Volume 4 Number 1
April 2001
ISSN 1094-8848

BUREAU OF TRANSPORTATION STATISTICS UNITED STATES DEPARTMENT OF TRANSPORTATION

JOURNAL OF TRANSPORTATION AND STATISTICS

JOURNAL OF TRANSPORTATION AND STATISTICS

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 4 Number 1
April 2001
ISSN 1094-8848

BUREAU OF TRANSPORTATION STATISTICS   UNITED STATES DEPARTMENT OF TRANSPORTATION

U.S. DEPARTMENT OF TRANSPORTATION

NORMAN Y. MINETA
Secretary

MICHAEL P. JACKSON
Deputy Secretary

BUREAU OF TRANSPORTATION STATISTICS

ASHISH K. SEN
Director

RICHARD R. KOWALEWSKI
Deputy Director

SUSAN J. LAPHAM
Associate Director for Statistical Programs and Services

JOHN V. WELLS
Chief Economist

Bureau of Transportation Statistics

Our mission is to lead in developing transportation data and information of high quality and to advance their effective use in both public and private transportation decisionmaking.

Our vision for the future: Data and information of high quality will support every significant transportation policy decision, thus advancing the quality of life and the economic well-being of all Americans.

The *Journal of Transportation and Statistics* is published by the

Bureau of Transportation Statistics
U.S. Department of Transportation
Room 3430
400 7th Street SW
Washington, DC 20590
USA
journal@bts.gov

Subscription information

To receive a complimentary subscription or a BTS Product Catalog:

mail	Product Orders Bureau of Transportation Statistics U.S. Department of Transportation Room 3430 400 7th Street SW Washington, DC 20590 USA
phone	202.366.DATA
fax	202.366.3640
internet	www.bts.gov

Information Service

email	answers@bts.gov
phone	800.853.1351

Cover and text design Susan JZ Hoffmeyer
Layout Dorinda Edmondson
Cover photo ©Getty Images/Eyewire

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 4 Number 1
April 2001

Contents

Letter to the Editor

I want to challenge the conventional wisdom concerning the post-WWII decline in urban transit ridership. It has long been held that the fall in transit ridership from 1946 forward could be explained by residential suburbanization, increased automobile commutation, low gasoline prices, as well as the decline of the central business district, among other factors. Suppose, however, that if account were taken of school bus "commutation," the decline turned out to be significantly less than reported.

Consider the case of Boston. There was a school car network operating all across the city up to the late 1940s. These school cars would pick up students headed to the seven central high schools from neighborhoods all over the city, using routes developed for that purpose. My memory of how we Latin School boys fought our way into the overcrowded cars is still quite vivid.

This special streetcar network operated by the Boston Elevated Railway was gradually replaced by school buses operated by the School Committee. I suspect that in Boston and elsewhere this change figures in the decline of public transit ridership.

Just as much to the point, by how much would U.S. urban transit ridership numbers increase if school bus ridership were added into the equation?

In any case, would a transit ridership statistical time series for Boston show such a steep postwar decline if it were adjusted for the changing character of the school-boy and school-girl commute?

CHARLES J. STOKES

Charles Anderson Dana Professor of Economics, Emeritus,
University of Bridgeport,
quondam Senior Fellow, The Brookings Institution, and
Director of Case Studies in Transportation

Assessing the Impact of Speed-Limit Increases on Fatal Interstate Crashes

SANDY BALKIN
Informed Analytics Group

J. KEITH ORD
Georgetown University

ABSTRACT

This study investigates the relationship between speed-limit increases and increases in the number of fatal crashes on U.S. rural and urban interstates. Past studies use expected historical trends to support claims that "speed kills." Using structural modeling, we assess the change in the average of the time series after a known change in speed limit occurs. The analysis is carried out separately for urban and rural interstates for each state. The results cast doubt on the blanket claim that higher speed limits and higher fatalities are directly related. After the initial speed-limit increases in 1987, the number of fatal accidents on rural interstates increased in some states but not in all. The 1995 round of speed-limit increases generally showed smaller increases in fatalities on rural interstates and slight to no increase on urban interstates. The approach also allows identification of seasonal effects that vary across the states.

INTRODUCTION

The relationship between the speed limit and the number of traffic-related fatalities is a subject of great interest to insurance companies, to the government at all levels, and to the general public. Historically, the government has taken an active role in the determination of speed limits, starting

J. Keith Ord, Georgetown University, McDonough School of Business, 320 Old North, Washington, DC 20057. Email: ordK@msb.edu.

with the establishment of the National Maximum Speed Limit (NMSL) by Congress in January of 1974. Prior to this legislation's setting the maximum speed limit at 55 miles per hour (mph), many states posted limits as high as 70 to 75 mph. In April of 1987, Congress passed legislation allowing states to increase speed limits to 65 mph on qualifying sections of interstate highways in rural areas with populations of less than 50,000. Within a few months, 38 states raised the speed limits on appropriate roads. More recently, the National Highway System (NHS) Designation Act of 1995 was signed into law on November 28, 1995. This act ended the federal government's involvement in the establishment of speed limits, putting the responsibility for speed-limit designation and compliance in the hands of the state governments. In most cases, state governments exercised their new rights and raised speed limits on rural and urban interstates.

The purpose of this study is to investigate the relationship between speed limits and traffic-related fatalities. Specifically, we aim to answer the question *Does an increase in the speed limit result in a higher incidence of fatal crashes?* Using a technique known as structural modeling, we are able to determine the impact speed-limit changes in the past have had on the number of fatal crashes on rural and urban interstates for each state based on its own past experiences. This method also gives information about the seasonal patterns in the number of fatal crashes.

The paper is organized as follows. The next section provides a review of the literature on the effects of speed-limit increases on the number of traffic-related crashes and fatalities. The third section presents the data and methodology used in this study. The fourth section demonstrates the analysis on a single state, and the fifth describes the results of the study for all states. The final section gives conclusions and perspectives for future work.

LITERATURE REVIEW

Various studies have attempted to determine the impact of speed-limit increases on the number of traffic-related crashes and fatalities. The following is a representative selection of the studies that motivated the study presented in this paper.

The report to Congress entitled "Effects of the 65 mph Speed Limit Through 1990" by the U.S. Department of Transportation (USDOT), National Highway Traffic Safety Administration (NHTSA) in May of 1992 looks at yearly interstate-fatality data split by rural and urban roadways. The analysis is based on "expected historical trends" (USDOT NHTSA 1992). These projected counts were derived from statistical models based on the historical relationship between rural-interstate fatalities and fatalities on other roadways. These results do not convey the impact of the speed-limit increase on traffic fatalities. Rather, the study relates interstate deaths to noninterstate deaths; it also assumes a stationary, or nonchanging, environment by fitting a global regression model. The authors then compare fatalities in 1986 with those in 1990 by computing percentage changes. This approach ignores historical trends and possible aberrant observations.

The paper does caution that care ought to be taken when interpreting the data. The authors note that results of individual states probably can not be generalized to the entire nation. They also point out that no statistical model is capable of controlling all factors affecting fatalities.

A 1997 paper entitled "Effect of 1996 Speed-Limit Changes on Motor Vehicle Occupant Fatalities" by Farmer, Retting, and Lund analyzes the effect speed-limit increases during and around 1996 had on interstates. This study employed linear regression models on trend and dummy variables to analyze the number of fatalities in states categorized by the time of their 1996 speed-limit increase (early, late, or none) and compares observed fatalities with projected values based on historical trends. They use percentage change between 1995 and 1996 to assess the impact of the 1996 legislation. They continually note that vehicle-miles traveled (VMT) may be able to explain the increase in fatalities but that the appropriate data are not available.

This study notes that while the national fatality toll for 1996 changed very little compared with 1995, the change in the fatality toll for individual states varied markedly between significant decreases and increases. The study also states that total interstate fatalities increased for the 11 states that had increased speed limits. The authors note that there has been an increase in the portion of the fatalities occurring on roads posted at 55 mph or greater and that some increase in fatalities on interstates is to be expected. Overall, this study presents a very thorough before and after comparison using percentage changes. A linear trend model with an intervention variable is used to compare actual 1996 fatalities with estimated 1996 fatalities based on historical trends. The restriction to annual data and the use of nonadaptive trends limit the value of the comparisons.

"The Effect of Increased Speed Limits in the Post-NMSL Era" is the title of another National Highway Traffic Safety Administration report to Congress (USDOT NHTSA 1998). This 1998 study also investigates the effect of the 1995 to 1996 speed-limit increases on rural and urban interstates.[1] It groups states into "changers" (12 count) and "nonchangers" (18 count) where the latter serve as a comparison for the former. The authors modeled the logarithms of fatality counts for each year during 1990 to 1996 as functions of time and type of state. Both linear and quadratic time variables were included. The impact of the speed increase was modeled using a dummy variable equal to one in 1996 and zero in previous years. They also included an interaction term between state group and the 1996 indicator to represent the difference between pre-1996/1996 changes for the two state types while accounting for the time trend. The authors claim that if this interaction term is significant, the 1996 departure from the time trend among the states that increased limits differs from the comparison states. The foremost problem with this analysis is that linear and quadratic trend models are not appropriate for these series. Including a quadratic trend may lower the residual variance for the in-sample fit, but it will damage the predictive ability of the model. Inspecting plots of the number of fatalities or fatal crashes shows that the addition of a global quadratic trend term typically does not provide a reasonable description for the whole length of the series.

Ledolter and Chan's article "Evaluating the Impact of the 65-mph Maximum Speed Limit on Iowa Rural Interstates" (1996) examines whether a significant change in the fatal and major-injury accident rates can be detected following the implementation of a higher speed limit on rural interstates in Iowa. The authors have access to quarterly data on traffic speed, traffic volume, and traffic safety. To answer the posed question, they fit a time-series intervention model relating number of accidents to traffic volume. They also include time trend, intervention variables for the May 1987 change, and quarterly seasonality. The authors find that expected numbers of fatal accidents in Iowa rose by two incidents per quarter on rural interstates, a statistically significant increase.

DATA AND METHODOLOGY[2]

The data used in the present study are the number of fatal crashes for each month from January 1975 to December 1998 for each state separated by rural and urban interstates. We used fatal crashes rather than number of deaths since we regard the accident data as a more reliable guide to road safety conditions. The number of fatal crashes was determined from the Fatality Analysis Reporting System (FARS) and is publicly available[3] and maintained by NHTSA. FARS provides monthly data on numbers of fatal crashes for each state with separate counts for rural and urban interstates.

The database was downloaded in *SAS*© format. It is possible to query the FARS database for yearly statistics for 1994 to 1998. Since our monthly values sum up to the yearly values reported by the online system, we are confident that we were able to successfully extract the appropriate data. Our yearly totals do not always exactly match the yearly totals given in the studies mentioned previously. These discrepancies can be attributed to the changing of the FARS database structure, to differences in opinion on which roadways were included, or to user error. Again, since our data set matches the online database query totals, we are satisfied with the quality of our data compilation.

We let y_t denote the number of fatal crashes that occur in month t and use time series models to examine the impact of an increase in speed limit on the number of fatal crashes. That is, we are mainly concerned with the modeling aspect of time series

[1] Specifically, only states with increases between December 8, 1995 and April 1, 1996 are considered.

[2] Further information regarding the data and their collection is available from the author.

[3] http://www-fars.nhtsa.dot.gov/

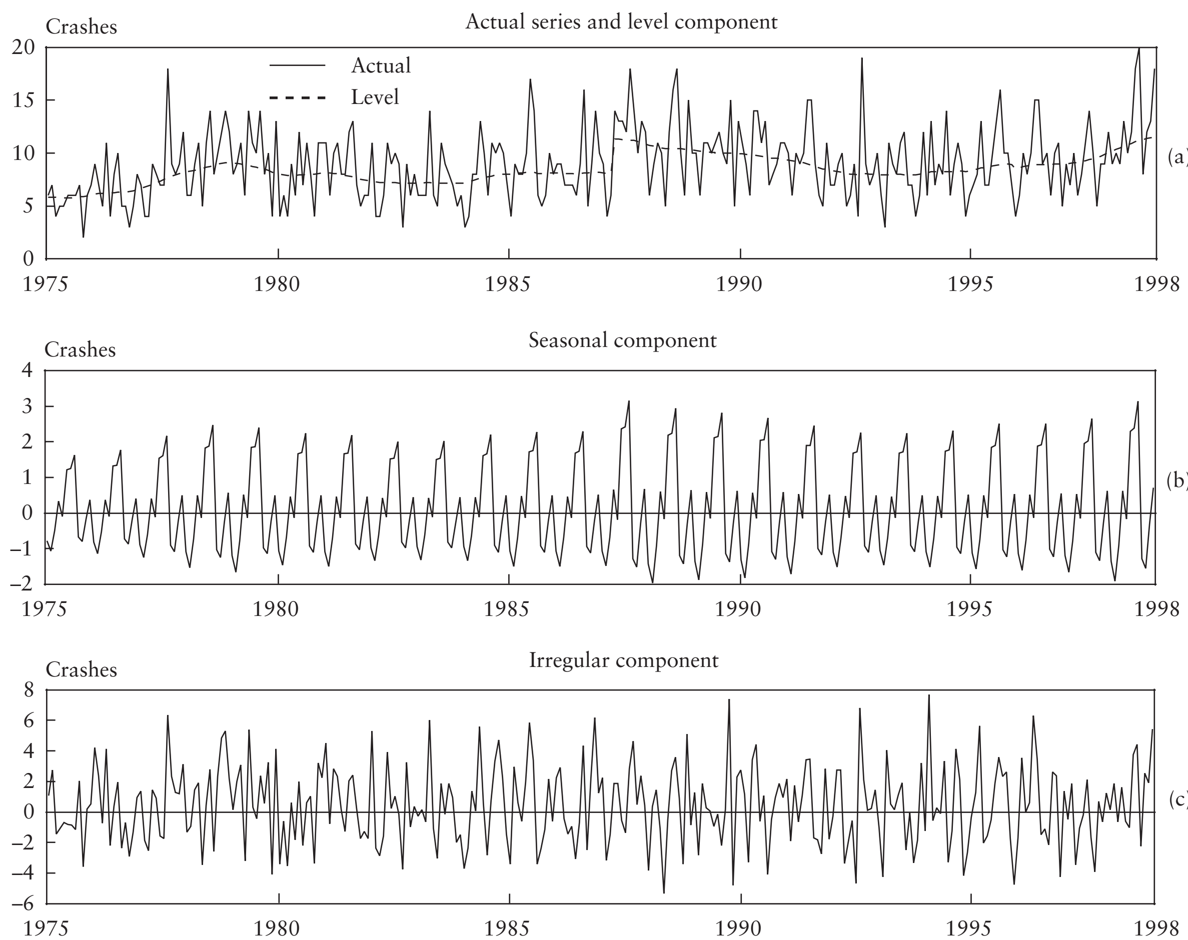
analysis, looking backwards in time for structural changes in the series. Since the accident data are collected over time in regularly spaced intervals and the timing of speed-limit changes is known, we use intervention analysis to examine these effects.

Intervention analysis is a time series technique used when a change in the environment occurs at a known time and affects the phenomenon of interest.[4] In this case, the known change is the speed limit. Since the change in speed limit is more or less permanent, a step intervention is most appropriate. We hypothesize that the change in speed limit results in a permanent shift in the number of accidents. To aid in the analysis and interpretation, we employ the logarithmic transformation. The use of logarithms allows us to consider percentage changes rather than absolute shifts and stabilizes the variance of the series. Since some of the months have zero fatal crashes, it is necessary to add one to each month prior to transforming the data. Thus, it is important to remember when looking at the plots of the data, as in figure 1(a), that the series is shifted up by one unit.

Motivated by some of the previous studies on this topic already discussed, we chose to employ a statistical modeling technique that could provide us with an explanation of the main features of the phenomena under investigation. Harvey and Durbin (1986) used structural time-series modeling to examine the effects of seat-belt legislation on British road casualties. In structural time-series modeling, models are set up explicitly in terms of the components of interest, such as trends, seasonals, and cycles. In addition, instead of assuming that these components remain constant over time, this approach allows them to evolve. The approach is intuitively appealing since environments that generate time series often do not remain constant

[4] See chapter 13 of Kendall and Ord (1990) or chapter 12 of DeLurgio (1998) for a description.

FIGURE 1 Rural Arizona Crashes



and an explicit description of how these components change can provide valuable insights.

The starting point for the construction of structural models is to represent an observed value as the sum of level, seasonal, and irregular components.

$$y_t = \mu_t + \gamma_t + \varepsilon_t,\ t = 1,\ldots,T,\ \varepsilon_t \sim NID\left(0,\sigma_\varepsilon^2\right) \quad (1)$$

where y_t, as previously defined, is the observation made at time t, which in our case is after a log transformation $\left(y_t = 1\mathrm{n}\left(x_t + 1\right)\right)$, and μ_t, γ_t, and ϵ_t are the level, seasonal, and irregular components. In this study, the level component is allowed to change according to a random walk process, and the seasonal component changes according to a trigonometric model. That is,

$$\mu_t = \mu_{t-1} + \eta_t,\quad \eta_t \sim NID\left(0,\sigma_\eta^2\right) \quad (2)$$

$$\gamma_t = \sum_{j=1}^{s/2} \gamma_{jt} \quad (3)$$

where, with s even (equal to 12 in this case) and $\lambda_j = 2\pi j / s$,

$$\begin{bmatrix} \gamma_{jt} \\ \gamma_{jt}^* \end{bmatrix} = \begin{bmatrix} -\cos\lambda_j & \sin\lambda_j \\ -\sin\lambda_j & \cos\lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_{jt} \\ \omega_{jt}^* \end{bmatrix},$$

$$j = 1,\ldots,\frac{1}{2}s - 1,\ \gamma_{jt} = \left(\cos\lambda_j\right)\gamma_{j,t-1} + \omega_{jt},\ j = \frac{1}{2}s,$$

and where the ω_{jt}s and the ω_{jt}^*s are both $NID\left(0,\sigma_\omega^2\right)$ and are independent of each other. This formulation allows the seasonal effects to vary over time.

It is possible to include a trend component within the level, but no such structure was found in any of the data used for this study, so it was omitted. Finally, in this study we are interested in testing whether a change in the speed limit results in a permanent change in the level of the number of fatal crashes for a given state on a given class of interstate. Thus, we can accommodate this type of analysis by extending the structural model to the form

$$y_t = \mu_t + \gamma_t + \lambda z_t + \varepsilon_t \quad (4)$$

where ε_t, η_t, and ω_t are mutually independent of each other, and each has zero mean and constant variance and is also serially independent. Formally, we write this as

$\varepsilon_t \sim NID\left(0,\sigma_\varepsilon^2\right)$, $\eta_t \sim NID\left(0,\sigma_\eta^2\right)$, $\omega_t \sim NID\left(0,\sigma_\omega^2\right)$

and

$cov\left(\varepsilon_t,\eta_t\right) = cov\left(\varepsilon_t,\omega_t\right) = cov\left(n_t,\omega_t\right) = 0$ for all t.

We refer to z_t as the intervention variable, defined as:

$$z_t = \begin{cases} 0, & t < \tau \\ 1, & t \geq \tau \end{cases} \quad (5)$$

Thus, z_t takes on a value of zero up until time τ, the month and year of the known speed limit change. The overall fit of the model might be improved by searching for possible interventions rather than pre-specifying their timing. Indeed, there may be a time lag before drivers adapted to the new limits. We decided to retain the more conservative strategy of using the timing of the legal changes and considering only pure level shifts at those times. With regard to potential time lags, the variables defined by equation (5) would differ for only one or two months. If a substantial effect exists, it would still be detected. As for the host of other potential interventions, we preferred to focus solely on the impact of speed-limit changes and to avoid concerns about mining the data. When the component parameters of the structural model are estimated, the intervention parameter λ can be used to assess the impact of the speed-limit change. The value $100 \times \left(\exp(\lambda) - 1\right)$ approximates the percentage increase in the number of fatal crashes after the speed limit was exposed. The exact value is more complex as a result of using the transformation $\ln(x_t + 1)$ rather than $\ln(x_t)$; the differences are slight unless the mean level is very low when percentage changes are rather unreliable anyway. The computer package STAMP 5.0, developed by Harvey and his associates, was used to perform the analyses presented in this study.

EXAMPLE: RURAL ARIZONA

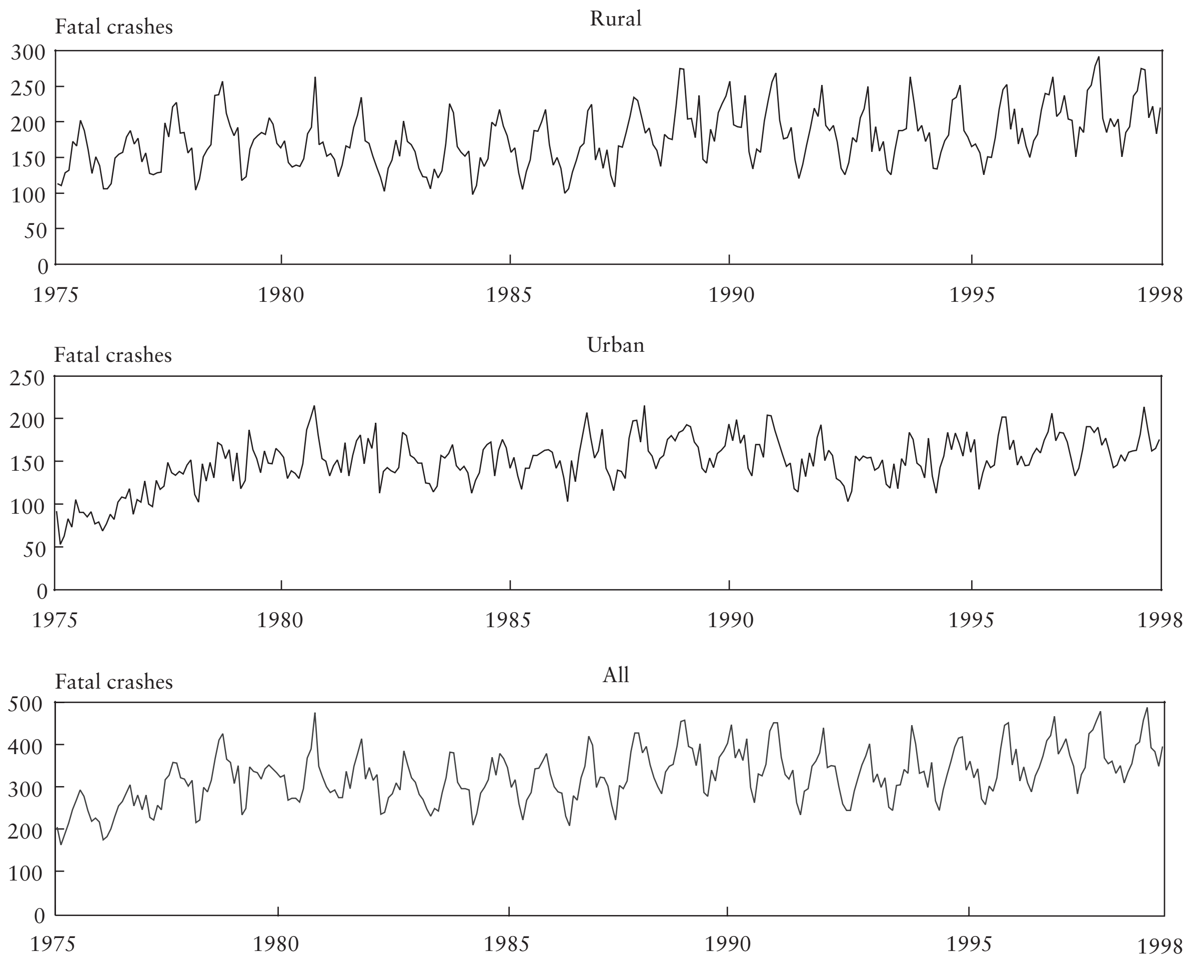
As an example of this method of analysis, consider rural interstates in Arizona. The speed limit was changed in April 1987 and December 1995. Thus, an intervention variable was specified for each of these months, defined as in equation (5). The original time series is shown in figure 1(a).

