the known, hypothesized, or estimated correlation between y_{ij} and y_{ik}. The covariance matrix of Y_i is modeled as

$$V_i = \phi A_i^{\frac{1}{2}} R(\alpha) A_i^{\frac{1}{2}} \tag{4}$$

where

A_i is an $n_i \times n_i$ diagonal matrix with $v(\mu_{ij})$ as the jth diagonal element.

ϕ is a dispersion parameter and is estimated by

$$\hat{\phi} = \frac{1}{N-p}\sum_{i=1}^{K}\sum_{j=1}^{n_i} e_{ij}^2, \qquad N = \sum_{i=1}^{K} n_i \tag{5}$$

R is the working correlation matrix. It is the same for all subjects, is not usually known, and must be estimated. The estimation occurs during the iterative fitting process using the current value of the parameter matrix β to compute appropriate functions of the Pearson residual

$$e_{ij} = \frac{y_{ij}-\mu_{ij}}{\sqrt{v(\mu_{ij})}}.$$

If $R_i(\alpha)$ is the true correlation matrix of Y_i, then V_i is the true covariance matrix of Y_i. If the working correlation is specified as $R = I$, which is the identity matrix, the GEE reduces to the independence estimating equation. The exchangeable correlation structure introduced by Liang and Zeger (1986) assumes constant correlation between any two choices within a subject/cluster. This exchangeable correlation structure can be used in the BGEE where the correlation matrix of each subject/cluster is defined as:

$$Corr(y_{ij}, y_{ik}) = \left\{ \begin{array}{l} 1,\ j = k \\ \alpha,\ j \neq k \end{array} \right\}$$

$$\text{e.g.,} \rightarrow \quad R_{3x3} = \begin{bmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{bmatrix} \tag{6}$$

where

$$\hat{\alpha} = \frac{1}{(N^*-p)\phi}\sum_{i=1}^{K}\sum_{j\neq k} e_{ij}e_{ik} \qquad \text{and}$$

$$N^* = \sum_{i=1}^{K} n_i(n_i-1) \tag{7}$$

DATA COLLECTION AND EXPERIMENT DESCRIPTION

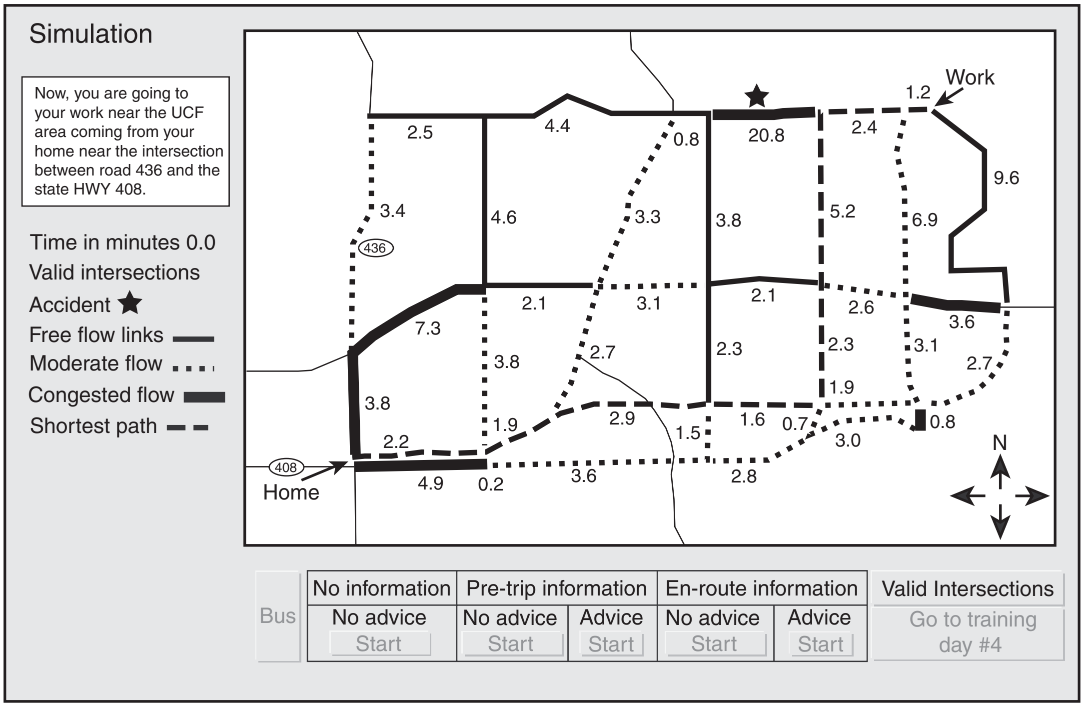
We used the travel simulator, Orlando Transportation Experimental Simulation Program (OTESP), to collect dynamic pre-trip and en-route route choice data. OTESP is an interactive windows-based computer simulation tool. It simulates a commuter home-to-work morning trip. OTESP provides five scenarios (levels) of traffic information to the subjects. In scenario #1, subjects receive no traffic information. Pre-trip information without and with advice are presented in scenarios #2 and #3, respectively. En route information, keeping the pre-trip information, without and with advice is presented in scenarios #4 and #5, respectively. The subject is required to choose his/her link-by-link route from a specified origin to a specified destination. The subject has the ability to move the vehicle on different segments of the network using the computer's mouse. Driving and riding one of two available bus routes are the travel modes used in OTESP. However, this study focuses only on the drive option.

In this study, we used a real network with historical congestion levels and weather conditions (figure 1). Intersections, recurring congestion, nonrecurring congestion (incidents), toll plazas, and weather condition delays are considered. The Moore's shortest path algorithm (Pallottino and Grazia 1998) was employed in the OTESP code to determine the travel-time-based shortest path, which is introduced as advice to the subjects in some scenarios. The simulation starts and ends with a short survey to collect the subjects' sociodemographic characteristics, preferences, perceptions, and feedback. A four-table database was created to capture all the information/advice provided and the traveler decisions. The program presents 10 simulated days (2 days for each scenario) after familiarizing the subjects with the system by introducing a training day for each scenario. Figure 1 shows a spot view of OTESP in its third scenario as an example.

Network

Figure 1 presents a portion of the city of Orlando network captured from a geographic information system database. The network has a unique origin-destination pair, where the assumed origin is the

FIGURE 1 Sample View of a Screen from the Orlando Transportation Experimental Simulation Program (Scenario #3)



subject's home and the assumed destination is the subject's work place. The network consists of 25 nodes and 40 links. This network portion was carefully chosen from the entirety of the Orlando network. It comprises different types of highways, including six-lane principle arterials, four-lane principle arterials, six-lane minor arterials, two-lane minor arterials, and local collectors. The network also includes two expressways.

Subjects

Subjects were recruited based on an experiment to guarantee the inclusion of groups of drivers that represent different incomes (two levels), ages (three levels), gender, familiarity with the network (two levels), and education (two levels). Because the subjects drove for their morning home-to-work trips, they were instructed that their main task was to minimize the overall trip travel time by deciding when and when not to follow the information and/or advice provided. Subjects were asked not to go through the simulation unless they had at least 30 minutes to devote to it (the average simulation took 23.77 minutes) and felt they could concentrate on it. Moreover, during the simulation, the subjects' response times were measured without notifying them, to ensure that they were paying attention. A total of 65 subjects participated in the simulation for 10 trial days each. Twenty-two subjects were under the age of 25 while 24 subjects were between 25 and 40 years of age, and 19 subjects were over 40 years old. Of the subjects, 24 were female and 41 were male. Two of the 65 subjects were excluded from this study, because their response times were outliers in the normal distribution (Z = 3.21 and 3.78, Z_{cr} = 2.57).

BGEE APPLICATION

Subjects viewed the level of congestion of every link in quantitative (travel time) and qualitative (green, yellow, and red links for free flow, moderate, and congested links, respectively) forms. The simulator also provided the shortest path from the subject's current position to the destination as advice. The information/advice level the subject received depended on the scenario, as mentioned above. At each node, the subject had to decide and choose between the two upcoming links. We considered this choice positive if the subject picked the link that had a lower level of congestion than the others (the delay on a link was equal to the difference between actual travel time at a specific movement—when a decision is made—and free flow travel time). A choice was considered negative if the subject picked the link with a higher level of congestion. We focused on the delay on a link when a particular movement occurred instead of travel time, because the links are different in length and speed limit.