The series is decomposed into level, seasonal, and irregular components represented graphically in figure 1, panels a, b, and c, after transformation back to the original units. We see a significant increase in the level around 1987 but none around 1995. This indicates that around 1987 the average number of fatal crashes significantly increased, but not so elsewhere. This increase occurs at the same time as a speed-limit increase. Statistically, it is estimated that the 1987 speed-limit increase resulted in a 41% increase in rural interstate crashes in Arizona (see table 1). There is no statistical evidence that the

TABLE 1 Significant Changes in Predicted Accident Rates Attributed to the Speed-Limit Increases on Rural Interstates

State	First change, percent	Second change, percent	1986 crashes	1987 crashes	1988 crashes	1995 crashes	1996 crashes	1997 crashes	1998 crashes
Alabama		24.8	62	53	57	49	62	79	65
Arizona	41.0		96	126	116	100	103	98	140
Arkansas	32.6		26	30	34	28	38	28	46
Florida		37.2	105	96	153	102	116	166	133
Georgia		30.0	66	52	65	50	100	73	85
Illinois	21.9		48	59	73	58	68	71	74
Iowa	35.8		12	21	29	20	21	22	21
Kansas	23.1		17	22	24	27	18	23	14
Maine	18.4		13	8	20	9	8	12	14
Maryland		37.4	20	17	19	16	15	26	18
Michigan	46.7		14	34	44	34	46	40	46
Minnesota	25.7		9	19	21	15	15	12	31
Missouri	13.0	42.2	58	53	54	66	87	93	98
Nebraska	35.5		12	12	20	19	21	22	28
Nevada		27.1	22	35	32	44	53	44	62
New Hampshire	21.4		5	6	19	6	10	10	8
New Mexico	15.8	25.5	66	103	78	87	85	105	102
North Carolina	42.6		40	58	76	55	71	50	53
Ohio	46.6		34	54	47	45	46	44	39
Pennsylvania		36.4	56	68	69	52	64	54	67
Tennessee	15.5	40.4	53	70	73	74	86	70	97
Texas		18.0	164	195	236	201	220	217	200
Virginia	31.6		39	39	64	74	59	71	62
Washington	24.5		26	32	40	35	43	30	39
West Virginia	46.2		14	13	29	38	48	40	34
Wisconsin	24.3		15	13	24	16	27	25	20
U.S. totals			**1,834**	**2,141**	**2,391**	**2,210**	**2,441**	**2,518**	**2,591**
Percentage change				**16.74**	**11.68**		**10.45**	**3.15**	**2.90**

FIGURE 2 Aggregate Fatal Crashes: Rural, Urban, and All

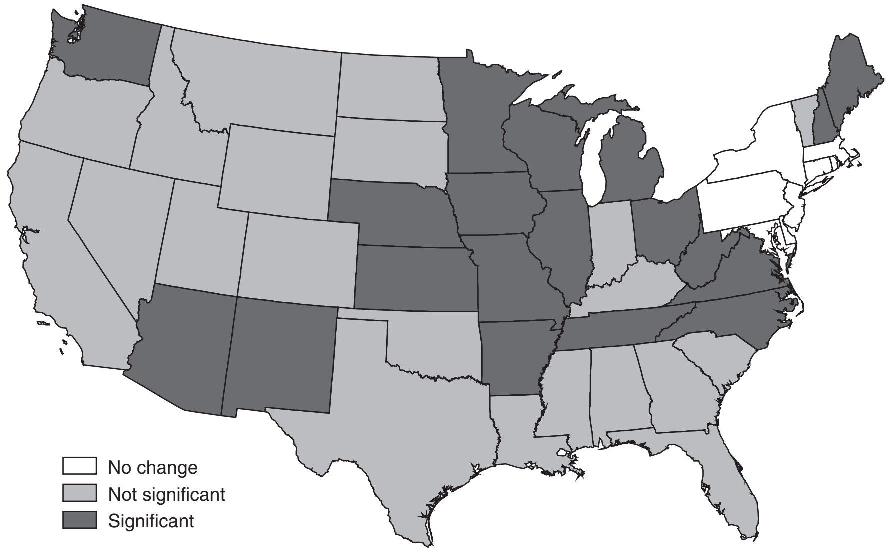


1995 speed-limit increase had any additional effect on the number of fatal crashes. This may change as more observations become available, better defining the impact of the policy change.

Next, we see from the seasonal component that there appears to be a strong monthly effect on the number of fatal crashes. For this example, there are considerably more crashes in June, July, and August compared with the other months. Such seasonal patterns exist for most states and reflect the higher traffic levels in summer months. The irregular component is simply what is left over after the level and seasonal components are taken into account.

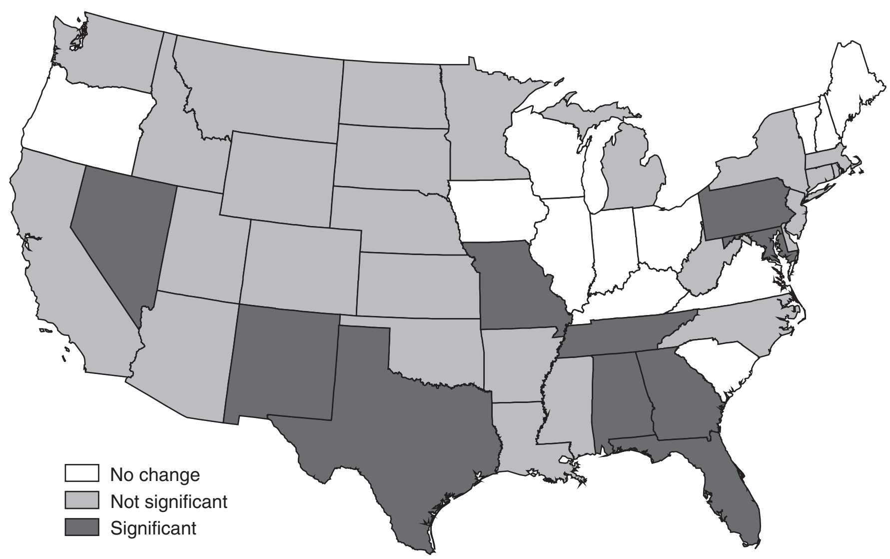
The Structural Time Series Modeling approach tells us that there is a strong seasonal effect on the number of fatal crashes and that there is a significant increase in the number of such crashes around the time the speed limit was changed. For this particular series, the plot of the level component suggests that, after the initial jolt of the speed-limit change, the trend gradually moves back to its original level. This phenomenon was observed for a number of states, but not all. Such a movement would be consistent with the slight overall decline in fatal accidents nationally over this time period, as shown in figure 2. This observation is made tentatively, since partial adjustment effects were neither modeled nor tested. The picture is further complicated as state laws were enacted at different times. Nevertheless, we view this as a question worthy of further exploration since several distinct hypotheses exist, with quite different policy implications. Such hypotheses include 1) drivers adjusted to driving at higher speeds, 2) states increased enforcement of driving laws, and 3) automobile safety was improved. However, we stress that our analysis was not designed to examine these

FIGURE 3 Significance Levels of Responses to the First Speed-Limit Change (1987) on Rural Interstates



Note: Signficant at the 10% level

FIGURE 4 Significance Levels of Responses to the Second Speed-Limit Change (1996) on Rural Interstates



Note: Signficant at the 10% level

questions; rather, they are important issues for further investigation.

STATISTICAL ANALYSIS

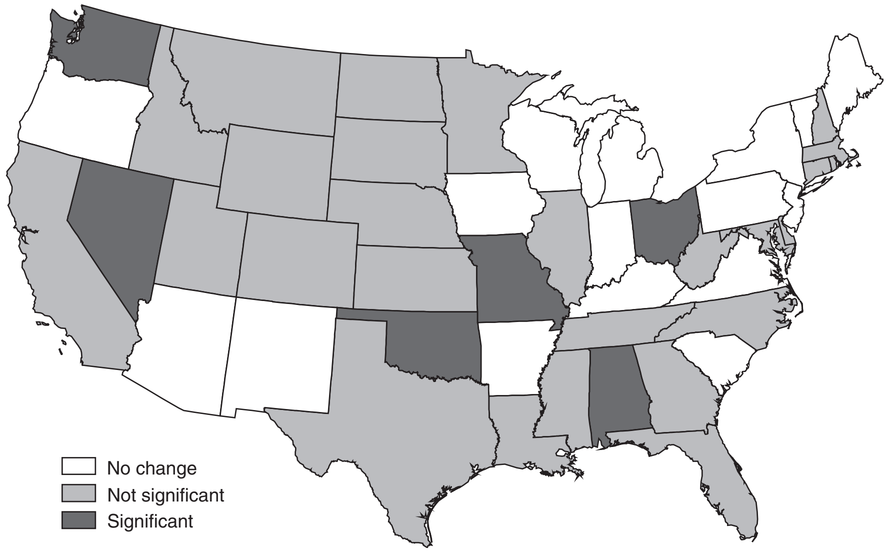
Each state's rural and urban interstates were analyzed using the structural modeling approach with deterministic step intervention variables at the time(s) of the speed-limit increases. Rural interstates are subject to 1987 and 1996 changes, while urban interstates were only changed around 1996. We will refer to the changes around 1987 as the FIRST speed-limit increases and those made around 1996 as the SECOND speed-limit increases.

Results for Individual States

We can summarize the findings as follows:

- 19 of 40 states experienced a significant increase in fatal crashes along with the FIRST speed-limit increases on rural interstates (figure 3).
- 10 of 36 states experienced a significant increase in fatal crashes along with SECOND speed-limit increases on rural interstates (figure 4).

FIGURE 5 Significance Levels of Responses to the Speed-Limit Change (1996) on Urban Interstates



Note: Signficant at the 10% level

- 6 of 31 states experienced a significant increase in fatal crashes along with the speed-limit increases on urban interstates (figure 5).

Table 1 shows the states with significant changes on rural interstates, the estimated monthly percentage impact of the speed-limit change, and the numbers of fatal crashes in 1986 to 1988. From this table, we can see the monthly percentage increase in the number of fatal crashes attributable to the speed-limit changes. The numbers of total fatal crashes for 1986 to 1988 are included for two reasons: 1) to interpret the percentages in terms of real numbers and 2) to see if the number of fatal crashes increases in the year after the speed-limit change. The purpose behind the first reason is to see, without minimizing the value of human life, what the significant increase translates to in terms of actual number of crashes. For example, suppose a state averages 36 crashes per year, or 3 per month, and the estimated monthly increase of fatal crashes is about 33%. The expected increase in the number of crashes is about one per month. Although statistically significant, such an increase is small in absolute numbers and may be attributable to other factors. The purpose behind the second reason is to assess whether drivers gradually adjust to new driving conditions. For example, Arizona, as graphically displayed in figure 1, had an increase in the number of crashes the year of the speed-limit change but a decrease from that level in subsequent years. This suggests that drivers in Arizona may have learned how to drive safely at the new limit. Such patterns are not consistent across states, and this issue requires further investigation.

Table 2 shows the same information for the urban interstates and also includes 1998 data but includes only states that experienced a statistically significant increase in the number of fatal crashes. During the 1996 set of changes, some states encountered a negative impact, a decline in the number of fatal crashes after the speed-limit increase. While this effect may be real, it is difficult

TABLE 2 Significant Changes in Predicted Accident Rates Attributed to the Speed-Limit Increases on Urban Interstates

State	Change, percent	1995 crashes	1996 crashes	1997 crashes	1998 crashes
Alabama	37.8	32	49	66	49
Missouri	35.0	56	85	82	94
Nevada	31.2	14	23	15	14
Ohio	40.8	60	64	64	78
Oklahoma	39.5	36	47	46	31
Washington	32.1	24	33	34	40
U.S. totals		**1,919**	**2,054**	**1,998**	**2,026**
Percent change			**7.03**	**–2.72**	**1.40**

to attribute it to the increase in speed limits. Therefore, the results are not included in table 2.

In order to get an idea of how many fatal crashes are associated with a particular speed-limit increase, we first remove from the fitted model the term relating to the increase for those states that had a significant increase in fatal crashes. We then analyze the difference between the modified expected and actual numbers. We approximate the predicted number of fatal crashes had the speed limit not increased by dividing the observed number of fatal crashes by one plus the percent change. Tables 3 and 4 show this information separated by rural and urban interstates. We see that the

TABLE 3 Predicted Number of Fatal Crashes Attributed to the Speed-Limit Increase on Rural Interstates

State	First, percent	Second, percent	Predicted 1988	Predicted 1996	Predicted 1997	Predicted 1998
Alabama	0.0	24.8	57.0	49.7	63.3	52.1
Arizona	41.0	0.0	82.3	103.0	98.0	140.0
Arkansas	32.6	0.0	25.6	38.0	28.0	46.0
Florida	0.0	37.2	153.0	84.5	121.0	96.9
Georgia	0.0	30.0	65.0	76.9	56.2	65.4
Illinois	21.9	0.0	59.9	68.0	71.0	74.0
Iowa	35.8	0.0	21.4	21.0	22.0	21.0
Kansas	23.1	0.0	19.5	18.0	23.0	14.0
Maine	18.4	0.0	16.9	8.0	12.0	14.0
Maryland	0.0	37.4	19.0	10.9	18.9	13.1
Michigan	46.7	0.0	30.0	46.0	40.0	46.0
Minnesota	25.7	0.0	16.7	15.0	12.0	31.0
Missouri	13.0	42.2	47.8	61.2	65.4	68.9
Nebraska	35.5	0.0	14.8	21.0	22.0	28.0
Nevada	0.0	27.1	32.0	41.7	34.6	48.8
New Hampshire	21.4	0.0	15.7	10.0	10.0	8.0
New Mexico	15.8	25.5	67.4	67.7	83.7	81.3
North Carolina	42.6	0.0	53.3	71.0	50.0	53.0
Ohio	46.6	0.0	32.1	46.0	44.0	39.0
Pennsylvania	0.0	36.4	69.0	46.9	39.6	49.1
Tennessee	15.5	40.4	63.2	61.3	49.9	69.1
Texas	0.0	18.0	236.0	186.4	183.9	169.5
Virginia	31.6	0.0	48.6	59.0	71.0	62.0
Washington	24.5	0.0	32.1	43.0	30.0	39.0
West Virginia	46.2	0.0	19.8	48.0	40.0	34.0
Wisconsin	24.3	0.0	19.3	27.0	25.0	20.0
Predicted total			**1,317.3**	**1,329.3**	**1,314.4**	**1,383.2**
Actual total			**1,516.0**	**1,530.0**	**1,525.0**	**1,596.0**
Approximate increase			**198.7**	**200.7**	**210.6**	**212.8**
Percentage increase			**15.09**	**15.10**	**16.02**	**15.39**

TABLE 4 Predicted Number of Fatal Crashes Attributed to the Speed-Limit Increase on Urban Interstates

State	Change, percent	Predicted 1996	Predicted 1997	Predicted 1998
Alabama	37.8	35.6	47.9	35.6
Missouri	35.0	63.0	60.7	69.6
Nevada	31.2	17.5	11.4	10.7
Ohio	40.8	45.5	45.5	55.4
Oklahoma	39.5	33.7	33.0	22.2
Washington	32.1	25.0	25.7	30.3
Predicted total		**220.2**	**224.2**	**223.8**
Actual total		**301.0**	**307.0**	**306.0**
Approximate increase		**80.8**	**82.8**	**82.2**
Percentage increase		**36.71**	**36.91**	**36.75**

estimated overall percentage increases are of the same order as the individual increases, resulting in approximately an additional 200 rural and 80 urban fatal crashes per year. It is important to note that these numbers only represent a crude approximation of the effect of the speed-limit increase.

Seasonality

One of the powerful benefits of using structural modeling is that instead of removing seasonality, the effect of a specific month is directly modeled. The strength of the seasonal pattern was one of the most surprising aspects of this analysis. Figures 6 and 7 show the following:

- 29 states exhibited seasonality at the 0.05 level of significance on rural interstates (figure 6)
- 18 states exhibited seasonality at the 0.05 level of significance on urban interstates (figure 7).

The extent of seasonality varies by state. Most states typically have a higher number of fatal crashes in August. Some states have different patterns with interpretations unique to that state. For instance, Florida tends to have more fatal crashes in March on its urban interstates. One possible interpretation of this could be the increase of traffic from college students traveling to Florida on spring break. In general, seasonal peaks appear to coincide with peak holiday seasons. Most states do not produce monthly data on vehicle-miles traveled, so we cannot adjust the data in a consistent manner for such effects.

Aggregate Analysis

Though the analysis is by state, it is of interest to generalize the effect of speed limit increases to the nation as a whole. To answer this question, we use a "Super t-Test." We first record the t-values of the intervention variables for all states. Positive t-values indicate a positive impact (increase) of the number of fatal crashes. Of fatal crashes, they determine the significance of the individual impact of the policy change. To answer the question whether or not fatal crashes increase along with speed-limit increases, we then perform a one-sided t-test to determine whether the mean of the t-values of all of the intervention variables is significantly greater than zero. If we reject the null hypothesis, we can conclude that there is indeed an increase in the number of fatal crashes. It does not tell us, however, how large this increase is, only if, on average, an effect exists.

The Super t-Test for Rural Interstates resulted in a t-statistic of 10.6 with 39 degrees of freedom (one-tailed p-value $\approx$ 0.000) for the FIRST set of speed-limit changes and a t-statistic of 4.0 with 36 degrees of freedom (one-tailed p-value = 0.0002) for the SECOND set of speed-limit changes. For urban interstates, the Super t-test gave a t-statistic of 1.373 with 30 degrees of freedom (one-tailed p-value = 0.090). We see from the Super t-Tests that rural interstates appear to be affected by speed-limit increases, while the effect for urban interstates is weak.

CONCLUSION AND FUTURE WORK

The purpose of the study is to investigate the relationship between speed limits and traffic-related fatalities. Specifically, we sought to discover if an increase in the speed limit results in a higher incidence of fatal crashes.

We carried out the data analysis using a time-series technique known as structural modeling. This approach enables us to partition a series into its level, trend, seasonal, and irregular (or residual) components and to evaluate the impact of major interventions such as speed-limit changes. Based on

FIGURE 6 Significance Levels of Seasonal Components for Fatal Accidents on Rural Interstates

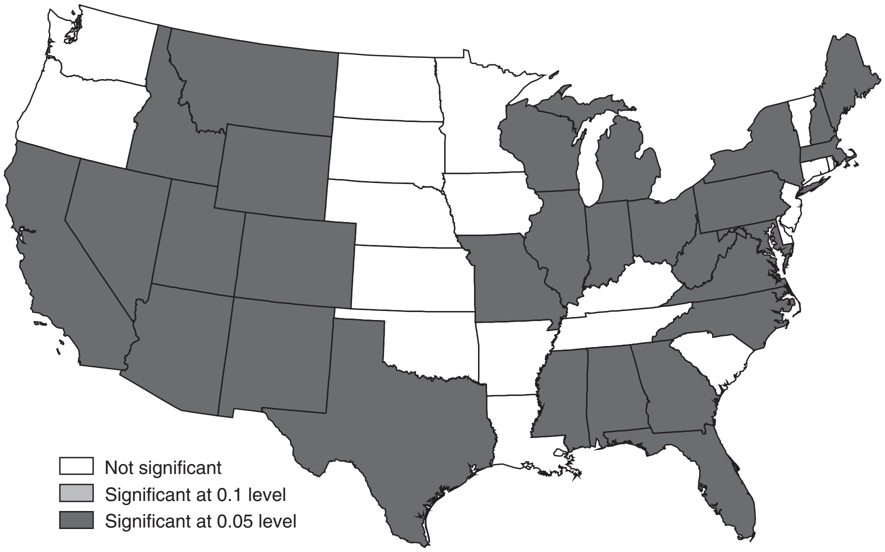
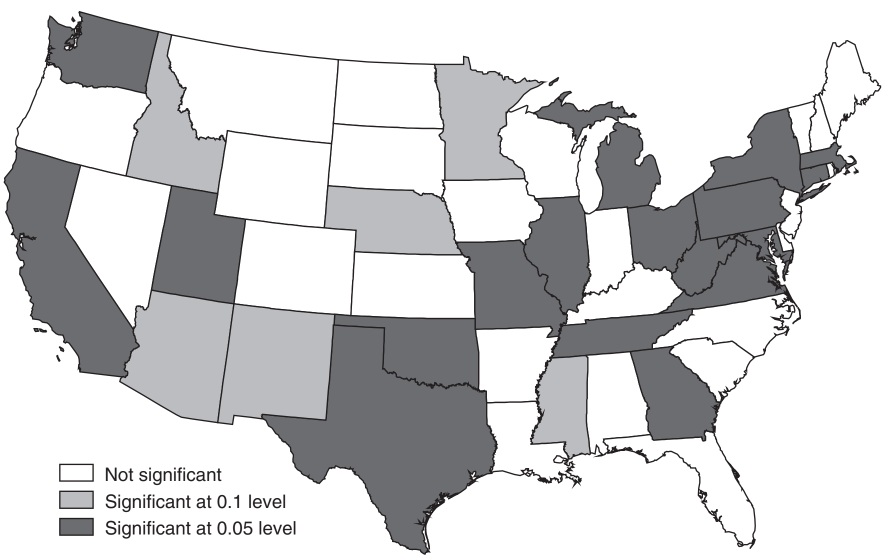


FIGURE 7 Significance Levels of Seasonal Components for Fatal Accidents on Urban Interstates



a review of the past literature, we formulated the impact of a speed-limit change as a one-time percentage increase in the number of accidents, after which the seasonal and trend patterns in the series would be expected to remain similar to those of past years. The analysis was performed for each state, separately for urban and rural interstates. Although the results are statistically significant as noted above, the numbers in some states may be small. The seasonal patterns probably reflect changes in the number of vehicle-miles traveled (VMT), with peaks occurring during holiday seasons. Seasonal analysis is critical to understanding any changes in pattern since unadjusted comparisons for a few months immediately before and after a change could be seriously in error. Our analysis allows comparisons to be made after proper adjustment for seasonal effects. Overall, increases were seen in some states following speed-limit changes. These increases were predominantly on rural rather than urban interstates.

ACKNOWLEDGMENTS

We wish to acknowledge the sponsorship of the Consumers Union for this research. The views expressed in this paper are those of the authors and

are not to be attributed in any way to the Consumers Union. We also wish to thank Tom Wassel of AstraZeneca, Eric Falk of Ernst & Young, and Eric Rosenberg of Consumers Union for data collection and analysis assistance.

REFERENCES

DeLurgio, S.A. 1998. *Forecasting Principles and Applications.* Boston: Irwin McGraw-Hill.

Farmer, C.M., R.A. Retting, and A.K. Lund. 1997. Effect of 1996 Speed-Limit Changes on Motor Vehicle Occupant Fatalities. *Insurance Institute for Highway Safety,* Washington, DC.

Harvey, A. and J. Durbin. 1986. The Effects of Seat Belt Legislation on British Road Casualties. *Journal of the Royal Statistical Society, Series* A 149:187–227.

Kendall, M. and J.K. Ord. 1990. *Time Series.* London: Edward Arnold.

Ledolter, J. and K.S. Chan. 1996. Evaluating the Impact of the 65-mph Maximum Speed Limit on Iowa Rural Interstates. *The American Statistician* 50, 1.

U.S. Department of Transportation (USDOT), National Highway Traffic Safety Administration (NHTSA). 1992. Effects of the 65-mph Speed Limit Throughout 1990. Report to Congress.

_____. 1998. The Effect of Increased Speed Limits in the Post-NMSL Era. Report to Congress.

Discussion

JOHANNES LEDOLTER
University of Iowa

I congratulate the authors for a very careful, statistically sophisticated, and impartial study on the impact of recent speed-limit changes on fatal interstate crashes. The findings have important policy implications as states face considerable pressure to increase maximum speed limits.

Any assessment of the impact of maximum speed-limit changes on traffic safety is difficult, and many reports have been written on this subject. The Balkin/Ord paper is an important contribution to this literature since it is comprehensive and current, covering all 50 states through 1998. Its findings are 1) the 1987 speed-limit change increased fatal accidents on rural interstates by about 200 crashes each year and the 1996 change added another 200 fatal crashes annually, for a combined total of 400 fatal crashes per year and 2) the impact of the 1996 change on the number of fatal accidents on urban interstates was not as strong, amounting to about 80 fatal crashes per year. The benefits of accelerated interstate travel come at the expense of safety though not all states are affected equally.

My comments on this paper have two purposes. First, the structural time series models which Balkin/Ord uses for characterizing the serial correlation among successive observations may not be familiar to readers of this journal. My comments address the relationship between these models and the more familiar Box-Jenkins ARIMA time series intervention models. Second, the Balkin/Ord study deals with a nationwide analysis of data from all 50 states. It is understandable that such an analysis can not be as detailed as studies that focus on specific states. My comments offer recommendations for model improvements, in particular suggestions for incorporating traffic volume and for using actual travel speeds instead of speed-change indicators.

Balkin/Ord uses structural time series intervention models for assessing the impact of maximum-speed-limit changes. This differs from other studies that use the ARIMA time series intervention models proposed by Box and Tiao (1975). The following discussion illustrates that these two model

Johannes Ledolter is the John F. Murray Research Professor of Management Sciences at the University of Iowa and Professor of Statistics at the Vienna University of Economics and Business Administration in Austria. He is a Fellow of the American Statistical Association and an elected member of the International Statistical Institute. Address: Department of Management Sciences, Tippie College of Business, University of Iowa, Iowa City, IA 52242. Email: johannes-ledolter@uiowa.edu.

families (structural time series intervention models and ARIMA time series intervention models) are closely related. Let us ignore, without loss of generality, the seasonal component in the structural model in equation (4),

$$y_t = \mu_t + \lambda z_t + \varepsilon_t \text{ with } \mu_t = \mu_{t-1} + \eta_t$$

After taking successive differences, $(1-B)y_t = y_t - y_{t-1}$ where B is the backshift operator $By_t = y_{t-1}$, we obtain

$$(1-B)y_t = \lambda(1-B)z_t + \varepsilon_t - \varepsilon_{t-1} + \eta_t = \lambda(1-B)z_t + n_t$$

The noise $n_t = \varepsilon_t - \varepsilon_{t-1} + \eta_t = (1-\theta B)a_t$ follows a first order moving average process.[1] The solution of the above difference equation is

$$y_t = \mu + \lambda z_t + \left[(1-\theta B)/(1-B)\right]a_t$$

This is like a regression of the number of fatalities on the indicator variable z_t with one important difference. The errors $\left[(1-\theta B)/(1-B)\right]a_t$ are no longer independent and follow an integrated first order moving average [ARIMA(0,1,1)] model, a very common, nonstationary time series model. Box and Tiao (1975) studies the estimate of the intervention effect λ under this particular model and show that it is a contrast between two weighted averages, one of observations before the intervening event and the other of observations afterwards. The weights are symmetric and decay exponentially according to the time distance of the observations from the intervening event. The rate of decay depends on the moving average parameter θ or, in terms of the parameters in the structural model, on the ratio of the variances of ε_t and η_t. Note that this estimate differs from the ordinary regression estimate in the model with independent errors, which is the difference of two (unweighted) averages; the observations before and after the intervening event are weighted equally. The equal weighting is inappropriate if observations are autocorrelated, and the analysis must be adjusted for the serial correlation in the observations. This adjustment can be achieved through ARIMA models as proposed by Box/Tiao or through the structural time series approach adopted by Balkin/Ord. Both approaches should lead to similar conclusions.

Balkin/Ord analyzes the number of traffic fatalities but fails to incorporate in its model a measure of risk exposure. Risk exposure can be measured through vehicle-miles traveled (VMT). Most states do have reliable estimates of traffic volume, typically obtained by sampling traffic flow at various continuous measurement stations spread throughout the state. An analysis of the ratio, the number of fatal accidents divided by VMT, is more meaningful since it adjusts accident numbers for traffic volume. The decreasing time trend in these ratios expresses the safety improvements of cars and roads.

The 1987 maximum-speed-limit increase was uniform across states. About one fifth of the states, mostly in the East, decided not to raise the maximum speed limit on rural interstates. The other states increased the maximum speed limit on rural interstates by 10 miles per hour (mph), from 55 mph to 65 mph. The response to the National Highway System Designation Act in 1995 was more varied. Appendix A of the National Highway Traffic Safety Administration Report to Congress (USDOT NHTSA 1998) gives a good overview of how states responded. About one fifth of the states did not increase the speed limit on rural interstates beyond the prior 65 mph limit. Most states raised the speed limits on rural interstates by 5 mph, to 70 mph. However, several other, mostly western, states raised it by 10 mph, to 75 mph. The fact that not all speed-limit changes were of the same magnitude was not incorporated in the analysis. The use of a single intervention dummy variable appears to be an oversimplification; it may have been better to incorporate the magnitude of the change.

Balkin/Ord models the speed-limit changes through 0/1 indicator variables. The model could be improved by including the actually traveled speeds. Admittedly, it would be difficult to obtain reliable traffic-speed data, certainly more difficult than obtaining reliable VMT data. However, many states do collect data on traffic speed. Iowa, for example, takes 24-hour measurements on 1 day

[1] See Abraham and Ledolter (1983) or Box, Jenkins, and Reinsel (1994).

each quarter at 4 rural interstate and 2 urban interstate stations. In earlier papers, Ledolter and Chan (1994; 1996) analyzed average traffic speeds, as well as the proportion of cars exceeding 55, 60, and 65 mph. Iowa had increased the maximum speed limit on its rural interstates from 55 mph to 65 mph in May of 1987. The average traffic speed on rural interstates did not change abruptly but increased gradually from about 59 mph in 1985/1986 to about 66 mph in 1990/1991. Our study shows that traffic speed does not change abruptly with the passing of a new rule but adapts gradually over a period of several quarters. The average actual travel speed (or a certain percentile of the distribution) is more indicative of driver behavior than the posted change in the maximum speed limit, making models that incorporate the actually traveled speeds preferable.

Several states also raised the maximum speed limit on rural primary roads. The safety impact of this speed-limit change is not reflected in the Balkin/Ord study since their analysis focuses solely on interstates. Its impact is found among the numbers of fatal accidents on the rural primary system. The impact may be substantial since most fatal accidents occur on rural noninterstates.[2] An investigation of the number of fatal accidents on rural primary roads, especially for those states that raised the speed limits on these road systems, is needed. Furthermore, speed increases may carry over to road systems not subject to the increased speed-limits. In our earlier analyses of the 1987 Iowa speed limit change on rural interstates, Ledolter and Chan (1994) and Ledolter and Chan (1996), we found small increases in the average actual travel speed on road systems that remained subject to the 55 mph limit. Hence, studies on the number of noninterstate fatal crashes for states that raised maximum speed limits on rural interstates but not on rural primary roads are also needed.

A brief comment on seasonality: for purposes of impact assessment, seasonality is a nuisance variable that needs to be excluded; seasonality by itself is of little interest. The number of fatal accidents is seasonal because traffic volume is seasonal. In addition, the number of fatal accidents per (million) VMT is seasonal because the risk of getting into an accident depends on seasonal weather and road conditions. Many studies model seasonality by including seasonal indicator variables or trigonometric functions (harmonics) of the seasonal frequency. The structural noise model in the Balkin/Ord paper goes one step further and allows for slowly changing seasonal components. This could have also been achieved by incorporating seasonal ARIMA components into a time series intervention model. Balkin/Ord's figure 1 shows that the seasonal fluctuations for Arizona are reasonably stable over time, indicating that a model with constant seasonal indicators would have been equally appropriate.

The Balkin/Ord paper analyzes *aggregate* data on the number of fatal interstate accidents. It does not use information to address if speed-related circumstances contributed to the accident, nor does it use such factors such as gender and age of the driver, road and weather conditions at the time of the accident, type of vehicle involved, and evidence of alcohol involvement. Aggregate analyses are always subject to the criticism that important factors have been overlooked. A follow-up analysis of individual fatal accidents would be worthwhile. Microdata on each fatal accident, with detailed information on various contributing factors, including whether speed was a contributing factor, is available for each accident. Admittedly, such analysis for 50 states would be an almost impossible task, and this comment is not a criticism of the Balkin/Ord study. Instead, it is offered as a recommendation that additional studies confirm the findings at the aggregate level with detailed analyses at the micro-level, at least for a few selected states.

[2] Rural and urban interstates are by far the safest road systems; only 5 to 10% of all fatal accidents occur on interstates.

REFERENCES

Abraham, B. and J. Ledolter. 1983. *Statistical Methods for Forecasting*. New York: Wiley.

Box, G.E.P., G.M. Jenkins, and G.C. Reinsel. 1994. *Time Series Analysis Forecasting and Control*. Upper Saddle River, NJ: Prentice-Hall.

Box, G.E.P. and G.C. Tiao. 1975. Intervention Analysis with Applications to Economic and Environmental Problems. *Journal of the American Statistical Association* 70:70–9.

Ledolter, J. and K. Chan. 1994. *Safety Impact of the Increased 65 mph Speed Limit on Iowa Rural Interstates.* Final Report. Midwest Transportation Center, University of Iowa.

_____. 1996. Evaluating the Impact of the 65 mph Maximum Speed Limit on Iowa Rural Interstates. *The American Statistician* 50:79–95.

U.S. Department of Transportation (USDOT), National Highway Traffic Safety Administration (NHTSA). 1998. *Report to Congress: The Effect of Increased Speed Limits in the Post-NMSL Era.* Washington, DC.

Discussion

MICHAEL D. FONTAINE
TONGBIN TERESA QU
Texas Transportation Institute

KARL ZIMMERMAN
CLIFFORD H. SPIEGELMAN
Texas A&M University

The Balkin/Ord paper has addressed the timely and controversial topic of whether speed-limit increases raise the frequency of fatal crashes. Earlier studies did not have sufficient data to determine the true relationship between speed-limit increases and fatal crashes. Enough time has now elapsed since the repeal of the national maximum speed limit (NMSL) in late 1995 to determine if there was a significant increase in fatal crashes. The structural model chosen by the authors certainly appears to be an improvement over models used in earlier studies.

One of the most important parts of scientific and engineering studies is the formulation of hypotheses. Studies of the safety effects of increased highway speed limits could focus on the total number of injuries, fatalities, injury crashes, fatal crashes, or many other measures. This study focuses on the number of fatal crashes. The authors fit the crash data to a structural model and do not attempt to explain the parameters estimated. Here we will try to extend the discussion begun by the Balkin/Ord paper in many important respects, such as giving alternative factors that might be modeled. We also provide a general discussion of study validity.

Clifford H. Spiegelman, Department of Statistics, Room A447 Blocker Bldg., Texas A&M University, College Station, TX 77843-3143. Email: cliff@stat.tamu.edu.

The study authors seem be looking for an effect independent of confounding factors and covariates. Confounding factors are common in field studies even when the study focuses on effects known to be practically important. For example, smoking is known to affect health, but there are several other factors such as diet, family health history, and exercise that should be taken into account when studying the effect of smoking. See, for example, Yano et al. (1977) as an example of a smoking-effect study that uses covariates. Including potentially important factors in an explicit manner was beyond the intended scope of the Balkin/Ord study.

It is unclear how the current study implicitly handles vehicle-miles of travel (VMT), increased availability of and different forms of airbags, or many of the other possible covariates we list in our discussion. The study authors appear to assume that changes in fatal crashes occur independently of all factors except the speed-limit increase. To some extent, they have shown that an increase in fatal crashes happening independently of other important factors may be a small, sporadic effect. This study, much like many studies of health effects, may not show as strong or convincing an effect because other important factors were not explicitly modeled.

STUDY VALIDITY

The purpose of the Balkin/Ord study was to determine whether an increase in the speed limit resulted in a higher incidence of fatal crashes. The authors hypothesized that the change in speed limit results in a permanent shift in the number of fatal accidents. The general approach used was first to explore trends in the data, then identify the discontinuity of the trend at the time of speed- limit increases, and finally draw conclusions on whether the speed-limit change resulted in higher numbers of fatal crashes. Two questions need to be answered for the purpose of this study. First, can a trend interruption be observed at the times of speed-limit increases? Second, was the interruption caused by the increases in speed limits? The authors attempted to answer the first question by applying their structural model. No attempt was made to determine if the increase in speed limit was solely responsible for any change in fatal crash frequency. The authors simply indicate whether the speed-limit increase resulted in a discontinuity. Their conclusions were made without considering possible effects from external factors such as weather, traffic, road user demographics, composition of the vehicle fleet and so on.

This study can be best described as a time-series quasi-experiment study. Compared to a true experiment, a quasi-experiment lacks full control over the events and subjects being studied, in terms of random assignment of treatments to subjects (Cook and Campbell 1979). In this study, the researchers could not control the population group, location, timing, or manner in which the speed limit was increased. The monthly data from January 1975 to December 1998 on fatal crashes both before and after the speed-limit increase forms the time-series quasi-experiment problem.

The concepts of internal and external validity are essential to obtaining meaningful results from any experimental design. "Internal validity is the basic minimum without which any experiment is uninterpretable," while "external validity asks the question of generalizability" (Campbell and Stanley 1966). In other words, any experimental design must be internally valid to yield reliable results and be externally valid to provide useful predictions about the effect of that treatment (in our case a speed-limit increase) to other populations and times.

The time-series design is generally internally valid with only one major limitation, namely, rival events (called "potential interventions" by the authors). Rival events occurring at the same times as speed-limit increases are the most serious threat to the internal validity of the authors' findings. These rival events could be responsible for any observed change in crash frequency or for masking changes in the fatal crash data. Rival events also provide potential alternative explanations for these findings. This problem can be overcome when the likelihood of rival events can be discounted. The authors did acknowledge the existence of potential interventions besides the speed-limit increase. The authors, however, did not give any explanation why potential interventions were not the causes for the observed effect or did not mask an effect that might have occurred.

In fact, rival events could plausibly cause the shift in fatal accident counts. For example, in the case of rural Arizona where increases in fatal crashes were found to be significant at the first speed-limit change, several issues can be raised: 1) Was there also an unusually high increase in VMT, which might be the cause of the increase in crashes? 2) Were there any severe weather events possibly causing bad road conditions, in turn causing more crashes? Winter weather could have played a role in short-term increases in the number of fatal crashes following the speed-limit increase in 1995 since raising speed limits was permitted as early as December. To ensure internal validity, these rival factors and combinations thereof must be ruled out as the causes for the significant increases in fatal crashes. The list of rival factors could be different for different states. However, such lists could possibly be extensive for all states. Considerable information about local conditions for each state is required to isolate the effect of the speed-limit increase from other rival factors.

The strength of the time-series approach is that fatal crash data before and after the speed-limit increases provide the possibility of exploring the existing trends and patterns in the data so that a discontinuity of the existing trend can be detected. As the authors pointed out, a simple before/after study would not be appropriate in this case. Simply comparing the count of fatal accidents immediately before to the count of accidents immediately after and then attributing the difference to the speed-limit increase would be misleading.

The external validity of the time-series design has serious problems. It is clear that the effect of the speed-limit increase is specific to the individual states. As shown in this study, some states have significant increases in fatal crashes, while some states have insignificant increases, and some essentially do not change. The legitimacy of the authors' conclusions generalized across states in this study is therefore uncertain.

The time-series approach is appropriate for this study. However, more study is needed to isolate the effect of speed-limit increase from the effects of other potential interventions. The authors' analysis does not sufficiently isolate the impact of the speed-limit increase from rival explanations.

DATA LIMITATIONS AND EXAMPLES OF A FEW CONFOUNDING FACTORS

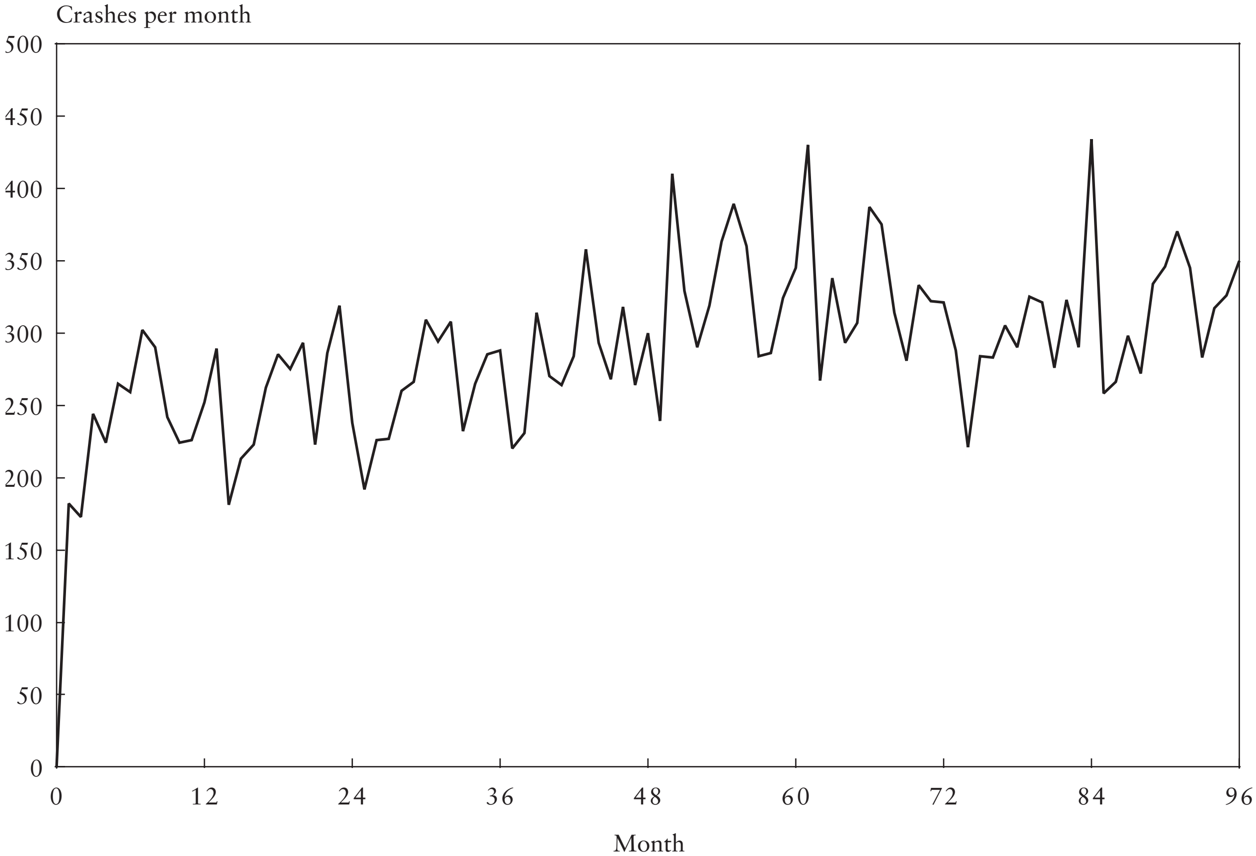
Fatal crashes represent a serious safety concern and are an important measure to examine. However, fatal crashes are relatively rare events and their counts may show quite a bit of instability from year to year and from month to month. For example, figure 1 in the paper shows that between 0 and 20 fatal crashes occurred each month on Arizona rural interstates. There is quite a bit of fluctuation from month to month in this figure, and it may be difficult to determine a significant trend based on fatal crashes alone. An examination of injury crashes would probably provide a more stable data set for this analysis. This paper provides a valuable indication of the possible impact of speed-limit changes on fatal crashes. However, there are some areas where this paper is unclear about the data used to perform this analysis. There are also potential opportunities to take this analysis in new directions that could provide, in our opinion, a more accurate examination of the impact of the speed-limit changes.

First, it is not clear from the paper whether the authors assumed that all states changed their speed limits in April 1987 and then again in December 1995. An assumption of uniform intervention dates across all states creates issues with their analysis. While many states did change their speed limits as soon as they were legally able, many other states did not enact a higher speed limit until much later. For example, Virginia did not choose to raise its speed limit to 65 miles per hour (mph) on rural interstates until July 1988 (Jernigan et al. 1994). Louisiana did not increase its interstate speed limit to 70 mph until August 1997 (USDOT NHTSA 1998). If these time periods were not correctly categorized in the analysis, the results could be inaccurate. It is not clear whether the authors changed the intervention dates on a state-by-state basis or used a uniform intervention date for all states.

In 1987, states were permitted to increase speed limits on rural interstates to 65 mph. This created a uniform speed-limit change in those states that chose to increase speed limits. In 1995, the NMSL was repealed, allowing states to set their own speed limits. Unlike the 1987 speed-limit change, this resulted in some variation in the interstate speed limits established across the nation. Some states raised speed limits to 70 mph, while others raised the limit to 75 mph. Some states, such as Texas, enacted differential speed limits for cars and trucks. Studies have shown that both the absolute speed of vehicles and large differences in speeds among vehicles in the traffic stream can be significant causal factors of crashes (Cirillo 1968; Beatty 1973). Given these studies, it appears that the impact of the both magnitude and type (differential or uniform) of speed-limit change should be considered when assessing the impact of the repeal of the NMSL. This could explain some of the differences in results observed by the authors.

In addition to differences in the magnitude of the speed-limit change, drivers also reacted differently to the speed-limit increases from state to state. Some states were experiencing a very high degree of motorist noncompliance with speed limits prior to the repeal of the NMSL. In these cases, actual speeds may not have changed very much following the repeal of the NMSL. Studies have shown that drivers in different states reacted very differently to the same speed-limit increases. For example, studies performed in Michigan and California showed relatively small increases in mean speed of only 1 and 2 mph, respectively, after the speed limit was increased 5 mph to 70 mph (Retting and Green

FIGURE 1 Injury Crashes per Month on Rural Interstate Highways in Texas, 1992–1999



1997; Nolf et al. 1998). In Texas, mean speeds were observed to increase by 5 mph when the speed limit was increased 5 mph to 70 mph (IIHS 1996). If all changes in crash frequency could be attributed to increased travel speed, it would be expected that states that experienced smaller increases in travel speed would exhibit smaller increases in crash frequency. The relationship between the magnitude of actual observed travel speeds and crash frequency bears further investigation.

Analysis of urban crashes also presents a number of additional concerns. Speed-limit changes were not always uniform in urban areas. It is a reasonable assumption that rural interstates would be posted at the maximum speed allowable by law. However, in urban areas road geometry, safety considerations, congestion, and high volumes of traffic may preclude posting the speed limit at the legal maximum. In many urban areas, a 55-mph speed limit was retained on some roads even though a higher speed limit was legally possible. Given that speed limits were not always increased on urban interstates, it may be difficult to determine if an increase in the speed limit was responsible for any observed increase in crash frequency in urban areas. In fact, roads where the speed limit was increased may actually represent a minority of the roads, which may dilute potential impacts of the increased speed limit in this paper's analysis. Factors such as increasing congestion and greater prevalence of work zones should be examined as possible alternative explanations for any observed crash increases in urban areas.

ADDITIONAL COVARIATES

Crashes occur for a wide variety of reasons, including driver error, vehicle breakdown or failure, poor roadway conditions, poor operating conditions, and all combinations of these factors.

For the sake of example, consider rollover crashes, one of the most severe types of crashes. The frequency of fatal rollover crashes is affected by any or all of the following:

- driver characteristics (age, gender, personality, experience, alcohol and drug use, fatigue)
- vehicle characteristics (type, loading, maintenance level)
- roadway characteristics (functional classification/design standard, curvature, delineation, illumination)
- environmental characteristics (time of day, visibility, precipitation, traffic volume)

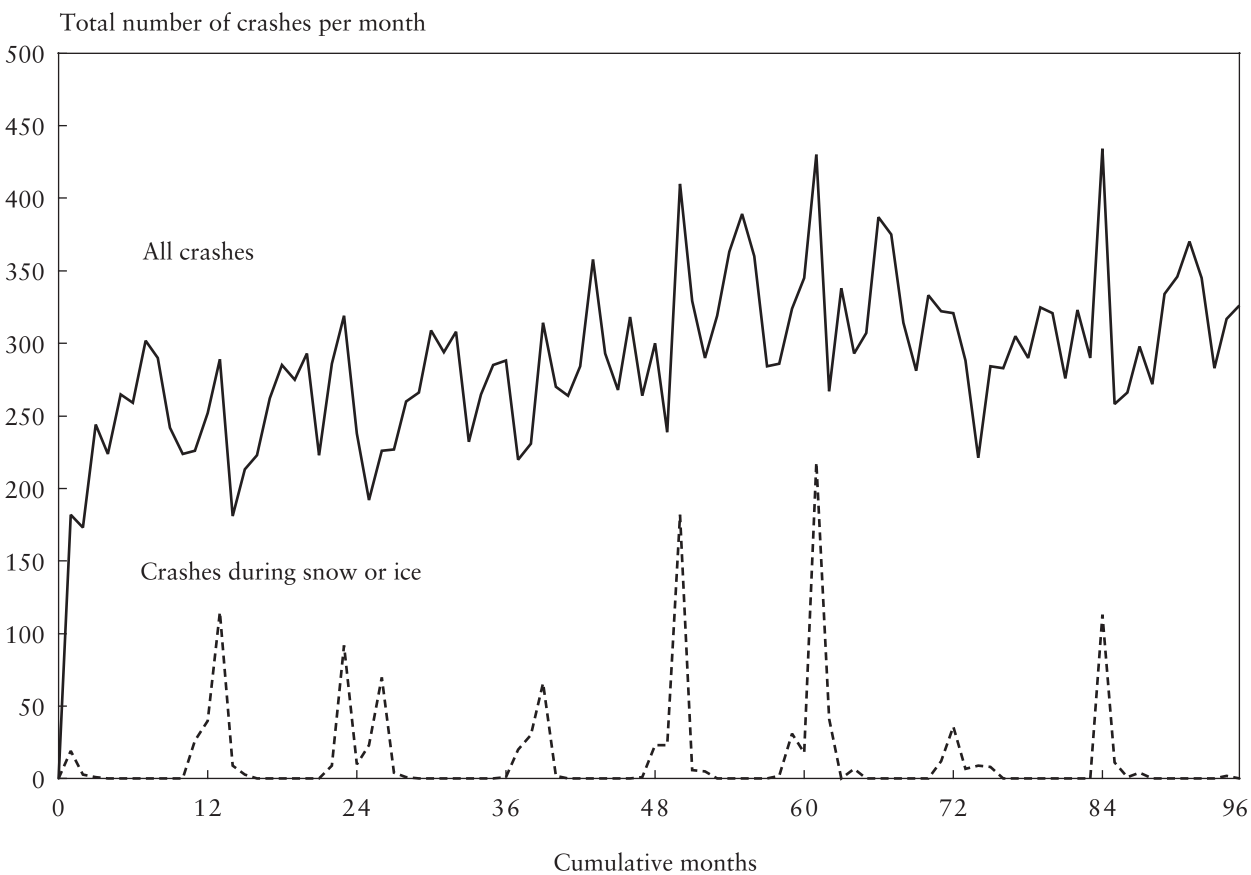
The above list is not all-inclusive. Some of these factors may not be very important and some difficult to quantify, but all of them can play a role in fatal crashes.

A few possible covariates are of particular concern because of their obvious effect on several of the factors shown above. These may include:

- effects of winter precipitation
- percentage of trucks on a particular roadway
- differences in the amount of speed-limit increase from state to state
- "spillover effects" in states that did not increase speed limits
- uneven changes within a state
- the general population's learning curve in adjusting to higher speed limits

The effects of winter precipitation, particularly sleet or snow remaining on the roadway for an extended period, may be very important. In Texas, there were three months with unusually high numbers for injuries and fatal crashes after the last speed-limit increase. These were February 1996, January 1997, and December 1998 (see figure 1). Each of these occurrences corresponded to a major winter storm, which tends to be a rare event in Texas (see figure 2). Two of these events early in the "after" period could skew the results and may have in at least one analysis (Griffin et al. 1998). How-

FIGURE 2 Injury Crashes per Month and Injury Crashes during Icy or Snowy Weather on Rural Interstate Highways in Texas, 1992–1999



ever, this effect will vary by location within the United States. A winter storm considered severe in Texas may be fairly normal winter weather in Minnesota, for instance.

Another environmental consideration is the percentage of trucks in the traffic stream. Not all interstate routes are equal in terms of truck traffic, with trucks composing over 50% of traffic on some routes in Texas and Nebraska. This has an effect on other traffic and means that trucks are more likely to be involved in a collision. Because trucks have a much larger mass than passenger cars, the risk to the occupants of the passenger vehicle is very high. Also, trucks are tall with high centers of gravity and are less stable under virtually all conditions than a passenger vehicle is. In Texas, the number of crashes involving a truck generally followed the total number of crashes, but this may not necessarily be true elsewhere in the country.