Sixty-three subjects completed 10 trial days each, for a total of 539 trial days in the drive mode (the remainder of the trial days were in the transit mode). During the trial days, 4,753 movements (decisions) were made on the 40 network links. Out of the 4,753 movements, 1,667 were excluded from the analysis, because the driver had no choice but to proceed onto a unique coming link. The remaining 3,086 link choices make up the data used for the BGEE model with binomial logistic function. The model was correlated because each subject had multiple choices in the data structure. The response variable was binary with the value of one for positive choices and zero for negative choices. The explanatory variables follow:

1. **Information familiarity:** one if the subject, in real life, uses pre-trip and/or en route information usually or everyday, zero otherwise.
2. **Information provision:** one for trial days where en route information was provided, zero otherwise.
3. **Same color:** one if the two coming links had the same color (qualitative congestion level), zero otherwise. This variable tests the effect of qualitative vs. quantitative information.
4. **System learning:** one for the second five trial days of the simulations, zero for the first five. This is based on the assumption that the subject in the last five simulation runs is more familiar with the information system and can use and benefit from it more effectively.
5. **Heavy rain:** one for heavy rain conditions; zero for light rain or clear sky conditions. Weather conditions were provided as part of the information.
6. **Number of movements from the origin:** representing the closeness to the destination.

Table 1 presents the results of the BGEE model for the independent case (no correlation is considered) and for the proposed exchangeable correlation. The differences in the results are due to the effect of correlation. By comparing the overall F statistic values for the two models, the exchangeable model was favored over the independent model. This indicates that the model has correlation that should be accounted for.

The modeling results showed that, in general, the provision of en route information increases the likelihood of making a positive link choice. This means that the en route short-term information has a good chance of being used. When the two coming links had the same qualitative level of congestion, drivers were less likely to make a positive choice. Thus, the qualitative information is more likely to be used than the quantitative information. Therefore, it is not enough to provide the driver with the expected travel time or that there is congestion, but providing the driver with information on the level of congestion is also necessary.

The following effects/interactions increase the likelihood of following the en route short-term information:

1. Being familiar with traffic information;
2. Learning and being familiar with the system that provides the information;
3. Heavy rain conditions;
4. Being away from the origin, that is, close to the destination (presented by the number of movements since the origin);
5. Providing qualitative information in heavy rain conditions; and
6. Being away from the origin and being familiar with the device that provides the information.

MGEE APPLICATION

The long-term route choices of the subjects in the experiment were used as the database for estimating this model. The 539 routes that were chosen during the 539 trial days (each subject chooses one route each trial day) were identified and categorized by the sequence of links that were traversed on a given trial day. The network used consists of four west-east expressway/arterials that connect the origin to

TABLE 1 Modeling Results for the BGEE Model with and without Correlation

	Without correlation		With correlation	
Parameter	**Coefficient**	***t* statistic**	**Coefficient**	***t* statistic**
Intercept	–3.682	–20.29	–3.699	–20.11
Information familiarity: 1 if subject uses pre-trip and/or en route traffic information usually or everyday, 0 otherwise	0.406	10.12	0.421	9.94
Information provision: 1 for scenario 4 where en route information is provided without long-term advice, 0 otherwise	0.236	2.07	0.243	2.11
Same color: 1 if the two coming links had the same color (qualitative congestion level), 0 otherwise	–0.467	–4.20	–0.476	–4.28
System learning: 1 for the second 5 trial days of the experiment, 0 for the first 5 trial days	3.469	15.75	3.598	15.70
Heavy rain: 1 for heavy rain condition, 0 for light rain or clear sky	0.415	3.24	0.402	3.17
Number of movements since the origin	0.610	9.78	0.634	9.61
Interaction terms				
Heavy rain × Same color	0.453	2.25	0.536	2.16
Number of movements since the origin × System learning	0.406	6.14	0.511	6.06
Summary statistics	**df**	***F* statistic**	**df**	***F* statistic**
Overall model	9	6,742.27	10	8,543.12

the destination: named here MR1, MR2, MR3, and MR4.

MR1 represents the expressway alternative on the network. MR2 is a six-lane arterial while MR3 is mainly a four-lane arterial with a relatively high number of traffic lights. MR4 is primarily a rural, two-lane, two-way arterial with a speed limit approximately equal to that of MR2 and MR3. MR1 has the highest speed limit among the four alternatives with few traffic lights, because it consists mainly of expressway links. The network has also five local collectors that allow the subject to divert from one main route to another.

In order to come up with a reasonable number of alternatives, in the analysis phase, the route choices made during the trial days were aggregated into the above four main routes. We considered that each chosen route belonged to a main route if most of the chosen route's links belong to this main route. That is, a chosen route was assigned to a certain main arterial if, and only if, the chosen route overlaps with this main arterial for a longer distance than it does with any of the other three main arterials. As a result, the four main routes MR1, 2, 3, and 4 were chosen 374, 99, 37, and 29 times, respectively.

The proposed MGEE method with a generalized polytomous logit function was employed to model correlated route choices. The categorical dependent variable has four alternatives, MR1, MR2, MR3, and MR4. These four alternatives form the fixed choice set available for all subjects at all trial days. The reference alternative for which all attributes in the analysis are set equal to zero is MR4. This route was chosen because it was picked with lesser frequency over the other three main routes. The dependent variable takes on a value of one to four. The independent variables include:

1. **Age:** one if the subject's age is over 30, zero otherwise;
2. **Income:** one if household income is greater than $65,000, zero otherwise;
3. **Education:** one if the subject has a graduate-level degree or higher, zero otherwise;
4. **Shortest 1:** one if MR1 was the shortest path, zero otherwise;
5. **Shortest 2:** one if MR2 was the shortest path, zero otherwise;
6. **Shortest 3:** one if MR3 was the shortest path, zero otherwise;
7. **Advised 2:** one if MR2 was the shortest path and the trial day was under scenario #3 or #5 (i.e., MR2 was the suggested route), zero otherwise;
8. **Advised 3:** one if MR3 was the shortest path and the trial day was under scenario #3 or #5 (i.e., MR3 was the suggested route), zero otherwise;
9. **Travel time 1:** travel time on MR1;
10. **Travel time 2:** travel time on MR2;
11. **Travel time 3:** travel time on MR3;
12. **Travel time 4:** travel time on MR4.

Tables 2 and 3 show the modeling results using the MGEE model for the independent case (no correlation is considered) and for the proposed exchangeable correlation, respectively. The differences in the results are due to the effect of correlation. By comparing the overall *F* statistic values for the two models, the exchangeable model was favored over the independent model (83,417.09 vs. 11,464.98). Also, as expected, the independent MGEE model underestimated the standard errors of the modeling effects that lead to inflated *t* statistic values (table 2).

In table 3, the *t* statistics were lower when compared with the corresponding values in table 2 (for most of the effects), indicating that the proposed methodology has also adjusted this error. This means that the proposed methodology overcomes the disadvantage of underestimating the standard errors for models that do not account for correlation. A number of studies reported this disadvantage (Louviere et al. 1983; Mannering 1987; Gopinath 1995; Abdel-Aty et al. 1997; and Stokes et al. 2000). The model produced three logistic equations for the four alternatives (MR1 vs. MR4; MR2 vs. MR4, MR3 vs. MR4). These equations are:

$$\log\left(\frac{\hat{\pi}_{MR1}}{\hat{\pi}_{MR4}}\right) = -65.12 + 3.42Age + 2.36Income + 5.00Education + 22.15S1 + 11.65S2 + 13.23S3 + 29.60A1 - 4.70A2 - 4.87A3 - 6.00TT1 - 2.35TT2 - 0.67TT3 + 9.56TT4$$

$$\log\left(\frac{\hat{\pi}_{MR2}}{\hat{\pi}_{MR4}}\right) = -35.39 + 2.41Age + 1.01Income + 1.99Education + 21.62S1 + 12.17S2 - 7.01S3 + 5.56A1 + 11.34A2 - 25.95A3 - 4.28TT1 - 2.18TT2 - 0.36TT3 + 7.36TT4$$

$$\log\left(\frac{\hat{\pi}_{MR3}}{\hat{\pi}_{MR4}}\right) = -3.63 + 9.38Age + 2.49Income + 12.37Education - 10.49S1 - 48.61S2 + 22.67S3 + 3.69A1 - 44.56A2 + 1.41A3 - 0.79TT1 - 0.10TT2 - 1.00TT3 + 1.89TT4$$

where the symbols *Sx*, *Ax*, and *TTx* refer to the effects "Shortest *x*," "Advised *x*," and "Travel time *x*," respectively, where *x* is the main route number. Using the above equations, the probability of choosing an alternative given a set of values for the independent variables is simple compared with using any probit link function (probit models). Moreover, computing a certain marginal effect of any variable on choosing an alternative is straightforward and simple regardless of the number of alternatives used in the model, which is not the case for the corresponding multinomial probit models.