Another effect might be the spillover speed effect caused when one state changes its speed limits and another one retains the old ones. An example of this is Nebraska, with interstate speed limits increased from 65/55 to 75/65, and Iowa, with the older 65/55 speed limits retained. It is possible that crashes in Iowa, for instance, may be affected more by the changes in neighboring states because of greater variations in speed. The spillover effect could be a source of some significant results in states that did not increase their speed limits in 1996, such as Maryland, Pennsylvania, and Tennessee. The idea of a learning curve deserves mention. In one section of the paper, the authors claim that the effects of such a change in driver behavior would be relatively short-term in nature and could safely be ignored. Later, the authors use it to explain the decrease in fatal crashes in Arizona after the initial increase in 1996. The same type of reduction occurs in Texas (rural), Oklahoma (urban), and Nevada (urban). The reduction might indeed be due to drivers becoming more accustomed to higher speeds. It could also be due to different weather patterns, changes in law enforcement (speed and otherwise), road construction, or any one or more other possible factors. Also, the effect is a year or two after the speed limit change, not a few months as the authors earlier claimed. If the learning curve is indeed short-term, then the later effects must be something else entirely.

SUMMARY

We thank Balkin and Ord for providing a means of discussing the impact of the 1995–1996 speed-limit increase. Their study represents an improvement in the series of studies of the effect of speed-limit increases. The study, however, is far from the final word on the impact of the speed-limit increase. In our opinion, their study did not pay enough attention to its conclusions' validity. We feel that a more comprehensive study that takes into account additional explanatory factors needs to be done if the public is to know the true effects of the speed-limit increase.

REFERENCES

Beatty, R.L. 1973. Speed Analysis of Accidents on Interstate Highways. *Public Roads* 37, no. 3:89–107.

Campbell, D.T. and J.C. Stanley. 1966. *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally.

Cirillo, J.A. 1968. Interstate System Accident Research–Study II–Interim Report II. *Public Roads* 35, no. 3.

Cook, T.D. and D.T. Campbell. 1979. *Quasi-Experimentation, Design and Analysis Issues for Field Settings.* Boston: Houghton Mifflin Company.

Griffin, L.I., O. Pendelton, and D.E. Morris. 1998. *An Evaluation of the Safety Consequences of Raising the Speed Limit on Texas Highways to 70 Miles Per Hour.* Texas Transportation Institute, Texas A&M University.

Insurance Institute for Highway Safety (IIHS). 1996. Urban Freeway Speeds Jump in First States to Raise Speed Limits. *Status Report,* 31, no. 5.

Jernigan, J.D., S.E. Strong, and C.W. Lynn. 1994. *Impact of the 65 Mile Per Hour Speed Limit on Virginia's Rural Interstate Highways: 1989–1992 Final Report.* Virginia Transportation Research Council, Charlottesville, VA.

Nolf, S.A., T.L. Maleck, W.C. Taylor, and T.S. Czewski. 1998. An Evaluation of the Michigan 70 MPH Speed Limit, in 77th Annual Meeting Preprint CD-ROM, Transportation Research Board. National Research Council, Washington, DC.

Retting, R.A. and M.A. Greene. 1977. Traffic Speeds Following the Repeal of the National Maximum Speed Limit. *ITE Journal* 67, no. 5.

U.S. Department of Transportation (USDOT), National Highway Traffic Safety Administration (NHTSA). 1998. *The Effect of Increased Speed Limits in the Post-NMSL Era.* Federal Highway Administration, Washington, DC.

Yano, K., G.G. Rhoads, and A. Kagan. 1977. Coffee, Alcohol, and the Risk of Coronary Heart Disease among Japanese Men Living in Hawaii. *The New England Journal of Medicine,* 297.

ANDREW HARVEY
Cambridge University

The paper by Balkin and Ord uses stochastic rather than deterministic trends to model the series on crashes. This is important since deterministic trends are rarely appropriate for economic and social time series, and their use can result in misleading inferences on the effects of interventions. One of the attractions of the structural time series modeling approach is that a deterministic trend emerges as a special case of a stochastic trend; this happens in equation (2) in Balkin and Ord when σ_n^2 is zero. The hypothesis that σ_n^2 is zero can be tested formally using the procedure of Kwiatkowski, Phillips, Schmidt, and Shin (1992). The amendments needed to allow for the effects of intervention variables are discussed in Busetti and Harvey (2001). This test has not yet been implemented in the STAMP package of Koopman et al. (2000), which Balkin and Ord uses to carry out calculations. However, evidence for the suitability of a random walk is provided by the Box-Ljung Q-statistic obtained when σ_n^2 is set to zero; for rural Arizona this results in Q(15,13) jumping from a statistically insignificant 12.85 to a highly significant 40.66. If the random walk is replaced by a first-order autoregressive process, the coefficient is estimated to be 0.98. Thus, for Arizona at least, the random walk level seems to be a reasonable model.

There are two ways in which the analysis could be improved. The first is by taking account of the fact that the data are in the form of counts, some of which are quite small. Rather than using the log(y+1) transformation, a count data structural time series model could be used. Harvey and Fernandes (1989) gives a procedure that can be used when only the level is stochastic, while Durbin and Koopman (2000) shows how simulation methods enable a general count data model to be estimated.

Andrew Harvey is a professor of Econometrics at the University of Cambridge on the Faculty of Economics and Politics. Address: Sidgwick Avenue, Cambridge, CB3 9DD England. Email: Andrew.Harvey@econ.cam.ac.uk.

The second suggestion is to make use of control groups. For example, the urban Arizona series can serve as a control for rural Arizona. If the series are correlated, one can go some way toward resolving the issue raised by Balkin and Ord when they say "…Arizona... had an increase in the number of crashes the year of the speed-limit change but a decrease from that level in subsequent years. This suggests that drivers in Arizona may have learned how to drive safely at the new limit. Such patterns are not consistent across states, and this issue requires further investigation." The structural time series framework for using control groups is discussed in some detail in Harvey (1996). In the present context, it simply involves setting up a bivariate time series model consisting of equations (1) and (4) of the Balkin/Ord paper and estimating them jointly with allowance made for correlations across the level, seasonal, and irregular disturbances. Thus

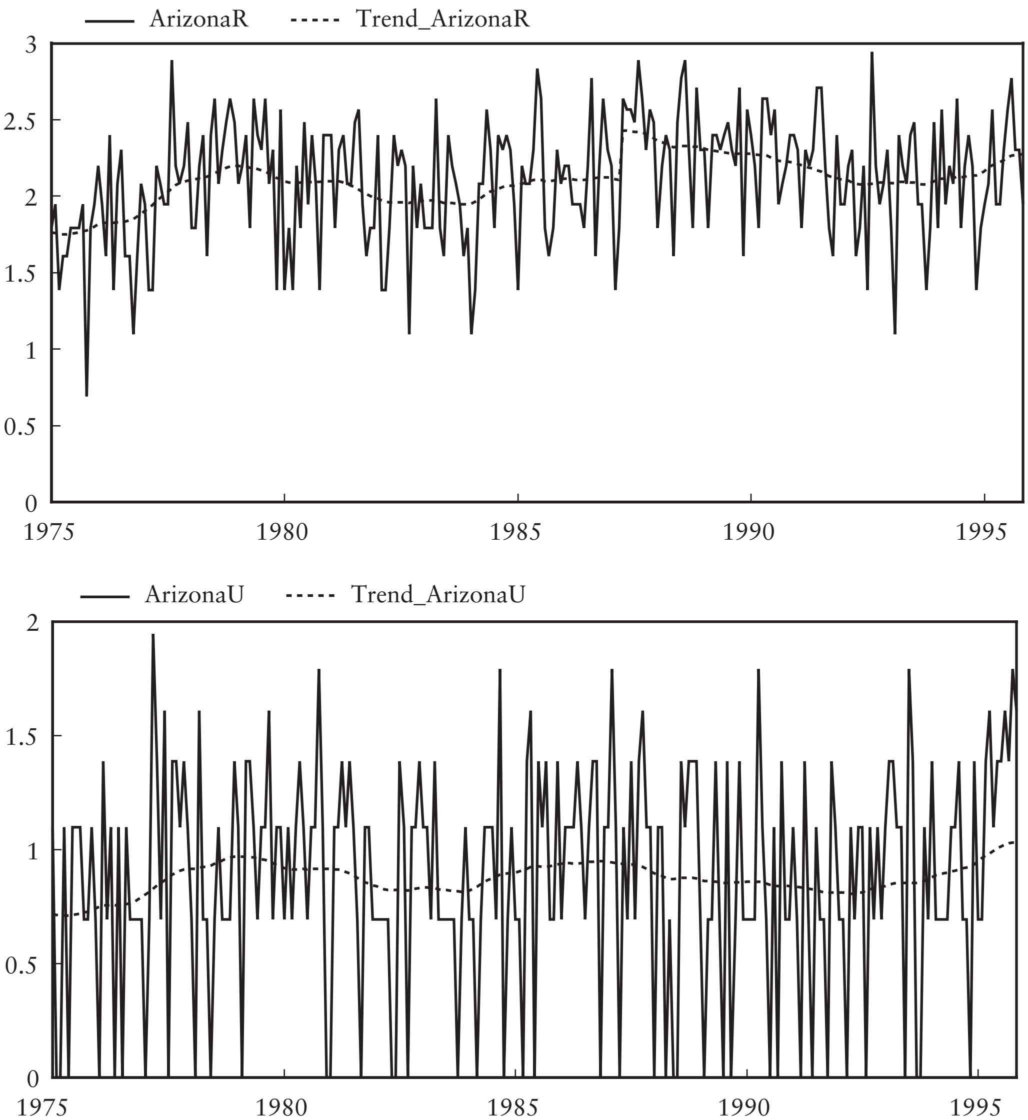
$$y_{1t} = \mu_{1t} + \varepsilon_{1t}, \qquad t = 1, \ldots, T, \tag{D1}$$
$$y_{2t} = \mu_{2t} + \lambda z_t + \varepsilon_{2t},$$

where the intervention variable, z_t, is defined as in (5), and

$$\mu_{it} = \mu_{i,t-1} + \eta_{it}, \qquad i = 1, 2 \tag{D2}$$

Such a model can be estimated in STAMP. Using data up to November 1995 to exclude the later change, the correlation between the level disturbances, η_{1t} and η_{2t}, is 0.81, while the correlation between the irregulars, ε_{1t} and ε_{2t}, is –0.07. This translates into a reduction in the root mean squared error (RMSE) of the level intervention in the rural series located at April 1987. The t-statistic correspondingly increases from 2.21 to 2.41. The gain is not dramatic, possibly because the number of crashes in the urban series is so small. However, figure 1 here clearly shows the connection between the series with the urban series and also shows the slight decrease noted by Balkin and Ord after 1987.

FIGURE 1 Series for Rural and Urban Arizona with Levels Estimated by a Bivariate Model



The ideal model for control group analysis would be a bivariate, count data model as in Fernandes, Ord, and Harvey (1993). A simpler option would be to aggregate the data to a quarterly level, thereby removing zeroes in nearly all the series and yielding a better Gaussian approximation in logarithms.

Balkin and Ord suggests the use of a "Super t-Test" to determine the significance of interventions for all states together. There may be a problem here insofar as the individual t-statistics are not independent of each other. An alternative approach, which also solves the small counts problem, is to aggregate all the crashes in states where there was a change in speed limit and then test the significance of the intervention variable. Taking the logarithm of the total number of crashes in the states where the speed limit was raised in April, May, or June of 1987 gives a t-statistic of 3.21 for a level intervention in May of 1987. Again, only observations up to November 1995 were used. The t-statistic increases when a control group series is formed from the urban series and the rural series where the limit was not raised in 1987. The bivariate model shows a correlation of 0.90 between the level disturbances and the intervention, λ, is estimated as 0.167 with a t-statistic of 4.35. The increase is clearly significant and translates into an 18% increase in crashes on roads where the speed limit was raised.

REFERENCES

Busetti, F. and A.C. Harvey. 2001. Testing for the Presence of a Random Walk in Series with Structural Breaks. *Journal of Time Series Analysis* 22:127–50.

Durbin, J. and S.J. Koopman. 2000. Time Series Analysis of Non-Gaussian Observations Based on State-Space Models from Both Classical and Bayesian Perspectives (with discussion). *Journal of Royal Statistical Society, Series* B 62:3–56.

Fernandes, C., K. Ord, and A.C. Harvey. 1993. Time Series Models for Multivariate Series of Count Data. *Developments in Time Series Analysis*. Boca Raton, FL: Chapman and Hall.

Harvey, A.C. 1996. Intervention Analysis with Control Groups. *International Statistical Review* 64:313–28.

Harvey, A.C. and C. Fernandes. 1989. Time Series Models for Count Data or Qualitative Observations. *Journal of Business and Economic Statistics* 7:409–22.

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

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, and Y. Shin. 1992. Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We that Economic Time Series Have a Unit Root? *Journal of Econometrics* 44:159–78.

Authors' Rejoinder

SANDY BALKIN

KEITH ORD

First, we would like to thank Johannes Ledolter, Andrew Harvey, and the team of Michael Fontaine, Tongbin Qu, Clifford Speigelman, and Karl Zimmerman (FQSZ) for their constructive and valuable comments. Without doubt, the issue at hand is a complex one, and the discussants have both carefully evaluated our study and suggested a number of directions for further research.

CHOICE OF DATABASE

With the benefit of hindsight, we recognize that the discussants raised some commentary in response to some decisions and implicit assumptions not clearly stated in the paper. We begin with our choice of data. We agree that fatal crashes (fortunately) represent a small proportion of the accidents on our highways, but we chose to restrict attention to this set of statistics because the reporting of such events is more uniform. State-by-state requirements vary considerably for lesser incidents, and we hoped to use data that were reasonably comparable across states. However, we accept that "injury crashes" might have produced more stable results, at least within a particular state. A similar discussion is appropriate for the choice of interstates or primary roads. We felt that the classifications of urban and rural interstates were more uniform across the country than the definitions of primary roads. Nevertheless, a study for primary roads would also provide valuable insights, particularly on the question of differential responses by state.

Vehicle-miles traveled (VMT) was indeed a variable that we would have liked to use to produce an accident *rate* rather than a pure count. However, an assessment of the availability and quality of monthly VMT figures at the state level led us to the conclusion that such an analysis was not feasible for all states. Thus, we reluctantly decided to work with pure counts. An analysis of those states for which good data are available, building on the earlier work of Ledolter and Chan (1996), would certainly be worthwhile.

There were several comments concerning our definition of the intervention variables. First, we would like to make clear that we set the indicator as a step function; that is, it was scored as zero in the months prior to that state changing the speed limit and as one for the month when the change took place *in that state* and for all succeeding months. We recognize that there may well have been different levels of preparation and compliance to the new limits in the states, but we did not have such information available. The references provided by FQSZ are helpful in exploring this question further. The precise coding of the interventions is difficult. Even when all increases are the same numerically, as in 1987, the proportions of interstate mileage affected vary by state, as do questions of enforcement. For example, in Pennsylvania during the era of the 55-miles per hour (mph) speed limit conventional wisdom held that one would not normally be ticketed for speeding when traveling at 64 mph or less but that during the era of the 65-mph limit, the "threshold" was raised only to 69 mph. Whether or not such folklore is true, it clearly has an impact on driving behavior.

Ledolter mentions their earlier study, which took into account changes in average traffic speed. Again, we were not able to find reliable monthly data on this variable for all states and so did not include it in the analysis. However, the Ledolter and Chan (1994, 1996) results suggest a gradual shift over time, whereas we hypothesized a sudden impact on accident rates. Thus, even if the data were available, we would expect the effects to be distinct. Of course, the statistical analysis would be more efficient if the average speed were taken into account.

When speed limit increases vary, should the intervention be scaled across states to match the amount of the increase? We note that such a scaling does not affect the statistical analysis for an individual state, provided separate indicators are used for each increase. As Ledolter notes, a case could certainly be made for using a single scaled indicator to cover the two increases, but the benefits of consolidating the "speed effect" need to be set against some of the "rival events" mentioned by FQSZ.

METHODOLOGY

We are grateful to Ledolter for summarizing the linkages between ARIMA and structural models. Hopefully, this description will make the paper and ensuing discussion accessible to a wider audience. We agree that similar results would be expected, whichever paradigm is adopted. Likewise, we concur that the seasonal patterns were usually quite stable; indeed, for a number of states the analysis did indicate fixed seasonals, but we did not report those details. Our reason for using structural models rather than the more widely used ARIMA framework was that we feel the direct specification of level, slope, and seasonal components is more intuitively appealing and allows the investigator to incorporate prior knowledge into the model selection process more readily. Granted, the ARIMA models can be decomposed into components, but this analysis is not provided in most software packages.

Harvey makes a number of valuable comments. The testing procedures developed by him and his co-authors over the years have brought the structural modeling approach to the point where it provides a completely viable alternative to ARIMA modeling. Indeed, as noted above, we feel that structural modeling is superior because of the intuitive understanding provided. As a theoretical aside, we note that the alternate approach to structural modeling developed by Ord, Koehler, and Snyder (1997) provides a system with the same parameter space as the ARIMA class, whereas the original system has a more restrictive parameter space. Thus, the methodological objections to using the structural modeling approach are gradually disappearing. We look forward to using these new developments in the next version of STAMP.

Harvey's comments about the use of a proper count-based model are well taken. We admit to using more accessible software in preference to the more correct but less computationally convenient count models. When the counts are small, this may lead to erroneous conclusions for a few of the smaller states, a point noted in our paper. Also, the idea of using control groups is an excellent one and would help to neutralize many of the "rival events" cited by FQSZ.

On the Super *t*-Test, we agree that the assumption of independence was not explored and that other approaches should also be considered. However, we feel that the general conclusions would remain valid.

FQSZ's comment that the study is "quasi-experimental" is perhaps too kind, and we would place it more on the "observational" end of the spectrum. The extent to which general conclusions can be drawn really rest on the precise definition of times in each state at which the interventions took place followed by a check for measurable effects at those times. In this sense, the study is quasi-experimental since it is highly unlikely that any of the rival events would match up with more than a few of the specified interventions.

ANALYSIS OF THE DATA

The commentaries suggest a variety of additional factors to be taken into account, and we are reminded of the old story about the statistician and the economist who jointly examined the results of a regression analysis. The statistician asked, "Why did you use so many variables?" to which the economist replied, "Why did you use so few?"

Our objective was to account for the broad trends and seasonal patterns in the data and, that done, to identify the effects of the changes in speed limits. Without doubt, incorporating some of these factors would serve to improve models for individual states. Further, key variables such as VMT would have been valuable had they been uniformly available. On balance, we believe that the "keep it simple" approach was appropriate for an initial study and that one of the objectives was, indeed, to stimulate thinking about more sophisticated analyses in the future.

The point about learning curves, raised by FQSZ, is an interesting one. We began with the simple intervention variable described earlier and did not hypothesize learning effects. These appear to exist in some states but by no means in all. We debated whether to modify the intervention variables to describe this more complex behavior but eventually decided to stay with our original formulation in order to avoid any charges of "data dredging." Again, this is an important question that deserves a properly designed study.

FQSZ observes "It is clear that the effect of the speed-limit increase is specific to the individual states" and later that "The generalization ability of the authors' conclusion across states in this study is therefore uncertain." Our tabulated results show that the effects were not uniform across states; the penultimate sentence of our conclusions states that "Overall, increases were seen in *some* [italics added] states following speed limit changes." Our conclusions might have been more forcibly stated, but the results surely point to some increase for rural interstates, even though the effects vary by state. A key question for further research is why states had different levels of response and how this information might be used to improve safety.

DIRECTIONS FOR FURTHER RESEARCH

We have already mentioned several possible directions for further research in conjunction with the comments made above. Of these, the most important ones would seem to be the study of differential responses by the states, the examination of learning effects, and the use of control groups. In addition, more focused studies that use more complete data from those states for which they are available would provide additional insights. Conversely, those states that do not provide key indicators such as VMT should give careful consideration to expanding their data collection efforts.

The other major point, made explicitly in some places and implicitly in others, is that the current analysis uses only aggregate data. The Fatality Analysis Reporting System (FARS) database provides detailed information on each fatal accident and could be used for micro-level studies to explore the impact of covariates such as those listed by FQSZ.

In conclusion, we would like to thank the commentators once again for their thoughtful and constructive suggestions, and we hope that our collective contributions will serve to advance understanding in this important area.

REFERENCES

Ledolter, J. and K. Chan. 1994. *Safety Impact of the Increased 65-mph Speed Limit on Iowa Rural Interstates.* Final Report. Midwest Transportation Center, University of Iowa.

Ord, J.K., A.B. Koehler, and R.D. Snyder. 1997. Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models. *Journal of the American Statistical Association* 92:1621–29.

Accounting for Uncertainty in Estimates of Total Traffic Volume: An Empirical Bayes Approach

GARY DAVIS
University of Minnesota

SHIMIN YANG
The St. Paul Companies

ABSTRACT

Predicted or estimated totals of traffic volume over one or more years are required in both highway pavement and safety engineering. While current recommended practices contain guidance on how to generate such estimates, they are less clear on how to quantify the uncertainty attached to the estimates. This paper describes an initial solution to this problem. Empirical Bayes methods are used to compute quantiles of the predictive probability distribution of the traffic total at a highway site, given a sample of daily traffic volumes from that site. Probable ranges and their associated probability values are readily found, and a point prediction of the total traffic can be obtained as the median, or 50th percentile, of the predictive distribution. The method of derivation can also be used to find the predictive density and the moments of the predictive distribution if needed. No data other than those routinely collected by statewide traffic monitoring programs are needed. A test comparing computed 90% credible intervals for annual traffic volume with the corresponding actual volume at 48 automatic traffic recorder sites showed that the actual coverage percentage was not significantly different from the nominal 90% value.

Gary A. Davis, Department of Civil Engineering, University of Minnesota, 122 CivE, 500 Pillsbury Drive SE, Minneapolis, MN 55455. Email: drtrips@tc.umn.edu.

INTRODUCTION

In engineering design, it is sometimes necessary to work with variables whose values are not known with certainty. In such cases, a rational compromise between over- and under-design requires first the determination of a probable range for a variable's outcome and then a design to accommodate values in this range. Implementation hinges on an acceptably accurate assessment of this probable range because if the assessed range is too narrow, the likelihood of a failure will be unacceptably high, but if the assessed range is too broad, resources will be expended in anticipation of improbable events. In addition, the users of scientific and engineering measurements often expect an assessment of the uncertainty attached to those measurements. The need for an assessment of uncertainty is revealed in the reporting of error bounds associated with opinion survey estimates, in the U.S. Supreme Court's recommendation that judges consider "known or potential rate of error" when determining the admissibility of expert scientific testimony (Foster and Huber 1999), and in the recommendation by the American Association of State Highway and Transportation Officials (AASHTO) that the "precision and bias" attached to traffic volume measurements be assessed and reported (AASHTO 1992).

The ability to justify an uncertainty assessment becomes especially important when either particular values of an estimate or the uncertainty range itself could be used by partisans to justify or oppose a controversial policy (Sarewitz and Pielke 2000). In a discussion of scientific predictions and social policy, Stewart (2000) found it useful to distinguish between two sources of uncertainty, which he called aleatory and epistemic uncertainties. Aleatory uncertainty arises when the outcomes of interest are governed by physically random processes that are at least in principle capable of generating stable long-run relative frequencies. Epistemic uncertainty, on the other hand, arises when our knowledge of underlying states of nature is incomplete. This duality in the notion of uncertainty appears to date back to the origin of modern ideas concerning probability (Hacking 1975) and is arguably the source of the current debate between Bayesian and frequentist views on the foundations of statistics (Howson and Urbach 1993).

Estimates or forecasts of the total traffic volume on a section of road for one or more years are used in both pavement design and traffic safety analysis and are also used to generate state- and nationwide estimates of total distance traveled. These are most often computed by multiplying an estimate of the road's mean daily traffic (MDT) volume by the number of days in the desired time horizon. For example, in pavement design (AASHTO 1993) estimates of MDT for each vehicle class are multiplied by 365 to obtain estimated yearly traffic totals, which are in turn used to predict the traffic loading over a pavement's design life. In traffic safety, the traffic exposure at a road site is computed by multiplying an estimate of MDT by the total number of days over which traffic accidents have been counted. For an intersection, these traffic totals are then summed over the intersection's approaches to give the total entering vehicles. For a highway section, the traffic total is multiplied by the section's length, producing an estimate of total vehicle kilometers of travel. Clearly, aleatory uncertainty is attached to an estimate of total traffic because, even if we knew a site's MDT exactly, the estimated traffic total and the actual total would likely differ, due to the unpredictable decisions of individual travelers. Epistemic uncertainty is present when the true MDT is not known exactly but has been estimated from a sample of daily traffic counts. AASHTO's (1993) recommended pavement design method explicitly allows for aleatory uncertainty as one of the components making up the overall variation term in the pavement design equation; however, epistemic uncertainty is not addressed. In traffic safety, current practices address the uncertainty in estimated accident rates due to the random nature of accident counts but do not appear to consider the contributions of either aleatory or epistemic uncertainty when estimating exposure (see Parker 1991).

Draper (1995) has illustrated how an accounting of multiple sources of uncertainty can be accomplished using Bayesian statistical methods, and in this paper we will consider the problem of assessing the uncertainty attached to estimates or forecasts of the total traffic volume. The second section will illustrate, using a simple example, how a Bayesian approach can be used to combine the contributions of aleatory and epistemic uncertainties into one assessment. The reasoning illustrated in that section will then be applied in the next section to develop an expression for the predictive distribution of a traffic total, which can be used to compute both point and interval estimates. The fourth section will then describe an initial empirical evaluation of this estimation method, and the final section will present conclusions. The development described in the third section draws heavily on past research into statistical models for time series of daily traffic counts (Davis and Guan 1996) and on weak convergence results for sums of lognormal random variables (Marlow 1967).

ILLUSTRATING THE SOURCES OF UNCERTAINTY IN TOTAL TRAFFIC ESTIMATION

Consider the problem of estimating the total traffic volume over a period of N days, using a short count collected with a portable traffic counter. Let

z_t = traffic volume on day t, t=1,...,N,

$z^N = \sum_t z_t$, the total traffic volume over days t=1,...,N,

z_l = traffic count on the lth sample day, l=1,…,n,

$\bar{z} = (1/n)\sum_l z_l$, the sample average,

$\mathbf{z} = \left[z_1, .., z_n\right]'$, n-dimensional vector containing the sample counts.

The ultimate objective is to estimate the total traffic volume z^N from the traffic count sample $\mathbf{z}$. For this example, we will assume that the daily traffic counts are independent and identically distributed normal, random variables with common mean μ and common variance σ^2. The mean value μ is the site's MDT, and we will assume that it is unknown. However, to avoid technical sidetracks, we will assume the variance of the daily volumes, σ^2, is known. We will also assume, for simplicity, that the sampled days are not part of the time period over which the total traffic is desired. Aleatory uncertainty will be present even when the MDT is known exactly. It is straightforward to show that given μ and σ, the probability distribution of the total traffic count z^N is normal with mean equal to $N\mu$ and variance equal to $N\sigma^2$. But, as noted earlier, we do not generally know μ but must estimate its value from the count sample, leading to epistemic uncertainty. If we assume that, prior to collecting our sample $\mathbf{z}$, our uncertainty concerning the MDT parameter is characterized by a prior probability density $f(\mu)$, Bayes Theorem can be used to find the probability density characterizing this uncertainty after collecting the sample $\mathbf{z}$. That is

$$f\left(\mu \middle| \mathbf{z}, \sigma^2\right) = \frac{f\left(\mathbf{z} \middle| \mu, \sigma^2\right) f(\mu)}{\int f\left(\mathbf{z} \middle| \mu, \sigma^2\right) f(\mu) d\mu} \tag{1}$$

In particular, if $f(\mu)$ is noninformative in the sense of being uniformly distributed on the real line, it can be shown that the posterior uncertainty concerning μ is characterized by a normal distribution with mean equal to the sample average and variance equal to σ^2 / n (Box and Tiao 1973). The joint effect of aleatory and epistemic uncertainty can then be determined by treating μ as a nuisance parameter and integrating it out of the joint density for z^N and μ

$$f\left(z^N \middle| \mathbf{z}, \sigma^2\right) = \int_{-\infty}^{\infty} f\left(z^N \middle| \mu, \sigma^2\right) f\left(\mu \middle| \mathbf{z}, \sigma^2\right) d\mu \tag{2}$$

(Draper 1995). Here $f\left(z^N \middle| \mathbf{z}, \sigma^2\right)$ denotes the predictive probability density of the total traffic given the count sample, while $f\left(z^N \middle| \mu, \sigma^2\right)$ denotes the predictive probability density of the total traffic when the MDT is known. For this example, closed form evaluation of (2) is possible (Box and Tiao 1973), leading to the conclusion that the predictive probability density $f\left(z^N \middle| \mathbf{z}, \sigma^2\right)$ is normal with mean equal to the $N\bar{z}$ and variance given by

$$\mathrm{var}\left[z^N \middle| \mathbf{z}, \sigma^2\right] = \frac{N(N+n)\sigma^2}{n} = N\sigma^2 + N^2\left(\tfrac{\sigma^2}{n}\right) \tag{3}$$

In this case, the contributions to the total variance attributable to aleatory and epistemic uncertainty can be separated, with $N\sigma^2$ as variance due to aleatory uncertainty and $N^2\left(\sigma^2 / n\right)$ as variance due to uncertainty concerning the MDT. Interestingly, while the variance due to aleatory uncertainty increases linearly with the number of days in the traffic total, the variance due to epistemic uncertainty increases quadratically. To see the relative contributions of these sources, suppose that we seek to predict one year's total traffic volume using a 10-day sample count and that the daily traffic volume has an MDT of 1,000 vehicles per day and a coefficient of variation equal to 0.1. The day-to-day variance would be $\sigma^2 = [(.1)(1{,}000)]^2 = 10{,}000$, so the standard deviation due to aleatory uncertainty would be $[(365)(10{,}000)]^{(\frac{1}{2})}$, which equals 1,910 vehicles. The standard deviation due to epistemic uncertainty would be $\left[(365)^2(10{,}000/10)\right]^{(\frac{1}{2})}$, equal to 11,540 vehicles. This example illustrates how epistemic uncertainty can be the dominant source of error and that neglecting its contribution can lead to a serious overstatement of a prediction's precision.

Because the main objective of this paper is to show how a more complete accounting of uncertainty can be added to current traffic monitoring practices, we describe these practices next. The chief purpose of a traffic monitoring program is to generate estimates of MDT on each of a jurisdiction's road segments. Ideally, this is done with year-round counting on each segment, but the cost of installing and maintaining such a comprehensive traffic monitoring system is prohibitive. Therefore, MDT estimates on the majority of road segments are obtained from samples gathered using portable traffic counters. Since traffic volumes vary systematically throughout the course of the year as well as across the days of the week, averages computed from short count samples are generally biased estimates of full year averages. However, if the magnitude of the bias is known, adjustments can be made. To determine these adjustments, most states employ a small number of permanent automatic traffic recorders (ATRs) placed on a representative sample of road segments. The daily traffic counts from the ATRs are used to cluster the ATRs into factor groups such that daily traffic volumes at sites in a factor group show similar seasonal and day-of-week variation patterns. The ATR counts are also used to estimate the seasonal and day-of-week factors characterizing each group. Each non-ATR road section is then assigned to one of these factor groups, and the variation factors characterizing the assigned group are used to adjust the short-count sample, providing a better estimate of the section's MDT. It is currently recommended that a suitably adjusted short count of 48 hours produces an estimate of MDT with acceptable precision (AASHTO 1992; USDOT FHWA 1995).

At least two sources of potential error can cause an estimated MDT to differ from a section's true (but unknown) MDT: sampling error, arising anytime the estimate is based on less than a complete census of the section's traffic volumes, and adjustment error, arising if the factors used to adjust the short-count sample differ from those which actually describe the sampled section's variation pattern. In a recent review of MDT estimation, Davis (1997b) pointed out that much of the earlier research used to justify the use of short counts for estimating MDT tended to underestimate the potential effect of adjustment error, and an analysis of the potential contributions from the two error sources indicated that adjustment error can plausibly be two to three times larger than sampling error. This analysis was consistent with recent empirical work by Sharma et al. (1996), which investigated the effect of adjustment errors in estimating MDT, as well as with work which highlighted the error caused by applying adjustment factors developed for traffic dominated by passenger cars to estimate the MDT of heavy trucks (Hallenbeck and Kim 1994; Cambridge Systematics 1994). The review also pointed out that both sampling and adjustment error can be explicitly accounted for within a hierarchical statistical model of the process generating the daily traffic counts and that this model can be used to develop an empirical Bayes (EB) estimator of MDT, which does not require that each roadway section be assigned a priori to a factor group (Davis and Guan 1996; Davis 1997a). Rather, a structure similar to that shown in equations (1) and (2) is used, in which the sample data are used to assess the posterior probabilities the sample site belongs to each factor group. The MDT is then estimated as a weighted average, with the factor group probabilities providing the weights. The next section describes how this hierarchical modeling approach can be extended to develop a method for computing the predictive distribution of a site's total traffic volume, rather than its MDT, given a traffic count sample at the site.

DERIVING THE PREDICTIVE DISTRIBUTION FOR TOTAL TRAFFIC VOLUME

Using Bayes Theorem to assess the information provided by a sample and then integrating out nuisance parameters, the two steps exemplified in equations (1) and (2) provide the basic framework for deriving the predictive distribution of traffic totals from more realistic assumptions. In the above example, we derived a predictive probability density $f\left(z^N|\mathbf{z}\right)$, but here we will focus on the corresponding predictive distribution function $F\left(z^N|\mathbf{z}\right)=P\left[Z^N \le z^N|\mathbf{z}\right]$. The distribution function is more useful from a practical standpoint since it leads immediately to a method for finding the quantiles of the predictive distribution by solving equations of the form $F\left(z^N|\mathbf{z}\right)=\alpha$. Since the expression for the cumulative distribution function turns out to have the form of a weighted average, an argument similar to that employed below could also be used to find the predictive probability density or the moments of the predictive distribution.

Aleatory Uncertainty

We will develop an explicit expression for $F\left(z^N|\mathbf{z}\right)$ in several steps. As in the example, the total traffic count z^N is determined as the sum of the daily counts z_t, so that the statistical properties of the daily traffic volumes will determine the form of the conditional distribution $F\left(z^N|\mathbf{z}\right)$. The first step is to characterize the statistical properties of the daily traffic volumes z_t. In order to develop plausible statistical models for the day-to-day fluctuations in traffic volumes at a particular site, a detailed analysis of daily traffic counts from 50 ATRs in Minnesota was undertaken, as described in Davis (1997a) and Davis and Guan (1996). This work indicated that variations in daily traffic volumes could be described using a lognormal regression model of the form

$$y_t = u + \sum_{i=1}^{12}\Delta_{t,i}m_{k,i} + \sum_{j=1}^{7}\delta_{t,j}w_{k,j} + e_t \qquad (4)$$

where

$y_t = \log_e(z_t)$, the natural logarithm of a daily count,

u = expected log traffic count on a typical day,

$\Delta_{t,i}=1$, if the count z_t was made during month i, i=1,...,12, (0 otherwise)

$m_{k,i}$ = correction term for month i, characteristic of factor group k,

$\delta_{t,j}$ = 1, if the count z_t was made on day-of-week j, j=1,…,7, (0 otherwise)

$w_{k,j}$ = correction term for day-of-week j, characteristic of factor group k,

e_t = random error.

If we let $\beta_k=\left[m_{k,1},\ldots,m_{k,12},w_{k,1},\ldots,w_{k,7}\right]'$ denote a column vector containing the monthly and day-of-week adjustment terms for factor group k, and $x_t=\left[\Delta_{t,1},\ldots,\Delta_{t,12},\delta_{t,1},\ldots,\delta_{t,7}\right]$, equation (4) can be written in a slightly simpler form:

$$y_t = u + x_t\beta_k + e_t \qquad (5)$$

In the above model, the mean value of the logarithm of the daily count varies according to month and day-of-week, and the magnitude of these variations depends on the factor group to which the site of interest belongs. Analysis of the regression residuals obtained after estimating the adjustment terms indicated that the error terms e_t were not independent but showed day-to-day dependencies, which could be described by a multiplicative autoregressive (AR) model of the form

$$e_t = \phi_1 e_{t-1} + \phi_7 e_{t-7} - \phi_1\phi_7 e_{t-8} + a_t \qquad (6)$$

Here the a_t are independent, identically distributed, normal, random variables with zero mean and common variance, and ϕ_1 and ϕ_7 are site-specific autoregressive coefficients.

The above model is parameterized by u, a mean-value parameter, β, the monthly and day-of-week adjustment terms, the variance of the e_t terms, which we will denote by σ^2, and the autoregressive coefficients ϕ_1, ϕ_7. In the next step, we will assume we know the values of these parameters but nothing else about the site. Properties of lognormal random variables (Shimizu and Crow 1988) can be used to show that the expected value of the total traffic volume is

$$\mu_N = E\left[z^N\right] = \sum_{t=1}^{N}\exp\left(u + x_t\beta_k + \tfrac{\sigma^2}{2}\right) \qquad (7)$$

and the variance of the total traffic volume is

$$v_N^2 = e^{(2\mu+\sigma^2)}\left(e^{\sigma^2} - 1\right)\sum_{t=1}^{N} e^{2x_t\beta_k}$$
$$+\left(e^{2u+\sigma^2}\right)\left(\sum_{t=1}^{N} e^{x_t\beta_k}\left(\sum_{s\neq t} e^{x_s\beta_k}\left(e^{\rho_{t,s}\sigma^2} - 1\right)\right)\right) \quad (8)$$

Here $\rho_{t,s}$ denotes the correlation between e_t and e_s, which can be computed from ϕ_1, ϕ_7 (see Brockwell and Davis 1991). Since z^N is a sum of lognormal random variables, characterizing its probability distribution turns out to be very difficult even for the case N=2, and no result for general N is known (Johnson, Kotz, and Balakrishnan 1994). In most situations of practical interest, however, N will be equal to the number of days in one or more years, so an asymptotic approximation might prove useful. In the Appendix, a result due to Marlow (1967) is adapted to show that the cumulative distribution function of the random variable

$$\frac{\mu_N}{v_N}\log_e\left(\frac{z^N}{\mu_N}\right) \quad (9)$$

converges to that of a standard normal random variable, implying that for large N, $\log_e(z^N)$ can be approximated by a normal random variable with mean equal to $\log_e(\mu_N)$ and variance equal to $\left(v_N^{\ 2} / \mu_N^{\ 2}\right)$. This, in turn, supports using a lognormal approximation for z^N,

$$P\left[z^N \leq \bar{z} \middle| u, \beta, \phi_1, \phi_7, \sigma^2\right] =$$
$$P\left[\log_e\left(z^N\right) \leq \log_e(\bar{z}) \middle| u, \beta, \phi_1, \phi_7, \sigma^2\right] \quad (10)$$
$$\approx \Phi\left(\frac{\log_e(\bar{z}) - \log_e(\mu_N)}{\frac{v_N}{\mu_N}}\right)$$

where $\Phi(\cdot)$ denotes the standard normal distribution function.

Epistemic Uncertainty

The sample $\boldsymbol{z}$ contains two types of information concerning z^N. On the one hand, it provides information concerning the model parameters. In principle, this information could be summarized by the posterior distribution function $F\left(u, \beta, \phi_1, \phi_7, \sigma^2 | z\right)$. On the other hand, the correlated noise in equation (6) implies that any given daily count z_t is correlated with its neighbors, so knowing z_t allows us to more accurately predict neighboring values. If the noise equation (6) is stationary, and if the sample counts are sufficiently separated in time from the counts comprising z^N (such as might occur when trying to predict the total volume for 1999 using a sample taken in 1997) we can take the sample and the total count to be independent of each other. Then the sample provides information on the total only by providing information on the model parameters. The predictive distribution $F\left(z^N | z\right)$ could then be found using a generalization of equation (2):

$$F\left(z^N | z\right) = \int F\left(z^N \middle| u, \beta, \phi_1, \phi_7, \sigma^2\right) dF\left(u, \beta, \phi_1, \phi_7, \sigma^2 | z\right) \quad (11)$$

When the sample counts are correlated with counts comprising the total z^N, the expression (11) will only be approximate, with the approximation deteriorating with increasing overlap between the sample counts and the counts entering into the total. In principle, smoothing algorithms could be used to include dependency on $\boldsymbol{z}$ in (7) and (8), so that a $\boldsymbol{z}$-dependent asymptotic approximation could be developed, but the necessary computational labor appears to be substantial. For the situations commonly encountered in highway and safety engineering, the number of counts entering into the total will be large compared to the size of the sample (for example, one or more years for N, compared to a 48-hour or two one-week counts for n), so that most of the aggregating counts will be separated from the sample counts by at least one month. For these cases, we conjecture that equation (11) will provide a suitably accurate approximation. Some empirical evidence supporting this conjecture will be presented in the fourth section.

The final steps involve characterizing the distribution $F\left(u, \beta, \phi_1, \phi_7, \sigma^2 | z\right)$ and then finding a computationally feasible way to evaluate the (multidimensional) integral in (11). It turns out, however, that this problem is very similar to the problem of computing Bayes estimates of mean daily traffic described in Davis (1997a) and Davis and Guan (1996), and a similar solution can be employed here. The essence of this approach is to assess the prior uncertainty concerning the model parameters, and then use Bayes Theorem to account for information provided by the data sample.

As in Davis (1997a), we will assume that the highway agency has divided its road segments into a set of m factor groups and that estimates of the adjustments factors for each group, β_k, k=1,...,m, are available. We will further assume that the agency maintains a total of M ATRs and that for each ATR estimates of the covariance parameters $\left(\phi_{1p}, \phi_{7p}, \sigma_p\right)$, $p = 1, \ldots, M$, are also available. Straightforward procedures for computing these estimates from ATR data, using commonly available software packages, are described in Davis (1997a). Prior to collecting any data for a site, we will assume that our uncertainty concerning that site's parameters is captured by the prior probability distribution

$$f(u, \beta, \sigma, \phi_1, \phi_7) = \left(\frac{1}{M}\sum_{p=1}^{M} I_{(\sigma_p, \phi_{1p}, \phi_{7p})}(\sigma, \phi_1, \phi_7)\right)\left(\frac{1}{m}\sum_{k=1}^{m} I_{\beta_k}(\beta)\right) \quad (12)$$

where $I_b(x)$ denotes the indicator function

$I_b(x) = 1$, if x=b, (0 otherwise).

Basically, this prior assumes that before collecting data we are completely uncertain of the value of u in the sense that our prior probability is uniformly distributed on the real line. For the adjustment term β, we are certain it takes on one of the values $\beta_1, \ldots, \beta_m$ characterizing our factor groups, but we are equally uncertain which of these is correct. Similarly, for (ϕ_1, ϕ_7, σ) we are certain one of the sets of values estimated from our ATR sites is correct, but prior to collecting data we are equally uncertain which one. Completing the specification of this prior by generating estimates of the β_k and $\left(\phi_{1p}, \phi_{7p}, \sigma_p^2\right)$ from ATR data results in an empirical Bayes (EB) method, in the sense of Padgett and Robinson (1978). That is, empirical distributions from samples are used to form the priors.

Because the logarithms of the traffic counts are normal random variables, the likelihood function of the sample is easy to specify. Letting **y** denote the vector containing the logarithms of the sample counts and V denote the correlation matrix of **y** (which can be computed once the value of the AR parameters ϕ_1 and ϕ_7 is known), then if we knew the site-specific values for the parameters u, β, σ, ϕ_1, and ϕ_7, the likelihood of the sample could be computed using the appropriate multivariate normal density.

$$f(y|\beta, u, \sigma, \phi_1, \phi_7) = (2\pi)^{-n/2}\sigma^{-n}|V|^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(y - X\beta_k - u1_n)' V^{-1}(y - X\beta_k - u1_n)\right) \quad (13)$$

Here X is a matrix, of dimension N×19, each row having elements equal to 0 or 1, according to the month and day-of-week of the corresponding sample count, while $\mathbf{1}_n$ is an n-dimensional column vector with each element equal to 1.0.

Predictive Distribution of Total Traffic

Applying Bayes Theorem to the prior and likelihood to obtain the posterior distribution for the parameters, substituting this into (11), and performing the indicated integrations produces, after some tedious algebra,

$$P\left[z^N \le \bar{z} \,|\, z\right] = P\left[y^N \le \tilde{y} | z\right] \approx \frac{\sum_{p=1}^{M} \sigma_p^{-(n-1)} \left|V_p\right|^{-1/2} \left(1_n' V_p^{-1} 1_n\right)^{-1/2} \sum_{k=1}^{m} \Phi\left(\frac{\tilde{y} - \hat{y}_{p,k}}{S_{p,k}}\right) \exp\left(-\frac{S_{pk}^2}{2\sigma_p^2}\right)}{\sum_{p=1}^{M} \sigma_p^{-(n-1)} \left|V_p\right|^{-1/2} \left(1_n' V_p^{-1} 1_n\right)^{-1/2} \sum_{k=1}^{m} \exp\left(-\frac{S_{pk}^2}{2\sigma_p^2}\right)} \quad (14)$$

where y^N=log$_e$(z^N), V_p denotes the sample correlation matrix computed using $\left(\phi_{1p}, \phi_{7p}\right)$, and

$$\begin{aligned} S_{pk}^2 &= \left(y - X\beta_k - \bar{y}_{pk} 1_n\right)' V_p^{-1}\left(y - X\beta_k - \bar{y}_{pk} 1_n\right) \\ \bar{y}_{pk} &= \frac{1_n' V_p^{-1}(y - X\beta_k)}{1_n' V_p^{-1} 1_n} \\ \hat{y}_{pk} &= \bar{y}_{pk} + \frac{\sigma_p^2}{2} + \log_e\left(\sum_{t=1}^{N} e^{x_t \beta_k}\right) \\ s_{pk} &= \left(\frac{\sigma_p^2}{1_n' V_p^{-1} 1_n} + \frac{\upsilon_{N,kp}^2}{\left(\mu_{N,kp}\right)^2}\right)^{-1/2} \end{aligned} \quad (15)$$

where $\upsilon_{N,kp}^2$ and $\mu_{N,kp}$ are as defined in (7) and (8) but evaluated using β_k and $\left(\Phi_{1p} \Phi_{7p} \sigma_p\right)$. The distribution given in (14) is a finite mixture of normal distributions where the weights given to the mixture components are the posterior probabilities that the sampled site has adjustment factors and covariance parameters characteristic of each the m factor groups and each of the M ATR sites. Although the expressions in (14) and (15) appear rather forbidding, the implied computations are readily carried out on a personal computer.

A CALIBRATION TEST

As noted above, the distribution (14) approximates the predictive distribution of a total traffic count, the approximation being appropriate when predicting the total of a large number of days (for example, a year or more) from a small sample (for example, two weeks or less). In an earlier study, Davis (1997a) used traffic counts from the year 1992 from 50 ATRs in outstate Minnesota to estimate monthly and day-of-week adjustment terms for the Minnesota Department of Transportation's (Mn/DOT) 3 outstate factor groups, as well as covariance parameters for each of the 50 ATRs. These estimates were then used to construct the discrete prior distributions for β and (ϕ_1, ϕ_7, σ), giving m=3 and M=50. In addition, daily counts from the year 1991 were available for 48 ATRs, and for each of these ATRs a sample consisting of a one-week count from the month of March and a one-week count from the month of July was drawn. The 1992 data were used for estimation, and 1991 data were used for validation because more ATRs had good data in 1992. A MATLAB (Mathworks 1992) program for evaluating (14) was written, and then for each of the 48 ATRs, the 5th and 95th percentile points of the predictive distribution of the logarithm of the 1991 total traffic volume were computed by embedding this routine inside MATLAB's root-finding algorithm. Finally, the logarithm of the total 1991 actual traffic volume was also computed for each ATR. The results of these computations are displayed in tables 1 through 3.

Note that the 5th and 95th percentile points describe the bounds of a 90% credible interval, and, clearly, if a large number of actual traffic counts fell outside the bounds of our intervals, we would have evidence for inaccurate prediction. On the other hand we would still expect a few actual counts to fall outside our bounds. If the intervals caught all actual volumes, we would be inclined to believe that the computed credible intervals were too large. If the approximation is acceptably accurate, we would expect the actual count to fall outside the bounds 10% of the time, and a test of the adequacy of the estimated credible bounds can be made by treating the number of missed totals as the outcome of a binomial random variable with 48 trials and a hypothesized miss probability of p=0.1. Inspection of the tables shows that for 8 of the ATRs (2, 8, 12, 204, 208, 217, 218, and 226) the actual count fell outside the estimated bounds, for a total of 8 binomial "successes." Since the probability of obtaining 8 or more successes by chance is 0.102, this result is not inconsistent with the hypothesis that equation (14) provides a reasonable approximation of the predictive distribution.

CONCLUSIONS

Predicted or estimated traffic totals are required in both highway pavement and safety engineering and are used to produce statewide and nationwide estimates of total distance traveled. Although recommended practices exist for estimating traffic totals as part of a traffic monitoring program, it is less clear how we should characterize the uncertainty associated with these estimates. This paper describes an initial solution to this problem, in which empirical Bayes methods are used to compute the quantiles of a traffic total's predictive distribution, given a sample of daily traffic volumes. Probable ranges and their associated probability values are readily found, and, if desired, a point prediction of the total traffic can be obtained as the median, or 50th percentile, of the predictive distribution. The method of derivation can also be used to find the predictive density and the moments of the predictive distribution. No data are required beyond that routinely collected by statewide traffic monitoring programs, and the estimates of the factor group adjustment parameters can be computed using standard linear regression methods. All other computations have been successfully implemented as MATLAB macros.

In conclusion, almost all engineering decisions must be made in the face of uncertainty, and the art

TABLE 1 Evaluation of 90% Prediction Intervals: First Factor Group

Mn/DOT ATR number	Fifth percentile of predictive distribution	Log total count	95th percentile of predictive distribution
2	**12.2954**	**12.5539**	**12.5531**
3	13.9783	14.1189	14.2267
7	12.2483	12.3857	12.4690
8	**10.9598**	**11.3207**	**11.2912**
9	11.8849	11.9756	12.1124
10	12.8934	13.0059	13.0748
12	**13.1947**	**13.1813**	**13.4464**
14	12.9814	13.0870	13.2310
50	13.4775	13.5663	13.6500
54	10.2983	10.5153	10.5339
56	11.4603	11.5531	11.6512
100	15.2954	15.4013	15.5066
102	16.4250	16.4932	16.6343
103	16.0997	16.1653	16.2748
104	15.2014	15.2137	15.2945
110	15.6586	15.7102	15.8464
164	13.8722	13.9628	14.0880
166	13.8270	13.8729	13.9741
170	13.8047	13.8419	13.9555
172	14.9882	15.1130	15.1175
179	13.0413	13.1812	13.2469
197	13.7666	13.8906	13.9826
199	12.4498	12.6153	12.6332
211	13.4673	13.5737	13.6714
213	14.0399	14.0975	14.1949
216	12.4219	12.5007	12.6338
217	**11.9371**	**12.1674**	**12.1390**
219	13.2525	13.2984	13.4746
225	12.0633	12.2808	12.3005
226	**11.5859**	**11.7921**	**11.7918**

NOTE: Bolding indicates ATRs that have counts falling outside the credible bounds.

of successful engineering requires cost-effective hedging against this uncertainty. It was argued earlier that standard methods for predicting total traffic ignore potentially important sources of error, and, hence, understate the resulting uncertainty characterizing estimates and predictions. Many of the statistical procedures used in highway engineering date to the middle part of the 20th century and

TABLE 2 Evaluation of 90% Prediction Intervals: Second Factor Group

Mn/DOT ATR number	Fifth percentile of predictive distribution	Log total count	95th percentile of predictive distribution
52	11.4197	11.5725	11.6027
175	15.2607	15.3471	15.4842
187	14.9642	15.0832	15.2037
200	15.7115	15.8787	15.9226
204	**14.0265**	**14.2724**	**14.2705**
207	12.8435	12.9302	13.0395
208	**14.8488**	**14.9436**	**14.9377**
215	11.7087	11.8455	11.9711
218	**11.5948**	**11.9728**	**11.8361**
224	13.1007	13.2427	13.3437

NOTE: Bolding indicates ATRs that have counts falling outside the credible bounds.