In the above equations, exponentiating the estimated regression coefficient yields the odds of choosing the corresponding alternative vs. choosing the base alternative MR4 for each one-unit increase in the corresponding explanatory variable. For example, the ratio of odds for a one-unit change in the travel time on MR2 is equal to $e^{-2.18} = 0.11$. This shows the ease of this model compared with the corresponding probit models.

Tables 2 and 3 also show the parameter coefficients for each equation with the corresponding *t*

TABLE 2 Modeling Results for the MGEE Model without Correlation

	Equation 1		Equation 2		Equation 3	
	MR1 vs. MR4		MR2 vs. MR4		MR3 vs. MR4	
Parameter	Coefficient	*t* statistic	Coefficient	*t* statistic	Coefficient	*t* statistic
Intercept	–41.31	–12.17	–24.93	–8.14	6.02	1.64
Age	0.04	1.15	0.42	0.55	6.70	5.50
Income	2.27	3.42	3.85	4.50	2.44	2.98
Education level	0.22	0.21	–2.36	–1.26	8.73	7.01
Shortest 1	12.71	9.92	13.37	10.64	–18.65	–17.52
Shortest 2	11.65	9.84	13.07	10.23	–23.48	–12.32
Shortest 3	1.65	1.83	–7.66	–5.63	22.67	16.71
Advised 1	7.93	9.24	21.79	16.86	3.69	3.67
Advised 2	7.26	5.12	12.95	8.22	–6.38	–2.98
Advised 3	–6.81	–7.04	–18.10	–11.33	1.41	1.35
Travel time 1	–6.07	–24.52	–4.35	–11.27	–0.89	–10.73
Travel time 2	–2.34	–19.66	–2.23	–20.18	–0.17	–2.82
Travel time 3	–0.67	–19.72	–0.36	–5.89	–0.96	–18.87
Travel time 4	9.44	26.15	7.36	13.34	1.89	22.90

Overall Summary Statistics

	df	*F* statistic
Overall model	42	11,464.98
Intercept	3	80.11
Age	3	12.10
Income	3	7.96
Education level	3	20.44
Shortest 1	3	197.39
Shortest 2	3	117.23
Shortest 3	3	167.53
Advised 1	3	96.60
Advised 2	3	44.95
Advised 3	3	57.78
Travel time 1	3	202.42
Travel time 2	3	151.75
Travel time 3	3	400.09
Travel time 4	3	478.48

statistic of each effect. Furthermore, tables 2 and 3 present the *F* statistic for each effect in the overall MGEE model. These values indicate the individual significance of every effect in the overall model and determine if changing the value of this effect statistically changes the probability of choosing a certain alternative. A certain effect may appear significant in one equation but be insignificant in another. All 13 effects included were found significant.

The parameter coefficients in table 3 show that older drivers (>30), those with larger household incomes, and those with a high level of education are, in general, more likely to choose MR1, MR2, or MR3 than MR4; that is, they are more likely to choose the expressways and/or the multilane arterials. Recall, MR4 is a two-lane, two-way rural arterial. However, the increase in this likelihood in some cases is not statistically significant. For example, these three socioeconomic factors above do not affect the probability of choosing MR2 vs. MR4 (t statistics = 1.07, 0.45, 0.74 < 1.96).

"Shortest 1," "Shortest 2," and "Shortest 3" measure the effect of providing information without advice to the subjects. The significance of "Shortest 1" in the first equation, with a positive coefficient parameter (22.15), shows that the probability of choosing the first alternative, MR1, increases if this route is the travel-time-based shortest route on the network, even with providing advice-free information. This means that the subjects were able to use and benefit from the qualitative and quantitative information provided to them. Moreover, they might be able to identify and then take the shortest route themselves using the travel times given to

TABLE 3 Modeling Results for the MGEE Model with Correlation

	Equation 1		Equation 2		Equation 3	
	MR1 vs. MR4		MR2 vs. MR4		MR3 vs. MR4	
Parameter	Coefficient	*t* statistic	Coefficient	*t* statistic	Coefficient	*t* statistic
Intercept	–65.12	–3.80	–35.39	–1.89	–3.63	–0.29
Age	3.42	2.56	2.41	1.07	9.38	3.87
Income	2.36	1.12	1.01	0.45	2.49	1.69
Education level	5.00	2.43	1.99	0.74	12.37	4.48
Shortest 1	22.15	2.74	21.62	3.51	–10.49	–1.62
Shortest 2	11.65	2.12	12.17	3.56	–48.61	–12.12
Shortest 3	13.23	5.04	–7.01	–3.13	22.67	6.57
Advised 1	29.60	3.85	5.56	0.76	3.69	0.67
Advised 2	–4.70	–1.22	11.34	3.79	–44.56	–10.03
Advised 3	–4.87	–0.64	–25.95	–2.98	1.41	0.24
Travel time 1	–6.00	–8.80	–4.28	–4.04	–0.79	–3.33
Travel time 2	–2.35	–5.61	–2.18	–6.03	–0.10	–0.67
Travel time 3	–0.67	–3.58	–0.36	–2.31	–1.00	–5.14
Travel time 4	9.56	7.03	7.36	3.99	1.89	3.24

Overall Summary Statistics

	df	*F* statistic
Overall model	43	83,417.09
Intercept	3	22.05
Age	3	10.31
Income	3	2.21
Education level	3	10.93
Shortest 1	3	79.65
Shortest 2	3	703.45
Shortest 3	3	57.34
Advised 1	3	21.51
Advised 2	3	48.55
Advised 3	3	19.91
Travel time 1	3	100.18
Travel time 2	3	36.34
Travel time 3	3	15.80
Travel time 4	3	81.22

them by the information system. The same interpretation applies to the coefficient parameters of the effects "Shortest 2" and "Shortest 3" in equations 2 and 3, respectively. By comparing these three coefficients (22.15, 12.17, 22.67), differences can be seen. This indicates that the marginal effects of these variables are not the same. However, they measure the same independent variable for different alternatives. Thus, it can be concluded that providing traffic information to drivers increases the likelihood that they will choose the shortest path (identified by them or given to them by an information system), but the odds differ between the shortest path and another, depending on the characteristics of each route.

To measure the effect of advising drivers to take a particular route, in addition to providing traffic information on all links of the network, the three effects, "Advised 1," "Advised 2," and "Advised 3" were employed. Advising MR1 or MR2 to the subjects increased the likelihood of their being their chosen (coefficients of 29.60 and 11.34, respectively). However, advising MR3 as the shortest path for a certain trial day does not affect its probability of being chosen (t statistic = 0.24). This result was not surprising, because MR3 is well known for its regular congestion due to its high accessibility and many traffic lights (most of the subjects were familiar with the network).

Similar to the effect of information without advice, the coefficient parameters "Advised 1" in equation 1, "Advised 2" in equation 2, and "Advised 3" in equation 3 (29.60, 11.34, 1.41, respectively) show that it is unclear that advising drivers to take a certain route increases the likelihood they will chose to do so. The characteristics of the route itself seem to be a factor in the decision. In this analysis, advising drivers to use an expressway or six-lane arterial increased the likelihood of it being chosen (MR1 and MR2). When drivers were advised to use a four-lane arterial with high density and traffic lights it did not affect the likelihood of that route being chosen. From these data, we can conclude that the characteristics of a certain route affect whether it is chosen even if the information advises drivers to use it.

The effect of travel time was represented in our model by the variables TT1, TT2, TT3, and TT4. The first three variables have negative coefficients in the three equations, with significant effects for TT1 in equation 1, TT2 in equation 2, and TT3 in equation 3. This clearly shows that the probability of choosing a certain route decreases as travel time increases. The effect TT4, the travel time of the base route MR4, showed up as a positive significant variable in the three equations. Therefore, the probability of choosing the other route (not choosing this base route) increases as travel time rises for the base alternative.

CONCLUSIONS

The proposed BGEE and MGEE techniques add new and useful methodology to the family of models that account for correlation in discrete choice models, especially for route choice applications. The literature review illustrated that a methodology was needed to account for correlation between repeated choices and/or between overlapping alternatives with simple computational effort and that can be applied to large networks. The proposed model proved to account for both types of correlation with simple computational effort and reasonable statistical efficiency for small and large networks. This makes BGEE and MGEE superior to the existing methodologies.