TABLE 3 Evaluation of 90% Prediction Intervals: Third Factor Group

Mn/DOT ATR number	Fifth percentile of predictive distribution	Log total count	95th percentile of predictive distribution
1	12.9936	13.0798	13.2428
51	11.4256	11.5018	11.6581
55	12.0286	12.1631	12.1852
57	12.4938	12.7573	12.7840
214	11.5181	11.6925	11.7181
220	13.2472	13.3545	13.4526
221	13.4560	13.4819	13.6290
223	12.8283	12.9375	12.9596

are based on simplified statistical models adapted to the computational constraints of those times. Statistical science has advanced considerably since then, and these advances can support and encourage the use of more realistic models in highway engineering. This paper proposes a modest step in this direction by providing a computationally practical method which accounts for uncertainty in traffic volume predictions. Of course, the importance of hedging against uncertainty depends on the consequences of error, and, fortunately, so far the consequences attached to using mistakenly precise traffic forecasts have not been too severe. Whether or not this state of affairs continues is of course another uncertain prediction about the future.

ACKNOWLEDGMENTS

The authors would like to thank Mark Flinner of Mn/DOT for providing the traffic-count data used in this study. This research was sponsored by Mn/DOT. However, all facts, conclusions and opinions expressed here are solely the responsibility of the authors and do not necessarily reflect the views of Mn/DOT.

REFERENCES

AASHTO. 1992. *Guidelines for Traffic Data Programs*. Washington, DC.

____. 1993. *Guide for Design of Pavement Structures*. Washington, DC.

Box, G. and G. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. New York: Wiley and Sons.

Brockwell, P. and R. Davis. 1991. *Time Series: Theory and Methods*. New York: Springer-Verlag.

Cambridge Systematics. 1994. *Use of Data from Continuous Monitoring Sites*. Report to Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

Davis, G. 1997a. *Estimation Theory Approach to Monitoring and Updating Average Daily Traffic*. Report 97-05, Minnesota Department of Transportation, St. Paul, MN.

Davis, G. 1997b. Accuracy of Estimates of Mean Daily Traffic: A Review. *Transportation Research Record* 1593:12–6.

Davis, G. and Y. Guan. 1996. Bayesian Assignment of Coverage Count Locations to Factor Groups and Estimation of Mean Daily Traffic. *Transportation Research Record* 1542:30–7.

Draper, D. 1995. Assessment and Propagation of Model Uncertainty. *Journal of the Royal Statistical Society* B 57:45–97.

Foster, K. and P. Huber. 1999. *Judging Science: Scientific Knowledge and the Federal Courts*. Cambridge: MIT Press.

Gallant, R. and H. White. 1988. *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. New York: Basil-Blackwell.

Hacking, I. 1975. *The Emergence of Probability*. Cambridge: Cambridge University Press.

Hallenbeck, M. and S-G. Kim. 1994. *Truck Loads and Flows, Task A Final Technical Report*. Washington State Transportation Research Center (TRAC), Seattle, WA.

Howson, C. and P. Urbach. 1993. *Scientific Reasoning: The Bayesian Approach*. Chicago: Open Court.

Johnson, N., S. Kotz, and N. Balakrishnan. 1994. *Continuous Univariate Distributions, Volume 1*. Second edition. New York: Wiley and Sons.

Marlow, N. 1967. A Normal Limit Theorem for Power Sums of Independent Random Variables. *Bell Systems Technical Journal* 46:2081–9.

Mathworks. 1992. *MATLAB Reference Guide,* Version 4. Mathworks, Inc.

Padgett, W. and J. Robinson. 1978. Empirical Bayes Estimators of Reliability for Lognormal Failure Models. *IEEE Transactions on Reliability* R-27:223–336.

Parker, M. 1991. *Highway Safety Engineering Studies: Procedural Guide*. Report No. FHWA-HI-88-039. Federal Highway Administration, U.S. Department of Transportation, Washington, DC.

Sarewitz, D. and R. Pielke. 2000. *Prediction in Science and Policy. Prediction: Science, Decision Making, and the Future of Nature*. Washington, DC: Island Press.

Sharma, S., B. Gulati, and S. Rizak. 1996. Statewide Traffic Volume Studies and Precision of AADT Estimates. *ASCE Journal of Transportation Engineering* 122:430–9.

Shimizu, K. and E. Crow. 1988. History, Genesis, and Properties. *Lognormal Distributions: Theory and Applications*. New York: Marcel Dekker.

Stewart, T. 2000. Uncertainty, Judgment, and Error in Prediction. *Prediction: Science, Decision Making, and the Future of Nature*. Washington, DC: Island Press.

U.S. Department of Transportation (USDOT), Federal Highway Administration (FHWA). 1995. *Traffic Monitoring Guide, Third Edition*. Washington, DC.

APPENDIX

Weak Convergence of Sums For a Class of Correlated Lognormal Random Variables

As above, let z_t, t=1,…,N denote a sequence of lognormal random variables with $z^N = \sum_t z_t$ denoting their sum. $\Phi(x)$ denotes the standard normal distribution function. As before, we will assume that the z_t follow the model

$$\log_e(z_t) = u + x_t\beta + e_t \qquad (A1)$$

and the error terms $\{e_t\}$ follow a stationary p-order autoregressive (AR(p)) process. Marlow (1967) showed that if there exists sequences of positive real number $\{\mu_N\}$ and $\{\nu_N\}$ such that

$$\text{(a)} \quad \lim_{N\to\infty} \frac{\nu_N}{\mu_N} = 0 \qquad (A2)$$

$$\text{(b)} \quad \lim_{N\to\infty} P\left[\frac{z^N - \mu_N}{\nu_N} \le x\right] = \Phi(x), \qquad (A3)$$

then

$$\lim_{N\to\infty} P\left[\left(\frac{\mu_N}{\nu_N}\right)\log_e\left(\frac{z^N}{\mu_N}\right) \le x\right] = \Phi(x) \qquad (A4)$$

To verify conditions (a) and (b), we will impose the restriction that the monthly and day-of-week factors for any given day are bounded from above and also bounded away from zero. That is, there exist constants Δ_1 and Δ_2, such that

$$0 < \Delta_1 \le e^{x_t \beta_k} \le \Delta_2 < \infty \tag{A5}$$

for all t and k. It is then possible to show that

$$\begin{aligned} v_N = &\left(e^{(2\mu+\sigma^2)}\left(e^{\sigma^2}-1\right)\sum_{t=1}^{N} e^{2x_t\beta_k} + \right.\\ &\left.\left(e^{2u+\sigma^2}\right)\left(\sum_{t=1}^{N} e^{x_t\beta_k}\left(\sum_{s\neq 1} e^{x_s\beta_k}\left(e^{\rho_{t,s}\sigma^2}-1\right)\right)\right)\right)^{1/2} \quad (A6)\\ &\le \left((N)e^{2u+\sigma^2+2\Delta_2}(1+\gamma)\right)^{1/2} = O\left(N^{1/2}\right) \end{aligned}$$

where

$$\gamma = \sum_{k=1}^{\infty}\left(e^{\rho_k\sigma^2}-1\right) < \infty \tag{A7}$$

whenever the ρ_k are the autocorrelations for a stationary AR(p) process. Similarly,

$$N\left(e^{u+\frac{\sigma^2}{2}+\Delta_1}\right) \le \mu_N = \left(e^{u+\frac{\sigma^2}{2}}\right)\sum_{t=1}^{N} e^{x_t\beta} \le N\left(e^{u+\frac{\sigma^2}{2}+\Delta_2}\right) \tag{A8}$$

so that

$$\frac{v_N}{\mu_N} = O\left(N^{-1/2}\right) \tag{A9}$$

and condition (a) is satisfied.

If the daily counts z_t were independent, we could use either the Lyaponuv or Lindbergh central limit theorems to verify condition (b), as was done by Marlow (1967). A more general central limit theorem, allowing for dependence of the sort generated by the AR(p) model for the errors e_t, is stated in Theorem 5.3 of Gallant and White (1988, 76). In particular, if we can show that

$$1)\ \left\|z_t - E\left[z_t\right]\right\|_4 \le \Delta < \infty, \text{ for all } t, \tag{A10}$$

$$2)\ \text{supremum}_t\left(\left\|z_t - E\left[z_t \mid e_{t+m},\ldots,e_t,\ldots,e_{t-m}\right]\right\|_2\right) = O(1/m).$$

3) the noise process $\{e_t\}$ is "α – mixing" of size – 4,

4) $v_N^{(-2)} = O(1/N)$,

then condition (b) will be satisfied, and we are done.

1) Let $\omega = \exp\left(\sigma^2\right)$, and since z_t is lognormal, its fourth central moment is known, so that

$$\begin{aligned} \left\|z_t - E\left[z_t\right]\right\|_4 &= \left(\omega^2(\omega-1)^2\left(\omega^2+2\omega^3+3\omega^2-3\right)e^{4(u+x_t\beta)}\right)^{1/4}\\ &\le \left(\omega^2(\omega-1)^2\left(\omega^2+2\omega^3+3\omega^2-3\right)e^{4u}e^{4\Delta_2}\right)^{1/4} < \infty \end{aligned}$$

2) This condition is satisfied trivially since

$$E\left[z_t \mid e_{t+m},\ldots,e_t,\ldots,e_{t-m}\right] = E\left[z_t \mid e_t\right] = z_t$$

implies

$$\text{supremum}_t\left(\left\|z_t - E\left[z_t \mid e_{t+m},\ldots,e_t,\ldots,e_{t-m}\right]\right\|_2\right) = 0$$

3) This condition is also satisfied trivially since the fact that the noise process $\{e_t\}$ is a stationary AR(p) process implies that it is α – mixing of all orders.

4) This follows from the fact, demonstrated above, that $v_N^{\ 2} = O(N)$.

Creating Land-Use Scenarios by Cluster Analysis for Regional Land-Use and Transportation Sketch Planning

JOSHUA SMITH
Gannett Fleming, Inc.

MITSURU SAITO
Brigham Young University

ABSTRACT

This study explores how cluster analysis can be used to categorize a large number of planning districts in a region into a smaller, manageable number of land-use scenarios consisting of planning districts of similar land-use patterns whose mean land-use distributions can be used as future land-use alternatives for those planning districts. We used Utah's Wasatch Front region for the analysis. After applying a family of cluster analysis methods, we were able to group the 343 planning districts in the region into 35 land-use planning scenarios. A combination of the Ward's linkage method, the Squared Euclidean distance measure, and the Z-score standardization of variables produced the most logical clustering of planning districts for the region.

INTRODUCTION

A recent survey on transportation planning issues and needs for planning research conducted by the Transportation Research Board indicated that one quarter of respondents identified research relating to land-use planning as a top-priority topic area (TRB 2000). Land-use and transportation systems interact to form an urban landscape, and the two components must be considered together in trans-

Mitsuru Saito, Department of Civil and Environmental Engineering, Brigham Young University, 368 Clyde Building, Provo, UT 84602. Email: msaito@byu.edu.

portation planning to create a livable urban area (Vuchic 1999). For many metropolitan planning organizations (MPOs), this ideal has been difficult, if not impossible, to carry out because local governments have jurisdiction over land-use planning, whereas regional transportation planning is often done by state agencies. Often, land-use and transportation planning done separately have resulted in undesirable urban sprawl and traffic congestion in urbanized areas. The development of urban planning procedures that integrate land-use and transportation planning while allowing all participants access to the decisionmaking process is needed for transportation planning in the new century.

In a study funded by the National Science Foundation, Balling and others developed a multiobjective genetic algorithm model to simultaneously optimize a land-use and transportation network within a city (Balling et al. 1999; Taber et al. 1999). This procedure quickly examines an extremely large search set of feasible plans and narrows the number of alternatives to be considered. The model optimizes land-use and transportation network plans with the objective of minimizing travel time on the street network, cost to the city, and change from the current status, typically politically infeasible. Other objective functions dealing with more current trends in land-use and transportation planning are currently being considered as additions to the model. The model was first applied to Provo, Utah, and then to the twin cities of Provo and Orem. See figure 1 (after page 42) for their locations in Utah.

This is a new paradigm in urban planning. The model produces a set of optimized land-use and transportation infrastructure plans, and the plans are presented to those involved in the planning effort, such as city council members, planners, and citizen groups. Each plan in the Pareto set is optimal for a different weighting of the competing objectives, allowing participants the opportunity to explore compromise solutions rather than being forced to choose from only a few plans.

The second phase of the model development expands this genetic algorithm model to regional urban planning. The proposed model aims to produce macro-level Pareto plans for a multi-city metropolitan region and optimize land-use and transportation corridors between the cities. These plans will not restrict micro-planning done at the city level. The proposed model will produce optimal land-use plans that give target scenarios for land-use distribution for each planning unit. In this study, the planning units were named "districts." The model might find, for example, that a district in a particular city would best benefit regional objectives if it had a mix of 40% low-density residential land use, 20% medium- and high-density residential use, and 10% each of commercial, industrial, and open-space uses. A city cooperating with the regional planning organization would try to meet such target scenarios but would be free to plan any conceivable layout of land use within the city's planning districts in order to optimize local objectives within the framework of the regional objectives.

SCOPE AND OBJECTIVES OF THE STUDY

The very nature of the genetic algorithm requires potential scenarios of land-use distribution to be discrete rather than continuous variables. Therefore, the objective of this study was to determine the suitability of cluster analysis for the creation of just such a scenario set. The Wasatch Front region of Utah, consisting of Weber, Davis, Salt Lake, and Utah counties, was selected as the study area (see figure 1). Geographic Information Systems (GIS) data defining suitable districts did not exist. Therefore, we gathered and manipulated land-use and other necessary data to create approximately 300 planning districts (modeler's discretion), each with an approximately known land-use distribution. Once the percentage of distribution of land use for each district was found, districts were grouped, or categorized, by cluster analysis to create a set of 20 or 30 land-use distribution scenarios. The cluster means for each scenario will be used by the aforementioned planning model as the discrete values for possible future land-use scenarios.

PREPARATION OF PLANNING DISTRICTS

Data necessary for accomplishing the objective of this study were collected from various sources: the Wasatch Front Regional Council, Mountain Land Association of Governments, State of Utah Automated Geographic Resource Center, Salt Lake

County, Utah Country, and various city planners and engineers.

Most of the information needed for this study is in GIS format. The GIS files contain map shapes representing parcels (individually owned plots of land), city boundaries, and boundaries for districts created for analysis. Associated with the parcel shapes are codes for various types of land use (for example: residential–low density, residential–high density, industrial, commercial, agricultural), from which percentages for each type of land use in each district were derived. Of particular advantage was the fact that the bulk of the land-use data was based not on zoning but rather on information collected for individual parcels from county recorders' offices. Only where gaps in the land-use data existed was zoning used as a rough approximation.

Land-use data for the Wasatch Front region exist in a variety of formats and data structures. While all the cities and counties maintain some type of data, the systems of classifications used vary in scope, detail, and accuracy. Making these different land-use classification systems congruent for the entire region would be a tremendous task. In order to get the best overall picture of land-use scenarios existing along the Wasatch Front, it was decided to use one source, the State of Utah Automated Geographic Resource Center (AGRC 1997), for all land-use data wherever possible. To draw appropriate district boundaries and find the land-use scenarios that exist in those districts, these data from AGRC needed to be combined with city boundary and parcel information available from the cities, counties, and MPOs.

In order to accomplish the first objective of the study, we followed a 14-step procedure. This procedure is briefly outlined in table 1, and detailed discussions of it can be found in Smith (2000). The GeoProcessing Wizard of ArcView (ESRI 1999) and user-written Avenue scripts were used as aids in constructing district boundaries. District boundaries were constructed using a combination of city boundaries, traffic analysis zone boundaries, and city-provided neighborhood boundaries.

An attempt was made to exclude from the districts most lands not considered candidates for future development. Undevelopable lands were defined as those lands covered by water or wetlands, with gradient slopes over 25%, or owned by certain public agencies such as the Forest Service or the Division of Wildlife Resources and not available for future development.

Figure 1 shows the result of the geoprocessing work; in total we created 343 planning districts. Figure 2 (after page 42) shows parcel-level land-use data for the area covered by the 343 planning dis-

TABLE 1 GeoProcessing® Steps Used in Preparing Land-Use Data for Cluster Analysis

Step 1	Create shapefiles that describe lands that are not developable.
Step 2	Intersect the undevelopable shapefile for each county (from Step 1) with shapefiles describing county and city boundaries.
Step 3	Create shapefiles for each county that describe the distribution of land use by parcel.
Step 4	Use Avenue script to split parcel/land-use shapefiles into separate files for each city.
Step 5	Intersect parcel/land-use shapefiles with undevelopable shapefiles for each city.
Step 6	Use Avenue script to re-code undevelopable lands with an appropriate land-use code.
Step 7	Estimate areas of "unknown" parcel land use by substituting zoning data.
Step 8	Intersect the new land-use shapefiles from step 7 with shapefiles describing boundaries within each city to be used in constructing district boundaries.
Step 9	Use Avenue script to write a new shapefile for each of the districts incorporated in step 8.
Step 10	Where necessary, aggregate districts together to form districts of approximately equal size.
Step 11	Clip the aggregated district shapefiles using the city boundaries.
Step 12	Dissolve shapefiles based on land use.
Step 13	Determine land-use statistics for each district.
Step 14	Create shapefiles that contain only information on district boundaries.

tricts, along with their land-use codes. The land-use categories shown in figure 2 are the ones used by the AGRC and form the basis for this study.

CLUSTERING PLANNING DISTRICTS TO CREATE SCENARIOS

Cluster analysis was chosen to categorize the districts due to the difficulty of creating intuitive groupings for the large number of land uses. If only two variables, such as percentages of residential and commercial uses, were to be considered, it would be a simple enough matter to make a plot of percentage residential versus percentage commercial, showing points for each of the 343 districts. Boundaries could then be drawn around groups of points to separate them into the desired number of categories for different scenarios. With the 13 variables involved in this study, however, such an exercise is impossible.

Cluster analysis is a family of methods that seeks to explore the structure of a data set by defining the relationships between individual observations in the set, such as planning districts in this study. Such analysis is particularly useful when no preconceived idea of the proper manner of data classification exists. The MINITAB software package (MINITAB 1999) was used to perform cluster analysis on the land-use data.

Criteria Used

The objective of the cluster analysis was to obtain land-use scenarios categorized by a reasonable system of classification that would include the following:

- unique and concise bounds on percentages for each of the land-use variables for each scenario
- mean values for each of the land-use variables in each scenario that could be used as values representative of the entire grouping
- scenarios that cover most of the full range of percentages found for each type of land use

Clustering Issues

The usual agglomerative clustering procedure was used for this study. This means that for a given data set, each step in the analysis agglomerates, or groups together, two clusters, which may each be either individual observations or sets of observations grouped together in a previous step. Thus, analysis of a data set with 100 observations would begin by treating each observation as its own cluster. The first step would reduce the number of clusters to 99 by grouping the two clusters closest together into a new cluster. Each step would then group two more clusters together until, after 99 steps, only one cluster of 100 observations remained. The analyst would then look at the various cluster groupings at different steps in the process to decide when the observations are most appropriately categorized. We acknowledge that the subjective nature of this step is a widely held criticism of cluster analysis as a technique. Nonetheless, we feel that the cluster analysis is a useful exploration tool for land-use data since this level of subjectivity is much lower than that usually employed by planners in classifying land-use scenarios.

The manner in which closeness of observations is measured is called the distance measure. While the distance between clusters is relatively straightforward if each cluster contains only a single observation, the matter of measuring distances between clusters of many observations becomes more complex. Consequently, various linkage methods exist to determine distances between clusters containing multiple observations. When calculating distances between clusters, each variable is assumed to be on the same scale unless some standardization technique is employed.

The Distance Measure

In two-dimensional space, the distance measure may be visualized by connecting two points representing two observations, *i* and *j*. The most widely used distance measure, the Euclidean distance, is the straight-line distance between the two points, calculated in *N*-space as

$$d_{ij} = \sqrt{\left(x_{1i} - x_{1j}\right)^2 + \left(x_{2i} - x_{2j}\right)^2 + \ldots + \left(x_{Ni} - x_{Nj}\right)^2} \quad (1)$$

The Euclidean distance may be squared in order to further reduce the likelihood of very dissimilar observations being clustered together. The Pearson distance, which may also be squared, is similar to the Euclidean distance, but incorporates the variances of each variable (x_1, x_2,..., x_N) in order to reduce the portion of the distance contributed by variables with high variance (MINITAB 1999).

FIGURE 1 Map of the Study Area and 343 Districts for the Largely Developable Areas of the Wasatch Front Region

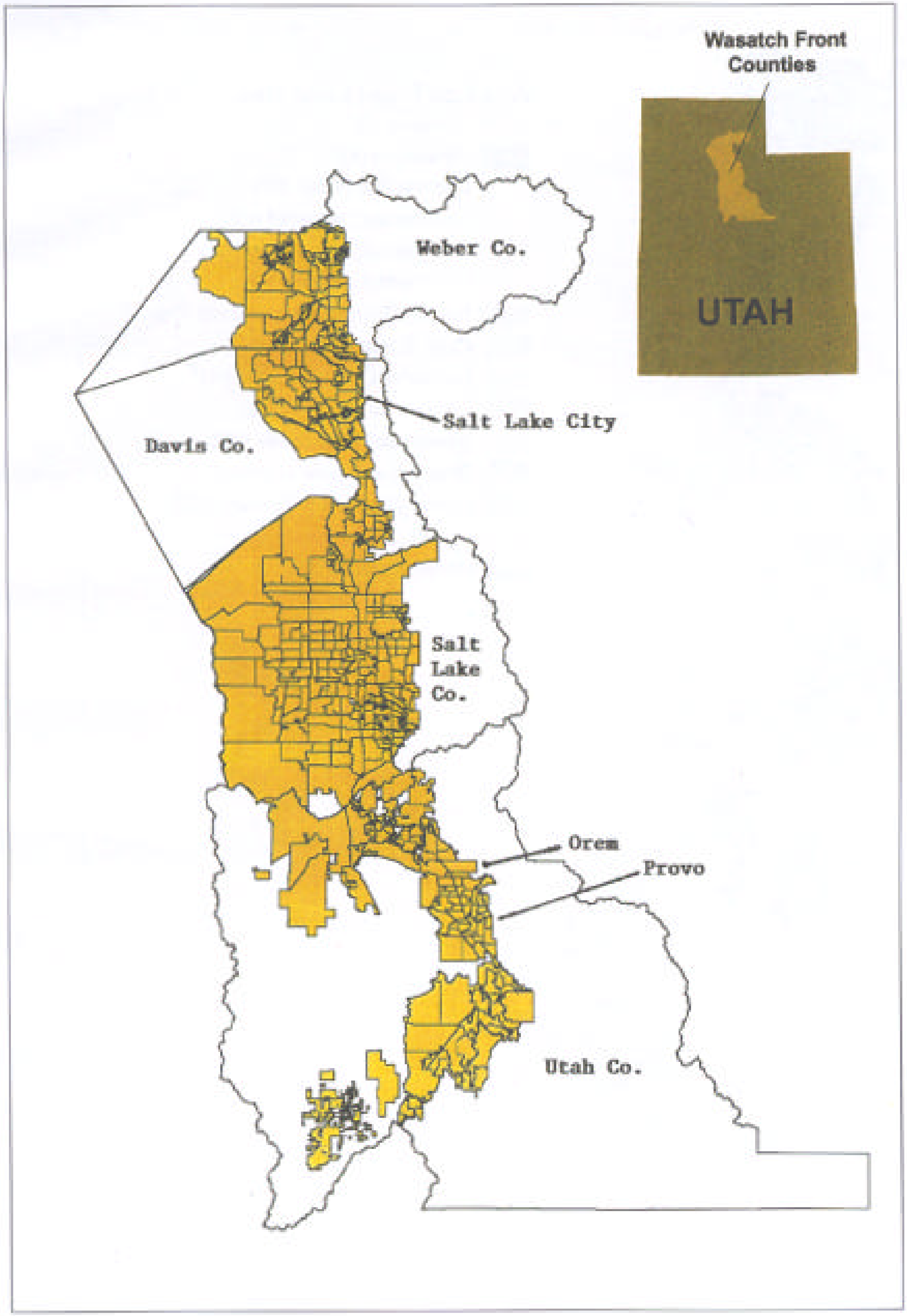


FIGURE 2 Land Use in the Wasatch Front Region

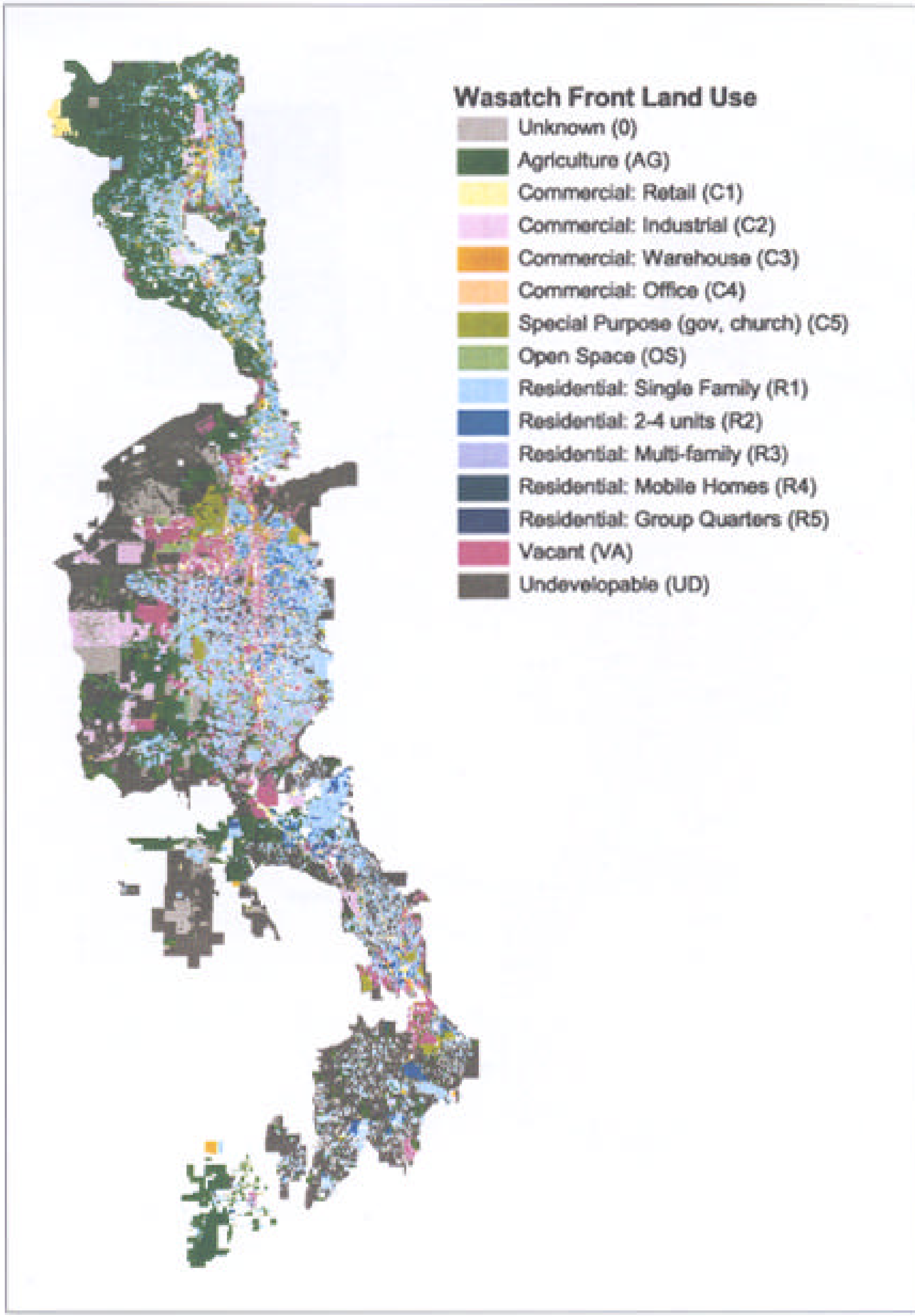


FIGURE 3 Classification of Land-Use Scenarios for 343 Districts Using Single Linkage