As a BGEE application, this paper presents a model of short-term route choice in compliance with ATIS. The paper also presents a multinomial route choice model (as an MGEE application). Both applications were developed with and without accounting for correlation. In both applications, the effect of correlation was tested statistically and found significant, which shows the importance of accounting for correlation in route choice models that may lead to different travel forecasts and policy decisions. This also shows the importance of our proposed methodology for large networks where the efficiency of the existing methodologies is questionable, as discussed in the literature review.

In this paper, we interpreted the modeling output of the BGEE and MGEE applications. The short-term route choice (BGEE) modeling results show that the provision of en route information increases the likelihood of making a positive link choice. The qualitative short-term information is more likely to be used than the quantitative information. Other effects were found to increase the usage of en route short-term traffic information: being familiar with the system that provides the information, heavy rain conditions, and proximity to the destination.

The multinomial route choice (MGEE) modeling results show that the subjects were able to use and benefit from the qualitative and quantitative information provided to them. Moreover, they might be able to identify the shortest route themselves using the travel times given to them. Finally, the odds of choosing a certain shortest route (advised or recognized by drivers using the advice-free traffic information provided) varied from one route to another and depended on the characteristics of the route itself. For example, the analysis in this paper showed that advising the use of the expressway or the six-lane arterial increase the likelihood of the route being chosen (MR1 and MR2). While advising the use of a four-lane arterial with a large number of traffic lights does not affect its likelihood of being chosen.

ACKNOWLEDGMENT

The author thanks Dr. M. Fathy Abdalla for his significant contribution to this paper. The results in this paper are based on a research project funded by the Center for Advanced Transportation Simulation Systems (CATSS) at the University of Central Florida (UCF) and the Florida Department of Transportation, and were included in Dr. Abdalla's Ph.D. dissertation at UCF, for which the author was the academic advisor (Abdalla 2003).

REFERENCES

Abdalla, M. 2003. Modeling Multiple Route Choice Paradigms Under Different Types and Levels of ATIS Using Correlated Data, Ph.D. dissertation. Department of Civil and Environmental Engineering, University of Central Florida.

Abdel-Aty, M., K. Vaughn, P. Jovanis, R. Kitamura, and F. Mannering. 1994a. Impact of Traffic Information on Commuters' Behavior: Empirical Results from Southern California and Their Implications for ATIS. *Proceedings of the 1994 Annual Meeting of IVHS America,* pp. 823–830.

Abdel-Aty, M., R. Kitamura, P. Jovanis, and K. Vaughn. 1994b. Investigation of Criteria Influencing Route Choice: Initial Analysis Using Revealed and Stated Preference Data, Report No. UCD-ITS-RR-94-12. University of California, Davis.

Abdel-Aty, M., R. Kitamura, and P. Jovanis. 1995a. Investigating the Effect of Travel Time Variability on Route Choice Using Repeated Measurement Stated Preference Data. *Transportation Research Record* 1493:39–45.

Abdel-Aty, M., K. Vaughn, R. Kitamura, P. Jovanis, and F. Mannering. 1995b. Models of Commuters' Information Use and Route Choice: Initial Results Based on Southern California Commuter Route Choice Survey. *Transportation Research Record* 1453:46–55.

Abdel-Aty, M., R. Kitamura, and P. Jovanis. 1997. Using Stated Preference Data for Studying the Effect of Advanced Traffic Information on Drivers' Route Choice. *Transportation Research C* 5(1):39–50.

Adler, J., W. Recker, and M. McNally. 1993. A Conflict Model and Interactive Simulator (FASTCARS) for Predicting En-Route Driver Behavior in Response to Real-Time Traffic Condition Information. *Transportation* 20(2):83–106.

Adler, J. and M. McNally. 1994. In-Laboratory Experiments to Investigate Driver Behavior Under Advanced Traveler Information Systems. *Transportation Research C* 2:149–164.

Allen, R., D. Ziedman, T. Rosenthal, A. Stein, J. Torres, and A. Halati. 1991. Laboratory Assessment of Driver Route Diversion in Response to In-Vehicle Navigation and Motorist Information Systems. *Transportation Research Record* 1306:82–91.

Ben-Akiva, M. 1973. Structure of Passenger Travel Demand Models, Ph.D. thesis. Massachusetts Institute of Technology.

Ben-Akiva, M. and D. Bolduc. 1996. Multinomial Probit with a Logit Kernel and a General Parametric Specification of the Covariance Structure, presented at the 3rd International Choice Symposium, Columbia University, New York, NY.

Bonsall, P. and T. Parry. 1991. Using an Interactive Route-Choice Simulator to Investigate Driver's Compliance with Route Guidance Advice. *Transportation Research Record* 1306:59–68.

Cascetta, E., A. Nuzzolo, F. Russo, and A. Vitetta. 1996. A Modified Logit Route Choice Model Overcoming Path Overlapping Problems: Specification and Some Calibration Results for Interurban Networks. *Proceedings from the 13th International Symposium on Transportation and Traffic Theory,* Lyon, France.

Conquest, L., J. Spyridakis, and M. Haselkorn. 1993. The Effect of Motorist Information on Commuter Behavior: Classification of Drivers into Commuter Groups. *Transportation Research C* 1:183–201.

Delvert, K. 1997. Heterogeneous Agents Facing Route Choice: Experienced versus Inexperienced Tripmakers, presented at IATBR '97, Austin, TX.

Garrido, R. and H. Mahmassani. 2000. Forecasting Freight Transportation Demand with the Space-Time Multinomial Probit Model. *Transportation Research B* 34(5):403–418.

Gopinath, D. 1995. Modeling Heterogeneity in Discrete Choice Processes: Application to Travel Demand, Ph.D. thesis. Massachusetts Institute of Technology.

Jou, R. 2001. Modeling the Impact of Pre-Trip Information on Commuter Departure Time and Route Choice. *Transportation Research B* 35:887–902.

Jou, R. and H. Mahmassani. 1998. Day-to-Day Dynamics of Urban Commuter Departure Time and Route Switching Decisions: Joint Model Estimation. *Travel Behavior Research*. New York, NY: Elsevier.

Khattak, A., A. Polydoropoulou, and M. Ben-Akiva. 1996. Modeling Revealed and Stated Pretrip Travel Response to Advanced Traveler Information Systems. *Transportation Research Record* 1537:46

Kim, K. and U. Vandebona. 2002. Understanding Route Change Behavior: A Commuter Survey in South Korea, presented at the 81st Annual Meetings of the Transportation Research Board, Washington, DC.

Kitamura, R. and D. Bunch. 1990. Heterogeneity and State Dependence in Household Car Ownership: A Panel Analysis Using Ordered-Response Probit Models with Error Components. *Transportation and Traffic Theory.* Oxford, England: Elsevier.

Liang, K. and S. Zeger. 1986. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika* 73:13–22.

Lipsitz, S., K. Kim, and L. Zhao. 1994. Analysis of Repeated Categorical Data Using Generalized Estimating Equations. *Statistics in Medicine* 13:1149–1163.

Liu, Y. and H. Mahmassani. 1998. Dynamic Aspects of Departure Time and Route Decision Behavior Under Advanced Traveler Information Systems (ATIS): Modeling Framework and Experimental Results. *Transportation Research Record* 1645:111–119.

Lotan, T. 1997. Effects of Familiarity on Route Choice Behavior in the Presence of Information. *Transportation Research C* 5:225–243.

Louviere, J. and G. Woodworth. 1983. Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data. *Journal of Marketing Research* 20:350–367.

Mahmassani, H. 1990. Dynamic Models of Commuter Behaviour: Experimental Investigation and Application to the Analysis of Planned Traffic Disruptions. *Transportation Research A* 24(6):465–484.

Mahmassani, H. and T. Hu. 1997. Day-to-Day Evolution of Network Flows Under Real-Time Information and Reactive Signal Control. *Transportation Research C* 5(1):51–69.

Mahmassani, H. and Y. Liu. 1999. Dynamics of Commuting Decision Behaviour Under Advanced Traveler Information Systems. *Transportation Research C* 7(2–3):91–107.

Mannering, F. 1987. Analysis of the Impact of Interest Rates on Automobile Demand. *Transportation Research Record* 1116:10–14.

Mannering, F., S. Kim, W. Barfield, and L. Ng. 1994. Statistical Analysis of Commuters' Route, Mode, and Departure Time Flexibility. *Transportation Research C* 2(1):35–47.