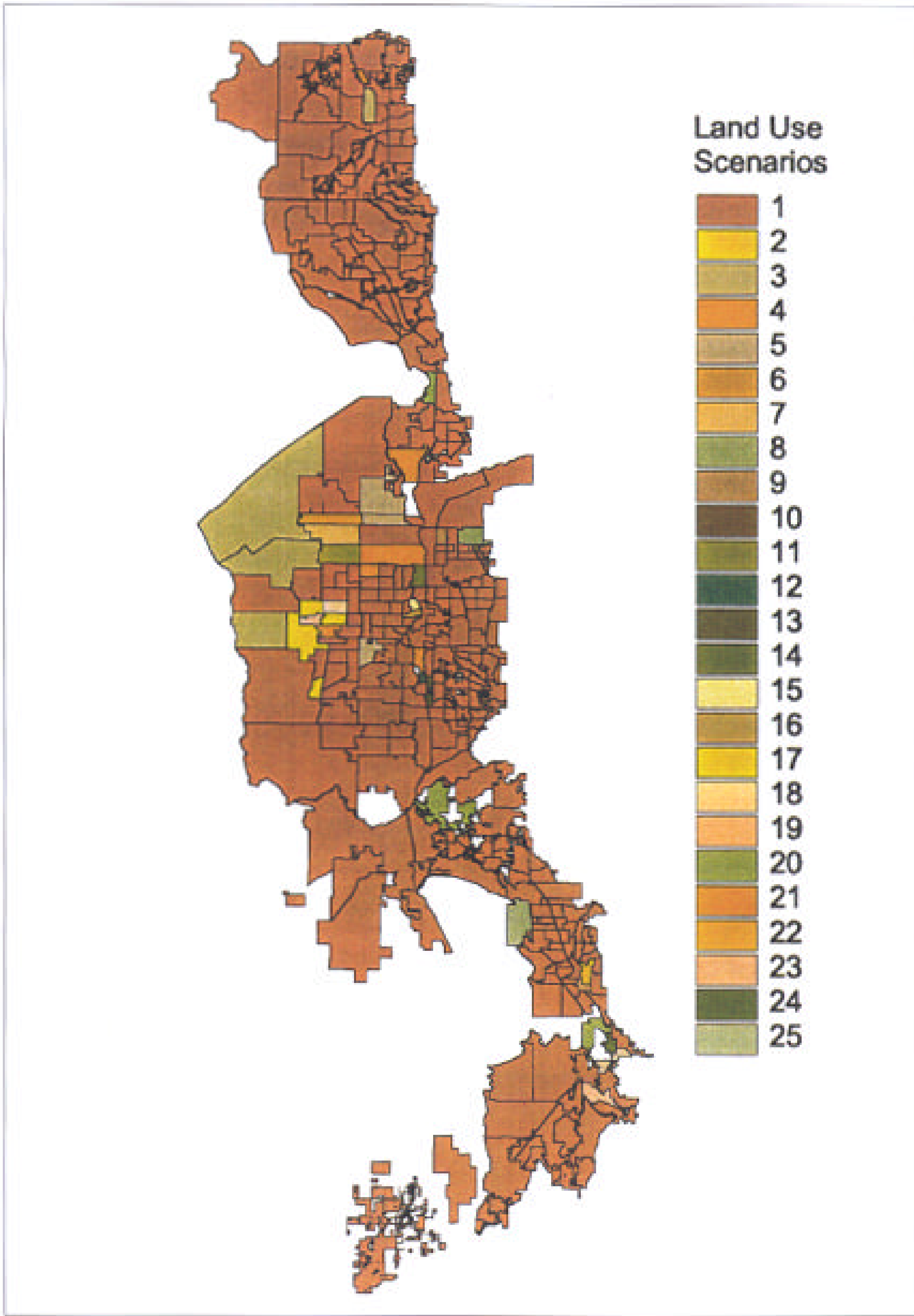
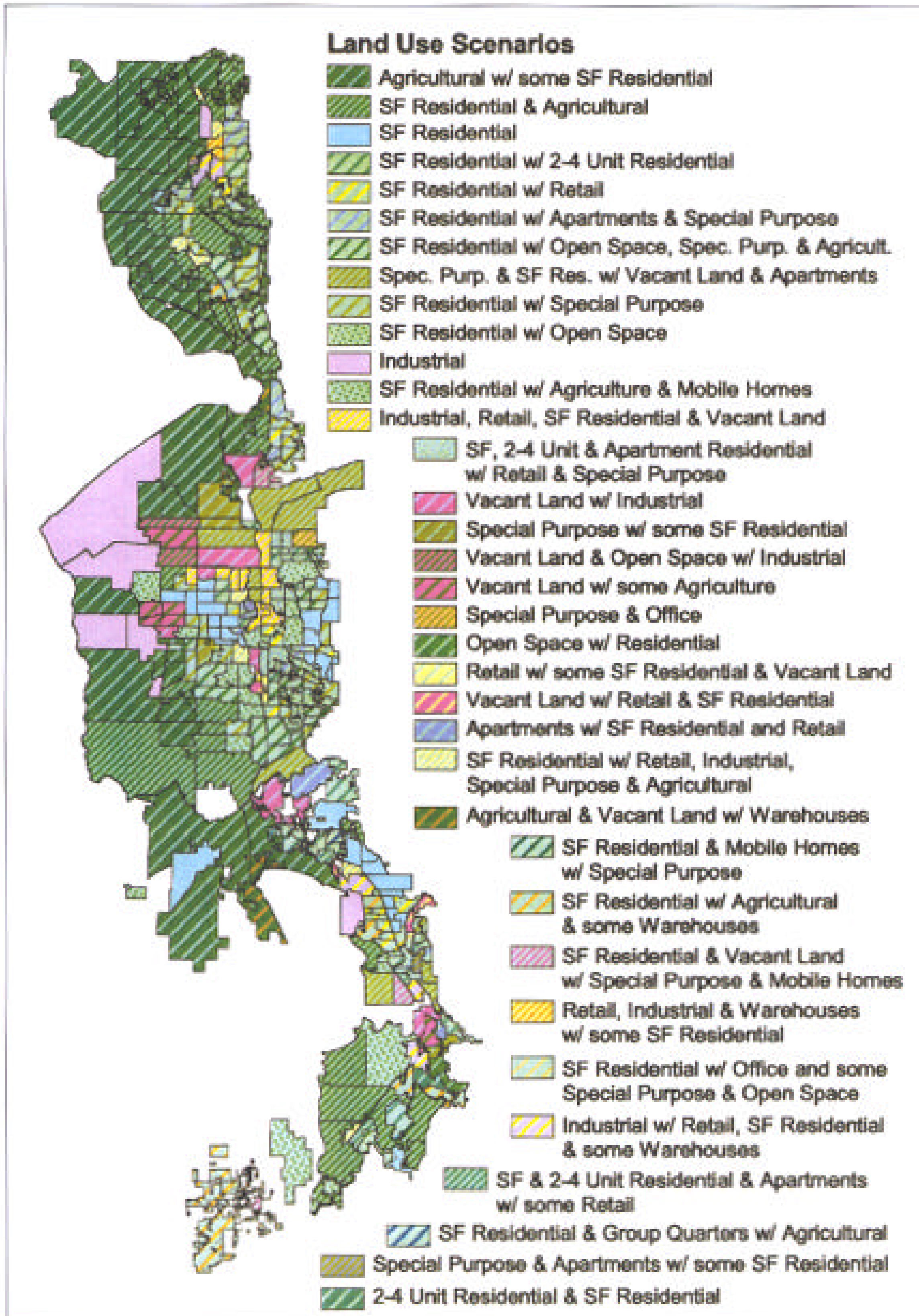


FIGURE 4 Final Classification of Land-Use Scenarios for 343 Districts Using Cluster Analysis



Another accepted measure of distance is the Manhattan distance, measured by summing the absolute values of the distances along axes between observations in *N*-space.

The Linkage Method

Seven different linkage methods available in the MINITAB software (MINITAB 1999) were used, including the single linkage, complete linkage, average linkage, centroid linkage, median linkage, McQuitty's linkage, and Ward's linkage. Detailed discussions of the linkage methods can be found in standard statistics textbooks and software user's manuals. Raising some issues relating to selecting a proper linkage method, this paper discusses two methods: single linkage and Ward's linkage.

The single linkage method defines the distance between any two clusters as the shortest distance between any observation in the first cluster and any observation in the second cluster. It was found that the chaining of observations tends to produce one very large cluster and other very small clusters. Consequently, the variable space within the large cluster is not very well explored.

In Ward's linkage, chosen as the final cluster method for this study, the distance between any two clusters is the sum of the squared deviations between the centroid and the points of the new cluster that would be formed by joining the two clusters. MINITAB uses an approximation to this distance. The objective of this method is to produce clusters with a minimal amount of within-cluster variance. Ward's linkage tends to produce clusters of similar numbers of observations since a disproportionately high number of observations in a cluster would result in a higher number of squared deviations to be summed, thus tending to increase the distance across which the cluster must be formed.

Standardization

Because the variables in this clustering problem have varying distributions, Z-score standardization was employed before calculating the distance matrix. Standardization of variables in a clustering problem can have both advantages and disadvantages. Consider the following simplification of the land-use scenario classification problem for two imaginary districts (refer to figure 2 for the land use codes shown here):

District	R1	R2	R3	R4	C1	C2	C3	AG
1	68%	1%	1%	0%	3%	2%	2%	23%
2	60%	2%	2%	9%	4%	2%	1%	20%

At first glance, these two districts may seem to be very similar in terms of their distribution of land use. Without any standardization of variables, these two districts would likely be clustered together in the same scenario type due to the low distance between them. Suppose, however, that the total range of percentages for R4, mobile homes, is only 0 to 10% for all of the districts being clustered. Also, suppose the average value for R4 is 0.5%, that its standard deviation is 1.5%, and that all of the other variables are very near their mean values. The 9% value for mobile homes is now clearly an extreme value for that variable. If such scenarios were always clustered together at relatively early stages, no diversity in mobile home land use would be apparent from the final cluster grouping. All of the cluster means for the different scenario types would reflect mobile home land use of about 0.5%. Transforming these percentages into Z-scores remedies this problem. The Z-score for R4 for district 2 would be $(9 - 0.5) / 1.5 = 5.67$, a relatively high value. Since all other variables are near their means, their Z-scores would all be near zero. The above comparison, now standardized, would look approximately like this:

District	R1	R2	R3	R4	C1	C2	C3	AG
1	0	0	0	0	0	0	0	0
2	0	0	2	5.67	0	0	0	0

Remembering that a standardized value of even one (one standard deviation) is a significant departure from the average, we can see that these districts would now be judged different enough based on the distance between them to remain in separate clusters until much later in the clustering process.

Consider, though, what might happen with these two hypothetical districts:

District	R1	R2	R3	R4	C1	C2	C3	AG
1	10%	2%	2%	9%	4%	2%	1%	70%
2	70%	2%	2%	9%	4%	2%	1%	10%

Both would have standardized R4 values of 5.76 and identical, near-zero values for R2, R3, C1, C2, and C3. Standardized values for R1 and AG would be non-zero but not nearly as extreme as 5.67 due to the approximately normal distribution for R1 and AG. These characteristics would likely cause these two districts to cluster together as districts with similarly high percentages of mobile-home use. In many ways, though, this categorization doesn't make much sense because single-family residential and agricultural uses, together accounting for 80% of the land use in the districts, are in opposite proportion to each other. The wide distribution of these variables compared to the narrow distribution for R4 is what makes these districts seem similar when standardization is applied. A primary objective in applying cluster analysis to the 343 districts was to balance these 2 effects of standardization.

RESULTS OF CLUSTER ANALYSIS

Thirty-eight cluster analyses were applied to the previously mentioned 343 districts using MINITAB statistical software in order to determine the best distance measure, linkage method, and standardization strategy for the data set in question. A table with cluster group assignments for each of the 38 analyses and 343 districts was created and joined to the district coverage's attribute table in ArcView so that results from the 38 analyses could be quickly viewed and compared. Table 2 summarizes the input parameters for the 38 analyses with comments on the results.

In the first 13 cases, different linkage methods and numbers of clusters were tried with the Euclidean distance measure with Ward's method giving the most appropriate distribution of land-use scenarios. Other methods generally sorted the districts into land-use scenario categories of mostly residential, mostly agricultural, mostly industrial, and so forth, without differentiating the distributions of the minor land uses in a district. Ward's method succeeded in partitioning many of these groups into separate scenario categories, particularly among the residential scenarios. The single linkage method was eliminated from consideration due to its tendency toward grouping the majority of the districts in one mega-cluster. With this method, most of the districts are grouped in one land-use scenario (scenario 1) as shown in figure 3 (after page 42). The median and centroid linkages were also judged to give poor enough results to eliminate the need for any further consideration of their use.

For cases 14 to 29, different distance measures were tried with the average, complete, McQuitty, and Ward linkage methods. For some reason, the Pearson and Squared Pearson measures produced a chaining effect similar to that of the single linkage for all but the Ward's linkage. For this reason, a standardization method other than the Pearson distance measure became necessary. The Manhattan and Squared Euclidean measures produced results comparable to the Euclidean.

For cases 30 to 33, Z-score standardization was applied to the variables before clustering for a few of the best combinations of linkage method and distance measure found thus far. With the complete and average linkages, the Z-score standardization also produced a chaining of observations. With Ward's method, the Z-score standardization proved adept in singling out extreme values of oddly distributed variables like mobile homes, apartments, and warehouses. However, doing this broadened the range of the more normally distributed variables like single-family residential and agricultural that could be included in the same scenario. Cases 34 and 35 were created with 30 clusters instead of 20 in an attempt to break up some of the more dissimilar clusters. Still, a few clusters exhibited odd groupings. As an example, the three districts Provo 3, Orem 2, and Utah County 3, were all grouped into the same cluster in case 34, largely based on their similarly high percentages of warehouse use (C3) (see table below).

District	R1	R2	R3	R4	R5	C1	C2	C3	C4	C5	AG	OS	VA
Provo 3	4.6	1.6	7.7	1.8	0.0	16.4	22.1	4.8	1.3	16.2	0.4	9.3	13.9
Orem 2	70.1	0.0	0.0	0.0	0.0	9.0	1.8	4.2	1.4	0.0	13.5	0.0	0.0
Utah Co. 3	8.8	0.0	0.0	0.6	0.0	0.0	0.0	4.1	0.0	0.0	86.5	0.0	0.0

TABLE 2 Input Parameters for 38 Trial Cluster Analyses on Land-Use Distress

Case	Distance	Linage	Standardization	No. of clusters	Comments
1	Euclidean	Single	None	20	Produced single large clusters of over 300 districts
2	Euclidean	Single	None	15	with no cluster larger than 5 districts; chaining
3	Euclidean	Single	None	25	effect took place
4	Euclidean	Average	None	20	Smaller main clusters and larger minor clusters
5	Euclidean	Average	None	30	but not much diversity in residential scenarios
6	Euclidean	Average	None	18	
7	Euclidean	Centroid	None	20	Relatively poor distribution of cluster sizes and
8	Euclidean	Centroid	None	30	poor diversity in residential scenarios
9	Euclidean	Complete	None	20	Somewhat more diversity in residential scenarios
10	Euclidean	Complete	None	17	
11	Euclidean	McQuitty	None	20	Little diversity in residential scenarios
12	Euclidean	Median	None	20	Large mega-cluster, like with single linkage
13	Euclidean	Ward	None	20	Relatively equal cluster sizes; good diversity
14	Pearson	Average	None	20	Pearson and Squared Pearson linkages created
15	Manhattan	Average	None	20	mega-clusters; Manhattan seemed almost as good as
16	Sq. Euclidean	Average	None	20	Euclidean; Squared Euclidean was nearly identical
17	Sq. Pearson	Average	None	20	to Euclidean
18	Pearson	Complete	None	20	Pearson and Squared Pearson linkages created
19	Manhattan	Complete	None	20	identical mega-clusters; Manhattan created slightly
20	Sq. Euclidean	Complete	None	20	more diversity than Euclidean; Squared Euclidean
21	Sq. Pearson	Complete	None	20	identical with Euclidean
22	Pearson	McQuitty	None	20	Changing the distance measure has similar effects
23	Manhattan	McQuitty	None	20	as with average and complete linkages
24	Sq. Euclidean	McQuitty	None	20	
25	Sq. Pearson	McQuitty	None	20	
26	Pearson	Ward	None	20	Pearson and Squared Pearson produced slightly less
27	Manhattan	Ward	None	20	diversity in residential and agricultural scenarios;
28	Sq. Euclidean	Ward	None	20	Squared Euclidean lumped agricultural scenarios
29	Sq. Pearson	Ward	None	20	together; Manhattan similar to Euclidean
30	Sq. Euclidian	Ward	Z-scores	20	Similar diversity to case without standardization (case 28) but oddly distributed variables better represented
31	Euclidean	Complete	Z-scores	20	Much poorer diversity than in case 9
32	Euclidean	Average	Z-scores	20	Forms mega-cluster; worse than case 4
33	Euclidean	Ward	Z-scores	20	More diverse in some areas than with case 13
34	Sq. Euclidean	Ward	Z-scores	30	Improved diversity over case 30
35	Euclidean	Ward	Z-scores	30	More diverse than case 33
36	Euclidean	Ward	None	30	Similar diversity to case 35 but oddly distributed variables like R3, R4 not as well represented
37	Euclidean	Ward	Scaled percentages	30	Oddly distributed variables well-represented but not enough of an improvement in variable bounds
38	Sq. Euclidean	Ward	Z-scores	35	Number of clusters increase to 35 to separate a few odd groupings

Clearly, Provo 3 is made up of largely commercial uses; Orem 2 is primarily single-family residential, and Utah County 3 is mostly agricultural. Ideally, the classification scheme chosen would allow for three different scenarios, distinct from those scenarios for residential, commercial, and agricultural that do not include warehouses, to represent these districts.

Case 36 used 30 clusters with no standardization applied to see if the increased number of clusters alone would provide for more distinct groupings without sacrificing representation of the more minor uses like mobile homes, apartments, and warehouses. The range of variables about the means decreased significantly, but most representation of extreme values for the minor variables was lost. For case 37, a compromise standardization procedure was tried before clustering. Instead of using Z-scores to standardize the variables, all variables were scaled such that the minimum value for the variable was 0.0% and the maximum value was 100%. Hence, a value of 5.8% for mobile homes (minimum 0% and maximum 12.1%) in West Valley City was scaled to 48.3% (5.8% / 12.1% × 100%). This method of standardization worked nearly as well as the Z-score method, but the improvement in within-cluster dissimilarity was small enough to make the effort unfruitful.

Consequently, the parameters chosen as best for cluster analysis on the land-use data were Ward's linkage method, the Squared Euclidean distance measure, and Z-score standardization of variables (case 38). Increasing the number of clusters to 35 broke up a few additional odd groupings. The cluster that included the three districts mentioned previously, for example, was broken into two clusters. It was decided that an additional reduction of dissimilarities would have required a greater number of clusters in the final analysis than was desirable; the final number of scenarios became 35.

Table 3 (on pages 10 and 11) lists the final 35 scenarios with cluster means. Since adding the mean percentages for each scenario did not always result in a total of 100% (sums varied from 92 to 100%), the mean values in table 3 represent the cluster means scaled such that their sums total 100% for each scenario. Figure 4 (after page 42) shows the district map coded for the results of the final cluster analysis with 35 scenarios. This figure shows which planning districts have similar land-use patterns. Table 3 and figure 4 show that the clustering by the Ward's linkage method performed well for the given data set.

CONCLUSIONS

This paper showed cluster analysis to be a viable tool for grouping planning districts into a smaller number of planning scenarios for regional land-use and transportation sketch planning. A comparison of figures 2 and 4 shows that cluster analysis reduces the amount of detail required to represent a wide variety of land-use scenarios but does so without significantly altering the big picture on a regional level. This finding is potentially significant to future land-use and transportation planning projects. Alternative land-use scenarios used in planning models need not be limited to a few hypothetical land-use cases. Rather, multiple scenarios can be generated as warranted to represent accurately all patterns of land use found in a given regional area. Scenarios currently not found in the study region, such as those taken from another region, may then be added to represent a fuller spectrum of possibilities.

Certain challenges must be met in order to apply cluster analysis to land-use data successfully. Chief among these is the proper selection of a distance measure, linkage method, and standardization procedure. The findings of this study do not eliminate the need to reiterate this process for data gathered from areas outside the Wasatch Front region of Utah. Additionally, measures must be taken to ensure that the input land-use data is current and accurate or the investigator risks magnifying the inaccuracies in the final set of derived scenarios. Uniformity among agencies in the categorized description of land use is indispensable in the gathering of timely and accurate data on regional land-use scenarios. Lastly, a good method for drawing district boundaries is needed to ensure that subjective concerns do not influence the process.

ACKNOWLEDGMENTS

This study was funded by the National Science Foundation (contract number CMS 9817690). The authors wish to thank Stuart Challender of the

State of Utah Automated Geographic Resource Center (AGRC), Mike Brown of the Wasatch Front Regional Council, Andrew Jackson and Andrew Wooley of the Mountain Land Association of Governments, and many other city employees and engineers who provided us with GIS-based data and other information essential for this study.

REFERENCES

Balling, R.J., J.T. Taber, M.R. Brown, and K. Day. 1999. Multiobjective Urban Planning Using Genetic Algorithm. *ASCE Journal of Urban Planning and Development* 125, no. 2:86–99.

Environmental Systems Research Institute (ESRI). 1999. *ArcView GIS 3.2 User Manual*. Redlands, CA.

MINITAB. 1999. *MINITAB Statistical Software Release 12 for Windows® User Manual*. State College, PA.

Smith, J. 2000. *Cluster Analysis as Part of a Land-Use Classification Scheme for Optimized Land Use and Transportation Planning*. Masters Thesis, Department of Civil Engineering, Brigham Young University, Provo, UT.

The State of Utah Automated Geographic Resource Center (AGRC). 1997. *Parcel-Level Land-Use Data*. Salt Lake City, UT.

Taber, J.T., R.J. Balling, M.R. Brown, K. Day, and G.A. Meyer. 1999. Optimizing Transportation Infrastructure Planning with a Multiobjective Genetic Algorithm Model. *Journal of the Transportation Research Board* 1685:51–6.

Transportation Research Board (TRB). 2000. Conference II Resource Paper: Summary of Survey Results. In *Refocusing Transportation Planning for the 21st Century,* Conference Proceedings 20:179–83.