McFadden, D. 1978. Modeling the Choice of Residential Location. *Transportation Research Record* 672:72–77.

____. 1989. A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration. *Econometrica* 57(5):995–1026.

Morikawa, T. 1994. Correcting State Dependence and Serial Correlation in the RP/SP Combined Estimation Method. *Transportation* 21(2):153–165.

Pallottino, S. and M. Grazia. 1998. Shortest Path Algorithms in Transportation Models: Classical and Innovative Aspects. Edited by P. Marcotte and S. Nguyen. *Equilibrium and Advanced Transportation Modeling*. Boston, MA: Kluwer Academic Publishers.

Papola, A. 2000. Some Development of the Cross-Nested Logit Model. *Proceedings of the 9th IATBR Conference*, July 2000.

Polydoropoulou, A., M. Ben-Akiva, A. Khattak, and G. Lauprete. 1996. Modeling Revealed and Stated En-Route Travel Response to Advanced Traveler Information Systems. *Transportation Research Record* 1537:38.

Reiss, R., N. Gartner, and S. Cohen. 1991. Dynamic Control and Traffic Performance in a Freeway Corridor: A Simulation Study. *Transportation Research A* 25(5):267–276.

Sengupta, R. and B. Hongola. 1998. Estimating ATIS Benefits for the Smart Corridor Partners for Advanced Transit and Highways, UCB-ITS-PRR-98-30. Institute of Transportation Studies, University of California, Berkeley.

Spyridakis, J., W. Barfield, L. Conquest, M. Haselkorn, and C. Isakson. 1991. Surveying Commuter Behavior: Designing Motorist Information Systems. *Transportation Research A* 25:17–30.

Stokes, M., C. Davis, and G. Koch. 2000. *Categorical Data Analysis Using the SAS System,* 2nd ed. Cary, NC: SAS Institute.

Streff, F. and R. Wallace. 1993. Analysis of Drivers' Information Preferences and Use in Automobile Travel: Implications for Advanced Traveler Information Systems. *Proceedings of the Vehicle Navigation and Information Systems Conference,* Ottawa, Ontario, Canada.

Swait, J. 2001. Choice Set Generation within the Generalized Extreme Value Family of Discrete Choice Models. *Transportation Research B* 35:643–666.

Vovsha, P. and S. Bekhor. 1998. The Link-Nested Logit Model of Route Choice: Overcoming the Route Overlapping Problem. *Transportation Research Record* 1645.

Wunderlich, K. 1996. An Assessment of Pre-Trip and En Route ATIS Benefits in a Simulated Regional Urban Network. *Proceedings of the Third World Congress on Intelligent Transport Systems,* Orlando, FL.

Yai, T., S. Iwakura, and S. Morichi. 1997. Multinomial Probit with Structured Covariance for Route Choice Behavior. *Transportation Research B* 31(3):195–207.

Zeger, S., K. Liang, and P. Albert. 1988. Models for Longitudinal Data: A Generalized Estimating Equation Approach. *Biometrics* 44:1049–1060.

Zhao, S., N. Harata, and K. Ohta. 1996. Assessing Driver Benefits from Information Provision: A Logit Model Incorporating Perception Band of Information. *Proceedings of the 24th Annual Passenger Transport Research Conference,* London, England.

APPENDIX

Multinomial Generalized Estimating Equations (MGEE)

Suppose a number of t repeated choices are made by subject i ($i = 1,\ldots,N$), the total number of repeated choices for subject i is T_i, and K is the total number of alternatives available for all subjects at all observations. Two-level indicator variables can be formed as y_{ikt}, where $y_{ikt} = 1$ if subject i had the choice k at time t, while $y_{ikt} = 0$, otherwise. A $(k - 1)$ vector $y_{it} = [y_{i1t},\ldots,y_{i,K-1,t}]'$ can be formed to show the choice of subject i at time t. Each subject has T_i covariate vectors x_{it}, where an x_{it} vector contains all the relevant covariates including the intercept, between- and within-subject covariates. Therefore, each subject has a matrix of covariates

$$X_i = \left[x_{i1},...,x_{iT_i}\right]'$$

of dimension $T_i \times p$, where p is the total number of covariates excluding the intercept.

The distribution of y_{it} is multinomial with the probability function

$$f(y_{it}|x_{it}, \beta) = \prod_{k=1}^{K} \pi_{ikt}^{y_i kt} \tag{8}$$

where $\pi_{ikt} = E(y_{ikt}|x_{it}, \beta) = pr\{y_{ikt} = 1|x_{it}, \beta\}$ is the probability that subject i had choice k at time t, and β is a $p \times 1$ vector of parameters. When y_{it} is binary, π_{ikt} is usually modeled with a logistic or probit link function (Zeger et al. 1988). When $k > 2$ with non-ordered response, the generalized polytomous logit link is appropriate (Lipsitz et al. 1994).

The matrix of coefficient parameters β is associated with the $[(K-1) \times 1]$ marginal probability vector

$$E(Y_{it}|X_i) = \pi_{it}(\beta) = \left[\pi_{it1},...,\pi_{i,(K-1),t}\right]' \tag{9}$$

These marginal probability vectors can be grouped together to form the $[T_i(K-1) \times 1]$ vector

$$E(Y_i|X_i) = \pi_i(\beta) = \left[\pi'_{i1},..., \pi'_{iT_i}\right]'$$

where

$$Y_i = \left[Y'_{i1},..., Y'_{iT_i}\right]' \tag{10}$$

The GEEs of the following form can be used to estimate β (Liang and Zeger 1986; Lipsitz et al. 1994)

$$u\left(\hat{\beta}\right) = \sum_{i=1}^{N} \frac{d[\pi_i(\beta)]'}{d\beta} \hat{V}_i^{-1}\left[Y_i - \hat{\pi}_i\right] = 0 \tag{11}$$

where V_i is the covariance matrix of Y_i. This covariance matrix, V_i, is a function of β and other nuisance parameters α, which is a function of the correlation between repeated choices made by the same subject i. Also, V_i depends on the correlation between overlapped (or correlated) alternative routes. This covariance matrix, V_i, has $[T_i \times T_i]$ blocks. Each block has $[(K-1) \times (K-1)]$ elements.

Estimating the Covariance Matrix

To get a general form of V_i, the correlation matrix of the elements of Y_i must be developed or estimated first. Therefore, the pairwise correlation between the (K – 1) elements of Y_{is} and Y_{it}, which accounts for correlation between observations s and t of subject i, must be determined. A typical element of the correlation matrix of the elements of Y_i is, for any pair of responsive levels j and k and pair of times s and t,

$$Corr(Y_{ijs}, Y_{ikt}) = E[e_{ijs}\ e_{ikt}],$$

$$\text{where } e_{ikt} = \frac{Y_{ikt} - \pi_{ikt}}{\left[\pi_{ikt}(1 - \pi_{ikt})\right]^{1/2}} \tag{12}$$

The element e_{ikt} is the residual for Y_{ikt}. This residual e_{ikt} is a typical element of the residual vector

$$e_{it} = A_{it}^{-1/2}\left[Y_{it} - \pi_{it}\right]$$

where A_{it} is a function of β and is equal to:

$$A_{it} = Diag\left[\pi_{i1t}(1 - \pi_{i1t}),..., ...,\pi_{i,K-1,t}(1 - \pi_{i,K-1,t})\right] \tag{13}$$

$$A_{it}^{-1/2} = Diag\left[(\pi_{i1t}(1-\pi_{i1t}))^{1/2},..., ...,(\pi_{i,K-1,t}(1-\pi_{i,K-1,t}))^{1/2}\right] \tag{14}$$

The correlation matrix of $Y_i = R_i(\alpha)$ with e_{ikt} as a typical element can be written as

$$Corr(Y_i) = R_i(\alpha) = \text{var}(e_i) = A_i^{-1/2}\text{var}(Y_i)A_i^{-1/2} \tag{15}$$

or

$$\text{var}(Y_i) = V_i = A_i^{1/2}Corr(Y_i)A_i^{1/2} \tag{16}$$

where

$$\hat{e}_i = \left[\hat{e}_{i1},...,\hat{e}_{iT_i}\right], \quad \text{and}$$
$$A_i = Diag\left[A_{i1},...,A_{iT_i}\right]$$

Then, var(Y_i) depends on β and $R_i(\alpha)$ where the latter takes the effect of correlation in computing the covariance matrix var(Y_i). The matrix $R_i(\alpha)$ is a T_i by T_i block diagonal matrix. Each block is a $[(K-1) \times (K-1)]$ matrix. The tth diagonal block of $R_i(\alpha)$ is $A_{it}^{-1/2} V_{it} A_{it}^{-1/2}$, also the *s*th-row and *t*th-column off-diagonal block $\rho_{ist}(\alpha)$ is

$$\rho_{ist}(\alpha) = A_{is}^{-1/2} E\left[\left(Y_{is} - \pi_{is}\right) \bullet \left(Y_{it} - \pi_{it}\right)'\right] A_{it}^{-1/2} \qquad (17)$$

where

$$V_{it} = \text{var}(Y_{it}) = Diag[\pi_{it}] - \pi_{it}\pi_{it}'$$

and $Diag[\pi_{it}]$ denotes a diagonal matrix with elements of π_{it} on the main diagonal and zero off-diagonal elements. The diagonal blocks of $R_i(\alpha)$ depend only on $\pi_i(\beta)$. In these diagonal blocks, the diagonal elements are:

$$Corr\left(Y_{ikt}, Y_{ikt}\right) = 1 \qquad (18)$$

and the off-diagonal elements are

$$Corr\left(Y_{ijt}, Y_{ikt}\right) = \frac{cov\left(Y_{ijt}, Y_{ikt}\right)}{\{\pi_{ijt}(1-\pi_{ijt})\pi_{ikt}(1-\pi_{ijt})\}^{-1/2}} = \frac{-\pi_{ijt}\pi_{ikt}}{\{\pi_{ijt}(1-\pi_{ijt})\pi_{ikt}(1-\pi_{ikt})\}^{-1/2}} \qquad (19)$$

Recall that these off-diagonal elements of the diagonal blocks of $R_i(\alpha)$ depend only on the *t*th choice of subject *i* from the *K* alternatives available. This clearly takes care of any correlation among the different alternatives of the multidimensional route choice model, usually due to overlapping distances between different routes. Thus, the unknown elements of $R_i(\alpha)$ are the elements of its off-diagonal blocks $\rho_{ist}(\alpha)$. This must be estimated.

If $\rho_{ist}(\alpha)$ is known, then $R_i(\alpha)$ is known. The only unknown term in equation 11 then is β. The estimated $\hat{\beta}$ can be obtained by a Fisher scoring algorithm until convergence,

$$\hat{\beta}^{m+1} = \hat{\beta}^m + \left[\sum_{i=1}^{N} \frac{d\left[\pi_i\left(\hat{\beta}^m\right)\right]}{d\beta}\left(\hat{\beta}^m\right)' \bullet \left[V_i\left(\hat{\beta}^m, \hat{\alpha}^m\right)\right]^{-1}\right]\left[\frac{d\left[\pi_i\left(\hat{\beta}^m\right)\right]}{d\beta}\hat{\beta}^m\right]^{-1} \bullet \sum_{i=1}^{N} \frac{d\left[\pi_i\left(\hat{\beta}^m\right)\right]}{d\beta}\left(\hat{\beta}^m\right)'\left[V_i\left(\hat{\beta}^m, \hat{\alpha}^m\right)\right]^{-1} \bullet \left[Y_i - \pi_i\left(\hat{\beta}^m\right)\right] \qquad (20)$$

where m is the iteration number. A starting β can be obtained by applying the regular MNL model. Iteration should continue until $\hat{\beta}^{m+1} = \hat{\beta}^m$ and $\hat{\alpha}^{m+1} = \hat{\alpha}^m$, where $\hat{\alpha}^m$ is the estimated $\rho_{ist}(\alpha)$ in the *m*th step.

Estimating the Off-Diagonal Blocks of the Correlation Matrix

Lipsitz et al. (1994) extended the exchangeable correlation structure, introduced by Liang and Zeger (1986), used in BGEE for multidimensional models. They used the same assumption that any two observations on the same subject/cluster *i* and category k are equally correlated. Under this assumption, $\rho_{ist}(\alpha)$ can be estimated as

$$\hat{\alpha} = \rho_{ist}(\hat{\alpha}) = \frac{\sum_{i=1}^{N} \sum_{t>s} \hat{e}_{is}\hat{e}_{it}'}{\left[\sum_{i=1}^{N} 0.5T_i(T_i - 1)\right] - p} \qquad (21)$$

where p is the total number of independent variables, including any interactions. The residual vector

$$\hat{e}_{it} = \hat{A}_{it}^{-1/2}\left[Y_{it} - \hat{\pi}_{it}\right],$$

which is estimated by plugging $\hat{\beta}$ from a previous step of iteration into A_{it} and π_{it}. It is worth mentioning that the elements of the *s*th-row and *t*th-column off-diagonal block $\rho_{ist}(\alpha)$ do not depend on the times *s* and *t*, but they do depend on the levels *j* and k.

BOOK REVIEWS

Economic Impacts of Intelligent Transportation Systems: Innovations and Case Studies
E. Bekiaris and Y.J. Nakanishi, editors
Elsevier
July 2004, 655 pages
ISBN 0762309784
$95 €150 £100

This new addition to the series, *Research in Transportation Economics,* is dedicated to the economic valuation of intelligent transportation systems (ITS) and telematics. The editors include the papers of close to 50 authors, thus ensuring comprehensive coverage of the topic. ITS technologies typically have short life cycles; they are rapidly evolving and have minimal historical data associated with them. Recognizing that ITS life cycles and cost structures differ greatly from that of regular infrastructure, techniques to address these issues are presented. The book covers a wide range of ITS technologies, including freeway management, electronic toll collection, advanced driver assistance systems, and traveler information systems.

The introductory chapter of this book recaps the goals and objectives of ITS as envisioned in landmark legislation. The chapter presents a Mitretek (2003) compilation of recent cost-benefit analysis that draws on experiences up to 2003. Also provided are definitions for acronyms directly downloaded from an ITS architecture website. Distinctly absent, however, is a critique of ITS implementations over the past two decades. This type of discussion would have given the readers some perspective on the topic.

The section on "Relevant Technologies and Market" identifies the marketplace for various ITS technologies. Two articles are included: the first, by Panou and Bekiaris, puts forth a taxonomy of technologies into clusters in the United States and overseas; the second, by Kauber, assesses the emerging markets through 2010.

The centerpiece of the volume is really the section on "Evaluation Technologies/Methodologies," commanding no less than 120 pages. Aside from traditional techniques such as cost-benefit analysis, I am pleased to see contributions on "Analytical Alternatives" (Haynes and Li), and I am equally delighted to see microsimulation being used in evaluating "Variable-Message-Signs Route Guidance" (Ozbay and Bartin).

While the methodologies section is cut and dried, the articles in the "Case Studies" sections make them come to life. They include:

- Incident Freeway Management,
- Electronic Toll Collection and Commercial Vehicle Operation,
- Public Transport, and
- Advanced Driver Assistance Systems (ADAS) and Driver/Traveler Information.

Aside from the "Evaluation Technologies/Methodologies" section, I view these 300 pages among the most useful parts of the book, covering the key components of ITS.

In the section on "Assessing the Impact of ITS on the Overall Economy," Kawakami et al. put forth a compact general equilibrium model to measure the implications of ITS. Gillen et al. evaluate the productivity gains from using Automatic Vehicle Location. I would have liked to have seen more contributors in this section, because it addresses a key objective of this volume; in fact it is the basic premise of the book.

Among the finishing touches, the book includes policy recommendations. Again, I rank this section of the book very highly. I am particularly partial to the guidelines laid down for implementations. For example, recommendations are made by Bekiaris et al. for two ADAS technology clusters: advanced cruise control and intelligent speed adaptation.

I am a bit disappointed, however, with the brevity of the conclusions chapter, which says nothing new or important. Considering the richness of the contributions of so many authors, surely the editors could have summarized them better and in more detail.

I am also disappointed in the lack of a subject index, which discounts the usefulness of the book. More importantly, it reinforces the impression that there is only a limited amount of editorial oversight.

Overall, this is a welcome addition to the *Research in Transportation Economics* series. Many of the 50 authors included in the volume are respected figures in this field, and the separate papers provide much useful information. Integrating the chapters into a book entitled *Economic Impacts of ITS,* however, falls a bit short of my expectations.

Reference

Mitretek Systems. 2003. ITS Benefits and Costs: 2003 Update, prepared for the Federal Highway Administration, U.S. Department of Transportation.

Reviewer address: Yupo Chan, Professor and Founding Chair, Department of Systems Engineering, University of Arkansas at Little Rock, 2801 South University, Little Rock, AR 72204-1099. Email: yxchan@ualr.edu.

Principles of Transport Economics

Emile Quinet and Roger Vickerman
Edward Elgar Publishing Company
2004, 385 pages
ISBN 1-84064-865-1
$120 hardback ($60 paperback)
£75 hardback (£35 paperback)

Instructors of upper-level undergraduate transportation economics courses have typically been frustrated searching for a suitable textbook. The traditional transportation books are long on institutional description and short on analytical method. A recent pair of books attempted to remedy this. Kenneth Boyer's *Principles of Transportation Economics* (Addison Wesley, 1998, $124.40) is well written but pitched at too low a level for students who have completed intermediate microeconomics. Patrick McCarthy's *Transportation Economics* (Blackwell, 2001, $108.95) is the opposite. There is too much detail for a typical undergraduate who is not specializing in transportation, and it is a tough read in places. In my class, I have been using José Gómez-Ibáñez, William Tye, and Clifford Winston's edited volume *Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer* (Brookings Institution, 1999, at the attractive price of $22.95 in paperback), which is a mixed bag and really does not function as a textbook.

To this mix is added a new text by Quinet and Vickerman. Actually, it is an updated English-language version of an earlier book in French by Quinet. The addition of Vickerman from England has given the book a more pan-European feel. The limitation of nearly all applied economics texts is that their examples are parochial. European students have limited interest in the detailed economics of the Mackinac Straits bridge in Michigan or the Lindenwold transit line in Philadelphia, and conversely the Channel Tunnel and the externality costs of truck traffic in Switzerland do not stir the imagination of American students. Therefore, Quinet and Vickerman face a tough task in breaking into the North American market.

The book is divided into three sections. The first fifth (in terms of pages) deals with putting transportation in its wider context in relation to economic activity, location theory, and urban economics. The next quarter deals with the basics of demand and cost. The remaining pages cover market equilibrium, market failure, and public policy intervention.

Overall, I consider the book to be a pedagogical failure. For example, within the first 40 pages there is a quick gallop through price elasticities (page 17), cost functions and the envelope theorem (page 27), and the throw away parenthetical comment on page 37 while discussing a model of interregional trade that "there is no trade initially (infinite transport costs)." All of this while the reader has yet to be introduced to basics of transportation demand and supply.

With my students, I consider it to be a major success if they can interpret and use elasticities, understand how a cost function is derived and estimated, and fully understand the role of transportation in trade equilibria and how a demand function for freight transportation can be derived. This book is not supportive of the crucial material that I feel is essential for my students to understand. While the first section contains important concepts, I felt that the material could have been better organized, and perhaps combined with topics covered later in the book, to produce a more logical and instructive narrative.

I found the second section dealing with demand and costs to be the strongest part of the book. The chapter on demand contains a particularly good description of the underpinning of traditional four-stage demand models. However, it deals with pure theory. There is no discussion of the practical econometrics of estimation or any empirical results. This is in contrast to McCarthy's text that has extensive practical discussion of this topic. However, the most amazing thing is that after 45 pages of discussing passenger demand models, freight traffic models are considered in under a page. Of course, this is not entirely the authors' fault. The profession has not lavished the attention on freight demand models the way that it has for passenger movement. This is a sad state of affairs in the North American context given the importance of freight railroads and competition with trucks, barges, and pipelines.

The chapter on costs is also quite good and discusses social costs as well as production costs. To my tastes, I felt there should have been a more extensive industrial organization-style analysis of production costs. After all, production costs drive firm behavior and market equilibrium. For example, on page 147 there is only passing reference to the difference between economies of firm size, economies of density, and economies of scope. And the discussion is based on ill-defined equations and not supported by any intuition. Yet, with my students, I find that driving home these distinctions produces greater appreciation of the differences between transportation modes and the debate concerning regulatory reform.

The final section of the book is a random walk through market equilibria and public policy. Again the book fails pedagogically by giving short shrift to important tools in transportation economics. Concepts such as Ramsey pricing (in the cost chapter, page 153), an undefined "Mohring effect" (also in the cost chapter, page 158), cost-benefit analysis (page 239), and reaction functions and the conduct parameter (page 266) are all mentioned in passing but not given the full treatment that undergraduates require. Overall, the chapters in the final section of the book are adequate and cover most of the important issues. However, I felt they lacked a structure that students could use to guide their studies.

Overall, one can say that the book touches on all of the topics usually found in an upper level transportation economics course, however, the ordering of material detracts from the learning experience. One can always quibble about the organization of the material, and I would accept that this is partly a matter of taste, and there will be deficiencies in almost any ordering. However, I would suggest that McCarthy and (especially) Boyer have got it right by using the structure of a typical intermediate microeconomics textbook as a guide to a logical sequence of topics.

The other failure of this book is in explaining and illustrating the principal tools of transportation economics in a way that undergraduates would find useful. One might consider that there are three main types of academic writing that do not report original research results. One is a handbook chapter presenting core reference material for a knowledgeable new researcher in a field. The second is a "handbook chapter on steroids" prepared for knowledgeable insiders and published in places such as the *Journal of Economic Literature*. The third is textbook writing designed for newcomers to the field who only have basic economic training. The authors of this book have failed; they have pitched the book somewhere in the middle and thereby satisfied none of the three markets.

How did this happen? The basic problem is that the market for a post-intermediate microeconomics text on transportation is very small. Publishers are, therefore, unwilling to invest resources in extensive use of referees, prepublication testing in a classroom setting, copy editors, graphic designers, and the preparation of end-of-chapter exercises. Edward Elgar, the publisher of this book, is not a major player in the mass-market textbook publishing business. Consequently the authors did not get the feedback they should have received at an early stage that would have allowed them to make this a more pedagogically useful tool.

Reviewer address: Ian Savage, Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston IL 60208, USA. Email: ipsavage@northwestern.edu.

Employment in the Airline Industry

Airlines report employment levels to the federal government on a monthly basis.[1] The release of the employment numbers on the Bureau of Transportation Statistics (BTS) website and through monthly press releases allows tracking of employment trends in the industry resulting from the financial crises experienced by many airlines after the terrorist and recession events of 2001. The employment numbers reflect the resultant reduction in the network carriers' operations and the subsequent growth of low-cost and regional carriers.

BTS makes year-end (December) airline employment data available online.[2] These data are supplemented by the monthly press releases, which BTS began reporting in March 2005.[3] Using these BTS data, analyses of the changes in the number of airline employees can be performed by business model, revenue size, or individual carrier.

The best representation of the current airline industry structure is a business model definition, which contains three carrier groupings: network, low cost, and regional. Network carriers use a traditional hub-and-spoke system for scheduling flights. Low-cost carriers operate under a generally recognized low-cost business model, which may include a single passenger class of service, standardized aircraft utilization, limited in-flight services, use of smaller and less expensive airports, and lower employee wages and benefits. Regional carriers provide service from small cities and primarily use smaller jets. Regional carriers are also used to support larger network carrier traffic into and out of smaller airports to the network carriers' hub airports.

The BTS online database uses an operating revenue classification system with three major categories: majors, nationals, and regionals.[4] Major carriers have annual operating revenues above $1 billion, national carriers have operating revenues between $100 million and $1 billion, and regional carriers have operating revenues less than $100 million. The regional category may not include the same carriers under the revenue size definition and the business model definition.

Using the business model classification, there are seven network carriers (Alaska Airlines, American Airlines, Continental, Delta, Northwest, United, and US Airways) and eight low-cost carriers (Air Tran, America West, ATA, Frontier, Independence, JetBlue, Southwest, and Spirit). December 2004 employment data provided by the airlines shows that American Airlines was the largest employer (71,232 full-time personnel) in the industry (table 1). American became the largest employer in 2001, the year it acquired the assets of Trans World Airlines. United was the second largest network employer in December 2004 with 54,460 full-time employees, followed by Delta (53,394). The remaining network carriers all employed less than 40,000 full-time personnel each.

[1] This includes all carriers that operate at least 1 aircraft with a carrying capacity of 18,000 pounds (passengers, cargo, and fuel). Regional carriers were excluded from reporting until 2003.
[2] Available at www.bts.gov/programs/airline_information/number_of_employees/.
[3] Available at www.bts.gov/press-releases.

[4] This system of classification emerged during the industry's pre-deregulation era.

Review by: Jennifer Brady, Analyst, Bureau of Transportation Statistics, Research and Innovative Technology Administration, U.S. Department of Transportation, 400 Seventh St. SW, Room 3430, Washington, DC 20590. Email address: jennifer.brady@dot.gov.

TABLE 1 Airline Employees by Business Model and Size Classification: 2004

Carrier	All employees	Full-time employees	Business model classification	Size classification
American	82,222	71,232	Network	Major
United	61,092	54,460	Network	Major
Delta	59,972	53,394	Network	Major
Northwest	39,784	37,572	Network	Major
Continental	35,395	28,227	Network	Major
US Airways	26,169	23,180	Network	Major
Alaska Airline	9,857	8,747	Network	Major
Southwest	31,274	30,749	Low cost	Major
America West	12,654	10,129	Low cost	Major
JetBlue	7,399	5,956	Low cost	National
AirTran	6,072	5,754	Low cost	National
ATA	6,268	5,625	Low cost	Major
Frontier	4,492	3,620	Low cost	National
Independence	4,301	4,014	Low cost	National
Spirit	2,627	2,339	Low cost	National

Source: U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics, *Certified Carriers (Full-time and Part-time),* available at www.bts.gov/programs/airline_information/number_of_employees.

All network carriers experienced employment growth throughout the 1980s. After a slowdown during the early 1990s, precipitated by the 1991 recession and the Gulf War, the number of network carrier employees began to increase again in 1995. However, immediately following the 2001 terrorist events and the simultaneous mild U.S. recession, employment decreased for the network carrier group.

Between December 2000 and 2004, all the network carriers reduced their numbers of full-time employees; however, certain carriers were particularly hard hit. The total number of US Airways full-time personnel decreased by 44% from December 2000 to 2004, and United experienced a 40% reduction in their full-time work force. The other five major carriers reduced their total number of full-time employees by approximately 20%, except Alaska Airlines, which sustained a 4% loss.

Southwest Airlines employs the most personnel among the low-cost carriers. Its December 2004 employment rolls indicate more than twice the number of full-time employees (30,749) worked for Southwest than the next largest employer, America West (10,129). The remaining low-cost carriers employed less than 10,000 full-time personnel each.

As a demonstration of the growing strength of low-cost carriers, Southwest, America West, and ATA generate sufficient revenues to qualify as major carriers. Furthermore, Southwest employs more people than two of the network carriers (US Airways and Alaska Airlines).

It is difficult to analyze historic trends for the airline industry because of an evolving carrier population and the financial demise of previously qualifying airlines. Among the network airlines operating in 1970, most had their highest employment levels before 2001. From December 2000 to 2004, the number of full-time personnel employed by network carriers decreased 31%[5] (table 2). During the same time period, employment among seven low-cost carriers reporting to BTS throughout the entire period (excluding Independence Air) increased 17%.

[5] December 2000 and 2001 data include statistics for TWA, which was purchased by American Airlines in 2001.

TABLE 2 Airline Employees: 2000–2004
(all data are from December of a given year)

Type of carrier	2000	2001	2002	2003	2004	Percentage change[1]
Network	399,971	348,882	334,444	286,940	276,812	–31%
Low cost	54,746	53,258	59,803	64,048	64,172	17%

[1] Among those airlines reporting in 2000 and 2004.

Source: U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics, *Certified Carriers (Full-time and Part-time),* available at www.bts.gov/programs/airline_information/number_of_employees.

New BTS monthly employment statistics show that the reduction in the number of network carrier employees has slowed, as has growth among low-cost carriers. In December 2004, January 2005, February 2005, and March 2005, full-time employment among network carriers showed decreases ranging from 3.5% to 5.3% (table 3). (The data comparisons presented here are for the same months in 2003 and 2004 or 2004 and 2005, not a full year of data.) Even low-cost carrier full-time employment experienced negative growth in the most recent period (March) at –0.3%.

Regional carrier data are also available in BTS press releases and online. However, not all of the currently reporting carriers were required to report until 2003, thus only 8 reported in 2000 and 13 reported in 2004, making comparisons over time problematic.

Raw data on airline employment are available on the BTS website (see footnote 2). *Certificated Carriers (Full-time and Part-time)* provides employment data beginning in 1970, and *P10—Annual Employee Statistics by Labor Category,* covering 1998 through 2003, provides detailed information for each airline by labor category and geographic groupings. Among the 15 labor categories detailed in the *P10* database are pilots and co-pilots, maintenance staff, and cargo handling and other flight personnel. The four geographic groupings in the *P10* database are domestic, Atlantic, Latin America, and Pacific. Low-cost carriers are almost exclusively domestic operators.

BTS will continue to issue monthly press releases with the most current data and employment trends. Using these press releases, it is possible to supplement the online databases with monthly data. Finally, the press releases provide documentation of business model changes by carriers, resulting in new groupings.

For further information on this topic, send email to answers@bts.gov or call 1-800-853-1351.

TABLE 3 Percentage Change in Airline Employment

Type of carrier	December 2003 and 2004	January 2004 and 2005	February 2004 and 2005	March 2004 and 2005
Network	–3.5	–4.5	–5.0	–5.3
Low cost	6.5	0.3	0.3	–0.3

Source: U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics, Table 1: Passenger Airline Employment, monthly press releases, 2005, available at www.bts.gov/press-releases.

JOURNAL OF TRANSPORTATION AND STATISTICS

Guidelines for Manuscript Submission

Please note: Submission of a paper indicates the author's intention to publish in the *Journal of Transportation and Statistics* (JTS). Submission of a manuscript to other journals is unacceptable. Previously published manuscripts, whether in an exact or approximate form, cannot be accepted. Check with the Managing Editor if in doubt.

Scope of JTS: JTS publishes original research using planning, engineering, statistical, and economic analysis to improve public and private mobility and safety in all modes of transportation. For more detailed information, see the Call for Papers on page 112.

Manuscripts must be double spaced, including quotations, abstract, reference section, and any notes. All figures and tables should appear at the end of the manuscript with each one on a separate page. Do not embed them in your manuscript.

Because the JTS audience works in diverse fields, **please define** terms that are specific to your area of expertise.

Electronic submissions via email to the Managing Editor are strongly encouraged. We accept PDF, Word, Excel, and Adobe Illustrator files. If you cannot send your submission via email, you may send a CD by overnight delivery service or send a hardcopy by the U.S. Postal Service (regular mail; see below). Do not send CDs through regular mail.

Hardcopy submissions delivered to BTS by the U.S. Postal service are irradiated. Do not include a disk in your envelope; the high heat will damage it.

The cover page of your manuscript must include the title, author name(s) and affiliations, and the telephone number and surface and email addresses of all authors.

Put the **Abstract** on the second page. It should be about 100 words and briefly describe the contents of the paper including the mode or modes of transportation, the research method, and the key results and/or conclusions. Please include a list of **Keywords** to describe your article.

Graphic elements (figures and tables) must be called out in the text. Graphic elements must be in black ink. We will accept graphics in color only in rare circumstances.

References follow the style outlined in the *Chicago Manual of Style*. All non-original material must be sourced.

International papers are encouraged, but please be sure to have your paper edited by someone whose first language is English and who knows your research area.

Accepted papers must be submitted electronically in addition to a hardcopy (see above for information on electronic submissions). Make sure the hardcopy corresponds to the electronic version.

Page proofs: As the publication date nears, authors will be required to proofread and return article page proofs to the Managing Editor within 48 hours of receipt.

Acceptable software for text and equations is limited to Word and LaTeX. Data behind all figures, maps, and charts must be provided in Excel, Word, or DeltaGraph (unless it is proprietary). American Standard Code for Information Interchange (ASCII) text will be accepted but is less desirable. Acceptable software for graphic elements is limited to Excel, DeltaGraph, or Adobe Illustrator. If other software is used, the file supplied must have an .eps or .pdf extension. We do not accept PowerPoint.

Maps are accepted in a variety of Geographic Information System (GIS) programs. Files using .eps or .pdf extensions are preferred. If this is not possible, please contact the Managing Editor. Send your files via email or on a CD via overnight delivery service.

Send all submission materials to:
Marsha Fenn, Managing Editor
Journal of Transportation and Statistics
BTS/RITA/USDOT
400 7th Street, SW, Room 4117
Washington, DC 20590
Email: marsha.fenn@dot.gov

U.S. Department of Transportation
Research and Innovative Technology Administration
Bureau of Transportation Statistics

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 8 Number 1, 2005
ISSN 1094-8848

CONTENTS