Vuchic, V.R. 1999. *Transportation for Livable Cities*. Rutgers: Urban Policy Research Press.

TABLE 3 Land-Use Distribution in Percent of the Clustered Land-Use Scenarios for 343 Wasatch Front Region Planning Districts

Scenario	No. of districts	Mean land-use distributions, percent													Name of scenario
		R1	R2	R3	R4	R5	C1	C2	C3	C4	C5	AG	OS	VA	
1	42	13.1	0.1	0.1	0.1	0.0	1.1	1.2	0.0	0.0	2.5	78.3	0.7	2.8	Agriculture with some SF residential
2	35	43.2	0.1	0.2	0.4	0.0	3.7	3.9	0.1	0.1	4.5	37.8	0.3	5.7	SF residential and agriculture
3	34	89.7	0.8	0.5	0.1	0.0	3.6	0.5	0.0	0.0	0.5	2.5	0.1	1.6	SF residential
4	23	68.8	7.5	0.8	0.1	0.0	4.3	1.0	0.1	0.3	6.2	6.3	0.9	3.6	SF residential with 2-4 unit residential
5	16	69.5	0.3	0.9	0.0	0.0	18.5	2.4	0.3	1.0	2.6	2.0	0.1	2.3	SF residential with retail
6	14	67.0	1.4	5.3	0.1	0.1	8.0	0.2	0.0	0.3	9.0	3.9	0.5	4.3	SF residential with apartments and special purpose
7	20	62.9	0.5	0.6	0.1	0.0	4.7	1.5	0.0	0.1	8.6	6.5	8.9	5.6	SF residential with open space, special purpose, and agriculture
8	11	25.1	1.3	2.8	0.0	0.1	4.5	8.9	0.1	0.1	27.1	3.6	5.3	21.0	Special purpose and SF residential with vacant land and apartments
9	28	70.6	0.3	0.4	0.1	0.1	2.4	0.6	0.0	0.0	14.3	6.4	0.3	4.6	SF residential with special purpose
10	14	52.9	1.4	0.8	0.0	0.0	4.6	1.2	0.0	0.1	7.5	3.2	22.9	5.4	SF residential with open space
11	7	2.5	0.0	0.0	0.0	0.0	0.1	82.3	0.0	0.0	0.4	14.5	0.0	0.1	Industrial
12	11	69.2	1.7	0.6	2.4	0.0	2.1	0.4	0.0	0.5	5.1	13.7	0.2	4.2	SF residential with agriculture and mobile homes
13	10	19.4	1.9	1.4	0.2	0.0	25.2	25.2	0.0	0.0	4.3	0.6	3.1	18.7	Industrial, retail, SF residential, and vacant land
14	9	42.0	6.1	9.2	0.1	0.1	18.6	1.0	0.1	0.4	15.3	0.0	0.7	6.5	SF, 2-4 unit and apartment residential with retail and special purpose
15	7	5.4	0.7	0.4	0.2	0.0	7.7	28.1	0.4	0.1	4.7	5.9	1.0	45.6	Vacant land with industrial
16	4	8.2	0.3	0.9	0.3	0.0	2.4	0.5	0.0	0.0	79.5	1.7	1.8	4.3	Special purpose with some SF residential
17	2	0.0	0.0	0.0	0.0	0.0	5.9	10.3	0.0	0.0	3.5	4.0	37.0	39.3	Vacant land and open space with industrial
18	6	4.9	0.0	0.0	0.0	0.0	0.9	2.6	0.1	0.0	1.0	14.3	1.8	74.5	Vacant land with some agriculture
19	1	1.2	0.2	2.9	0.0	0.0	0.2	0.0	0.0	29.8	55.8	0.0	9.9	0.1	Special purpose and office
20	1	34.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	64.6	0.9	Open space with residential

TABLE 3 Land-Use Distribution in Percent of the Clustered Land-Use Scenarios for 343 Wasatch Front Region Planning Districts (*continued*)

Scenario	No. of districts	Mean land-use distributions, percent R1	R2	R3	R4	R5	C1	C2	C3	C4	C5	AG	OS	VA	Name of scenario
21	3	10.8	0.0	1.1	0.6	0.0	65.6	5.0	0.0	0.0	2.9	4.9	0.0	9.1	Retail with some SF residential and vacant land
22	3	20.4	0.3	0.2	1.2	0.0	21.4	5.0	0.0	2.7	6.0	2.9	1.1	38.7	Vacant land with retail and SF residential
23	5	38.7	1.9	28.5	0.5	0.1	18.1	1.9	0.4	0.1	0.0	1.4	0.1	8.1	Apartments with SF residential and retail
24	7	40.8	0.8	4.0	4.0	0.0	16.5	8.9	0.8	0.4	9.2	8.3	1.4	4.9	SF residential with retail, industrial, special purpose, and agriculture
25	2	13.2	0.0	0.0	0.0	0.0	1.5	3.2	12.7	0.1	6.0	35.3	0.0	28.0	Agriculture and vacant land with warehouses
26	4	35.4	3.3	3.0	10.5	0.0	4.6	4.0	0.1	0.9	20.1	10.9	0.6	6.5	SF residential and mobile homes with special purpose
27	6	56.7	1.8	0.0	0.9	0.0	3.5	4.8	4.2	0.5	0.6	26.2	0.0	0.8	SF residential with agriculture and some warehouses
28	3	37.7	1.4	1.6	4.0	0.3	1.8	0.2	0.1	0.0	11.4	5.2	1.0	35.3	SF residential and vacant land with special purpose and mobile homes
29	2	16.6	1.8	0.2	0.0	0.0	33.8	21.2	10.3	2.6	6.2	1.0	2.5	3.7	Retail, industrial, and warehouses with some SF residential
30	2	41.7	4.5	1.7	0.0	0.1	6.6	0.2	0.4	8.1	14.0	2.5	14.2	5.9	SF residential with office and some special purpose and open space
31	5	16.2	2.6	1.7	0.9	0.0	17.3	29.5	3.8	0.7	9.8	9.5	3.4	4.6	Industrial with retail, SF residential, and some warehouse
32	3	39.1	18.8	9.5	1.7	0.0	14.5	3.2	0.5	0.6	3.0	4.0	1.2	4.0	SF and 2-4 unit residential and apartments with some retail
33	1	47.3	0.7	0.0	0.0	13.9	0.0	0.0	0.0	0.0	0.0	36.2	0.0	1.9	SF residential and group quarters with agriculture
34	1	13.9	2.2	21.8	0.0	0.0	0.7	0.1	0.0	0.0	53.9	0.4	3.8	3.2	Special purpose and apartments with some SF residential
35	1	41.7	46.4	0.0	0.0	0.0	2.8	0.0	0.3	0.2	0.1	8.6	0.0	0.0	2-4 unit residential and SF residential

SF = Single family

Three Faces of Eve: How Engineers, Economists, and Planners Variously View Congestion Control, Demand Management, and Mobility Enhancement Strategies

ERIK FERGUSON
ETF Associates

ABSTRACT

The political acceptability (*A*) of public policy measures correlates positively with program effectiveness (*E*) and negatively with program cost (*C*) and other obstacles to implementation (*I*) under normal circumstances. Ferguson (1991) observed that the political acceptability of many demand management strategies seemed to correlate negatively with implied program effectiveness. Engineers, economists, and planners each have their own unique professional standards. Increased effectiveness is the primary goal of engineering. Improved efficiency is the generally accepted standard in economics. Process issues are of vital concern in planning. A review of the literature indicates few studies that rate demand management strategies in terms of all four variables of interest (*A, E, C,* and *I*) simultaneously. Three relevant studies were identified: one each by an engineer, an economist and a planner. Raw data, regression results, bivariate correlations, and model output reveal that two of the three studies support the Ferguson hypothesis. The other supports a more traditional public policy model. *E* is the most influential variable in the engineer's data. *C* is the most influential variable in the economist's data, while *I* is the most influential variable in the planner's data. These revealing results suggest the

Erik Ferguson, ETF Associates, P.O. Box 888729, Dunwoody, GA 30356. Email: etfassoc@bellsouth.net.

subtle manner in which professional training and experience may alter perceptions of transportation policies and programs in professional practice.

INTRODUCTION

Transportation planning is a capital-intensive process. The Federal-Aid Highway Act of 1956 removed gasoline tax revenues from the federal budget and set up a dedicated Highway Trust Fund. The Act further specified that all local return monies from this source must be allocated to the construction of new highway facilities (Weiner 1999). Project evaluation of highway capital investments was relatively simple in those days. Capital costs were fixed and low, at least by later standards. Operating costs were negligible from the government's perspective. Benefits included reduced out-of-pocket travel costs and time. Reduced costs were sufficient to justify most proposed projects, making concerns about the imputed value of travel time irrelevant. In the 1960s, rising land acquisition, highway construction costs, and a diminished potential to shorten travel distances made saving time more important in justifying new highway investments (Walters 1961).

In the 1970s, transportation system management (TSM) came into vogue. Even with savings in travel time properly accounted for, new highway construction became more expensive and difficult to justify, particularly in the face of local citizen opposition. The purpose of TSM is to soften the blow of transportation planning by making it more short-range, user-friendly, and demand-oriented. TSM actions include improved vehicular flow, preferential treatments for high occupancy vehicles, reduced peak period travel, parking management, promotion of alternative modes of transportation, and transit and paratransit service improvements (UMTA 1977). TSM's most ardent admirers predicted its possible demise from political opposition, the projected outcome of institutional inertia, and professional apathy (Gakenheimer and Meyer 1979). However, many TSM strategies have done much better than expected.

In the 1980s, travel demand management (TDM) became the watchword of the day. TDM is the demand side of TSM, making it more suitable for private sector participation. TDM operates even closer to end users such as individuals, households, and firms. TDM strategies include alternative modes and hours of travel, alternative locations for specific activities, as well as economic incentives and institutional arrangements that may be required (Orski 1987).

Due to the wide range of strategies available to deal with traffic problems, the complexity of decisions associated with congestion management has increased in recent years. Sorting through the bewildering array of alternative policies, programs, and projects can be a daunting task. Word of mouth, hearsay evidence, and the occasional case study can only go so far. The more comprehensive the evaluative outlook, the better the implied advisement should be. Comprehensive comparative assessments of traditional planning, TSM, and TDM techniques, however, remain relatively few and far between (Ferguson 2000).

PREVIOUS STUDIES

One study by the Urban Mass Transit Association (UMTA) (1977) categorizes TSM measures without identifying any of the potential impacts. Wagner and Gilbert (1978) evaluates TSM measures in terms of effectiveness and cost but omits any discussion of implementation issues. Misch and Margolin (1981) categorizes the "feasibility or prospects" of TSM actions taken in support of ridesharing in a limited fashion. Schonfeld and Chadda (1985) focuses on the effectiveness of travel reduction options with particular emphasis on parking management strategies. Levinson et al. (1987) evaluates the effectiveness of TSM strategies in terms of connected "impact chains." Bhatt and Higgins (1989) focuses mainly on results, measured in terms of mode split (table 1).

Two studies by the Institute of Transportation Engineers (ITE) (1989 and 1997) present information on impacts, costs, and obstacles to implementation associated with a wide range of congestion control and mobility enhancement tools but discuss political acceptability only tangentially. Ferguson (1991) observes that the public acceptability of many TDM strategies correlates negatively with their implied effectiveness. Road and parking pricing are more effective but less popular strategies. However, voluntary efforts to promote alternative

TABLE 1 Performance Measures

Study	Measure(s) of				Strategies
	Effectiveness	Cost	Implementation	Acceptability	
Wagner and Gilbert (1978)	7 applications + 2 impacts	2 (capital + operating)			13 TSM strategies
Misch and Margolin (1981)	2 (effect + compatibility)	1 (cost per VMT or VHT reduced)	1 (feasibility or prospects)		20 TSM actions
Schonfeld and Chadda (1985)	4 effects + 1 application	2 (user + general)	1 (problems and requirements)		50 travel reduction options
Levinson et al. (1987)	17 (goals or impacts)				10 TSM strategies
Bhatt and Higgins (1989)	1 (results)				25 TSM programs
ITE (1989)	1 (impacts)	1 (costs)	1 (implementation)		55 congestion management tools
Ferguson (1991)	0 (implied)			1 (public)	4 TDM strategy groups
Downs (1992)	2 (extent + impact)	2 (commuter + social)	2 (institution + administration)	1 (political)	23 congestion-reducing policies
Zupan (1992)	1 application + 4 impacts	3 (employee + employer + public)	1 (index)	4 (employee + employer + municipal + public)	22 TDM solutions
Arnold (1993)	1 (scalar)	1 (scalar)	1 (scalar)	1 (number of agencies)	53 congestion-reducing measures
Comsis (1993)	1 (vehicle trip reduction)				22 TDM programs
OECD (1994)	8 applications + 8 impacts				35 congestion control strategies
McBryan et al. (1996)	2 applications + 2 impacts	2 (incremental cost + who pays)	4 (enabling + implementing authority + time + difficulties)		33 TDM strategies
ITE (1997)	1 (benefits/costs)		1 (implementation)		82 mobility enhancement tools
Dueker et al. (1998)	3 applications + 2 impacts		1 (administration)	1 (political feasibility)	10 parking management strategies
Booz-Allen & Hamilton (2000)	9 direct + 12 indirect effects		1 (practical feasibility)	1 (political)	32 TDM strategies
Totals	**91**	**15**	**14**	**11**	**489**

TDM = travel demand management
TSM = transportation system management
VHT = vehicle hours of travel
VMT = vehicle-miles traveled

modes and hours of travel are more popular but less effective.

Downs (1992), Zupan (1992), and Arnold (1993) evaluate various mixes of congestion control and demand management strategies in terms of four generalized performance measures: effectiveness, cost, ease (or difficulty) of implementation, and political acceptability. Dueker et al. (1998) perform a similar analysis on a smaller number of parking management strategies. Comsis (1993) and the Organisation for Economic Cooperation and Development (OECD) (1994) revert to an earlier emphasis on program effectiveness. McBryan et al. (1996) balances effectiveness, cost, and implementation issues but deliberately excludes political acceptability as "too arbitrary" (Shadoff 1997).

The remainder of this paper focuses on an analysis and evaluation of data derived from Arnold (1993), Downs (1992), and Zupan (1992), hereinafter referenced to as such.

HYPOTHESES

There are two major hypotheses tested here, one rather precise in nature, the other far less so.

1. Do Arnold, Downs, and Zupan confirm or deny Ferguson's (1991) assertion that TDM program effectiveness and political acceptability are negatively rather than positively correlated? An objective comparison of the signs and significance of relevant parameters determines the outcome of this test.
2. Are there any other differences in data or results among the three studies examined? How do such differences relate to the professional outlook of the respective authors? This is a much more subjective test but intriguing nonetheless.

AUTHORS

In order to evaluate the possible contribution of professional perspectives to the understanding of complex policy issues such as congestion control or demand management, it is important to know with whom one is dealing. The three studies in question were each authored by a single individual (table 2).

Eugene Arnold is a Senior Research Scientist at the Virginia Transportation Research Council, Virginia Department of Transportation (VDOT) in Charlottesville, Virginia. The Commonwealth of Virginia is one of the most conservative states in the Union. VDOT ranks among the more innovative state transportation agencies, thanks in no small part to its proximity to the nation's capital. Arnold is an active member and national leader in the Institute of Transportation Engineers (ITE). For example, he chaired the committee that prepared the most recent update of *Trip Generation* (ITE 1997). It should come as no surprise to find that Arnold's structure of the research problem closely parallels that identified in ITE (1989).

Anthony Downs is a Senior Fellow in Economic Studies at the Brookings Institution in Washington, DC. His expertise extends to topics as diverse as democracy, demographics, housing, metropolitan

TABLE 2 Authors and Data

	Author		
Attribute	Arnold	Downs	Zupan
Professional orientation	Engineering	Economics	Planning
Institutional affiliation	Virginia DOT	Brookings Institution	Regional Plan Association
Expertise	Vehicle trip generation	The costs of sprawl	Public transit and land use
Variable components	4	7	12
Latent variables	4	4	4
Observations	53	23	22
Data type	Quantitative	Qualitative	Mixed
Data source	Mail survey	Author	Author

policy, real estate, real estate finance, smart growth, suburban sprawl, and urban policy. This is his first foray into transportation planning in more than a quarter of a century. Downs (1992) expands on the ideas originally set forth in Downs (1962).

Jeffrey Zupan is Senior Fellow for Transportation at the Regional Plan Association in New York City. His research on the relationship between land use and public transportation is well known and respected. Pushkarev and Zupan (1977) showed 1) what type of transit works best, 2) where it works best, 3) which support policies are most effective, and 4) how to estimate transit demand and costs (see Lee 1978). The connection between Zupan (1992) and Pushkarev and Zupan (1977) is clear.

DATA

The data used in this analysis are widely available and based on previously published research results. They are reproduced here mainly for the reader's convenience (see tables A-1 to A-3). However, Arnold's original survey data were obtained directly from Arnold for this study. Table A-1 includes more information than was published in the original paper.

Arnold

Arnold surveyed 85 local, regional, and state transportation agencies in Virginia regarding 53 congestion-reducing measures. The ratings and rankings shown in table A-1 reflect the collective judgment of traffic engineers and transportation planners in Virginia, not necessarily Arnold's personal opinions.

Arnold treats the ordinal scales he uses to measure effectiveness, cost, and implementation as interval scales in his analysis: a testable proposition. Arnold's data are purely quantitative in presentation but largely qualitative in nature. This may reflect a preference among engineers for data and methods that are more objective. Arnold's data include 53 observations and only 4 variables, a case of problem overidentification. This is positive, of course. Engineers prefer larger margins of error.

Arnold's principal research findings include the following.

1. TSM measures have almost caught up with supply-side measures in terms of actual use. Supply-side measures are more effective but extremely costly and face many obstacles to implementation.
2. TDM measures are much less effective and slightly more costly than TSM measures. With obstacles to implementation about the same, it is no surprise that TSM is much more popular than TDM in the Commonwealth of Virginia.

Downs

Downs evaluates 23 congestion-reducing policies using 7 items corresponding to 4 variables (table A-2). Whereas the Arnold data appear purely quantitative, the Downs data appear purely qualitative in nature. Downs uses different semantic scales to describe most of his items. The two cost items share an identical scale.

All of the scales constructed by Downs are ordinal in nature, with the possible exception of the institution required for implementation, which may be categorical. The Downs data are overidentified but not nearly as much as the Arnold data. Downs assigns a non-ambiguous descriptive adjective to each attribute for every policy he considers.

Downs' principal research findings include the following.

1. Supply-side policies are more expensive but require less institutional change than demand-side policies.
2. Demand-side policies have broader effects but are less politically acceptable than supply-side policies.

Zupan

Zupan evaluates 22 TDM solutions using 12 items corresponding to 4 variables (table A-3). Zupan is the everyman of performance measurement. He incorporates a little bit of everything in his evaluation matrix, including both words and numbers, scales both absolute and relative, as well as suitable descriptors of measurement variability, non-applicability, and the unknown.

Zupan's 12 items are just barely identified by the 22 observations in his matrix, providing another indication of his tolerance for uncertainty. There is not one item in Zupan's matrix that is ranked consistently using a single semantic scale. The only possible exception is ease of implementation, which still includes two "unknowns" and one "not applicable" rating. Planners may be less technical than engineers and less consistent than economists, but one thing stands clear: they are much more comfortable with uncertainty.

Zupan's principal research findings include the following.

1. Strategies that reduce traffic congestion, have no negative effect on transit, and are politically popular should be pursued.
2. Strategies that have little effect on traffic congestion, negative effects on transit, or are politically unpopular should be avoided.
3. Strategies for which too little is currently known should be studied more carefully.

Comparing the methodological results of these three studies would be intriguing but difficult to accomplish. Observed differences between the three databases seem much greater on the surface than any prospective similarities might be. The problem of compatibility must be resolved in order to make any more meaningful comparisons between these studies and their results.

Semantic Scales

The next objective is to evaluate a model of the following general form:

$A = f(E, C, I)$

where

A = (political) acceptability
E = (program) effectiveness
C = (program) costs
I = (obstacles to) implementation

It is hypothesized that (political) acceptability should be

1. Either positively or negatively correlated with program effectiveness, depending on whether a traditional public policy model applies, or Ferguson's 1991 conjecture is correct.
2. Negatively correlated with program costs and obstacles to implementation.

In order to test these hypotheses, it is necessary to convert the data found in tables A-1 to A-3 to standard form. Certain assumptions are required; the most important ones follow.

1. Observed semantic scales correspond to unobserved, logically consistent, and mathematically comparable probability surfaces that can be mapped.
2. Observed variable components correspond to unobserved or latent variables that may be combined.

Given these assumptions, composite variables representing *A, E, C,* and *I* may be computed and subsequently analyzed using regression. The initial assignment of values to each semantic scale in the data is straightforward (table 3).

1. Arnold's use of semantics requires little comment. He assigned values of 0, 1, and 2 to his 3 ordinal scales. This has been changed to 1, 2, and 3 for consistency with the Downs and Zupan data but otherwise should have no effect.
2. Downs employs a wide variety of different labels. By inspection, all of his labels correspond to a single four-point scale of zero to three, with relatively minor variations as noted.
3. Zupan employs a bewildering array of positive and difference scales, often simultaneously. He occasionally uses different labels that mean the same thing (for example, he uses "same" and "none" on a difference scale). He uses three different labels to represent aspects of uncertainty ("varies," "unknown," and "not applicable"). Despite these obstacles, it remains possible to assign plausible initial values to all of his labeling conventions.

REGRESSION RESULTS

With an initial assignment of semantic values in hand, it is possible to calculate regression equations for each dataset. This does not imply that the results will maximize the potential of any available information, of course. For that to happen, several more questions must be answered.

1. Semantic values: the initial assignment of semantic values is unambiguous in the case of Arnold and Downs. The accuracy of the semantic values assigned to Zupan's mixed scales and uncertainty measures is open to question.

TABLE 3 Semantic Scales

Author	Variable	Component	Numeric values (initial assignment) –2	–1	0	1	2	3	4	5
Arnold	Effectiveness					Minimal	Average	Maximum		
	Cost					Inexpensive	Average	Expensive		
	Implementation					Easy	Average	Difficult		
	Acceptability				No	Yes				
Downs	Effectiveness	Extent				Narrow	Variable	Broad		
		Impact				Little	Moderate	Great		
	Cost	Direct to commuters To all of society			None	Minor	Moderate	Great		
	Implementation	Required institution			None	Cooperative	Regional			
		Ease of administration				Easy	Moderate	Hard		
	Acceptability					Poor	Moderate	Good		
Zupan	Effectiveness	SOV reduction Peak trip reduction VMT reduction Transit impact	Highly negative	Negative	None / varies / unknown	Low Positive	Medium	High		
	Cost	Employee Employer Public capital		Lower	Same / none / varies / unknown / not applicable	Higher				
	Implementation				Unknown / not applicable	Difficult <—	—	—	—	—> Easy
	Acceptability	Employee Employer Municipal Public		Negative	Varies / unknown / not applicable	Low Positive	Medium	High		

SOV = single-occupancy vehicle
VMT = vehicle-miles traveled

2. Semantic differences: Arnold and Downs employ a series of ordinal scales. Semantic differences between values on an ordinal scale may be linear or non-linear: a testable proposition. It is difficult, if not impossible, to evaluate semantic differences where the semantic values themselves are open to question, leaving Zupan out of this discussion.
3. Component weights: where more than one item describes a variable, these must be combined to provide a unitary performance measure. For simplicity, it is assumed here that each item bears the same weight: a testable proposition.
4. Variable transformations: none of the variables used in this analysis, with the possible exception of Arnold's measure of political acceptability, is anything more or less than a dimensionless performance indicator. All variables were standardized under an assumption of normality in recognition of this fact and to facilitate comparisons. The limited sample sizes involved do not allow for any other assumption.

Arnold

The Arnold data generally support the traditional public policy model (table B-1):

1. Arnold's measure of political acceptability is positively correlated with program effectiveness and negatively correlated with program cost and obstacles to implementation. The cost variable is not statistically significant in any of the equations although it does approach significance as the model's assumptions are relaxed.
2. Supply-side measures are most politically acceptable, followed by TSM and growth control, with TDM last. The only statistically significant difference is between supply-side and TDM measures.
3. Semantic differences between values on Arnold's ordinal scales are clearly nonlinear.
4. The extent of nonlinearity varies. Implementation is least, and cost is most nonlinear. Low and medium costs are not semantically distinguishable in Arnold's data.

Downs

The Downs data generally support the Ferguson hypothesis (table B-2).

1. Political acceptability is negatively correlated with program effectiveness, cost, and obstacles to implementation. There is a statistically significant bias toward supply-side policies.
2. Most of the semantic differences in the Downs data are highly correlated with their respective component weights. Given the limited sample size, it is not possible to test simultaneously for both semantic differences and component weights under these conditions. The sole exception to this rule is the difference between no cost and minor costs, which seems to operate somewhat more independently.
3. Component weights are of somewhat greater interest than semantic differences, at least from a public policy perspective. Component weights that maximize overall model goodness-of-fit include the following:
   a. Effectiveness: impact is more important than extent, but the difference is fairly slight.
   b. Cost: social costs are twice as important as commuting costs.
   c. Implementation: required institution (which operates efficiently as a linear variable) weighs in at an order of magnitude greater than ease of administration.
4. Simultaneous estimation of low-cost semantic differences and variable component weights produce results roughly equivalent to those found separately, indicating a stable solution to the model.

Zupan

The Zupan data generally support the Ferguson hypothesis (table B-3).

1. Political acceptability is negatively correlated with program effectiveness and obstacles to implementation and is, paradoxically, positively correlated with program cost.
2. "Varies" is an unknown quantity, normally assigned to the midpoint of any given range of values. Zupan uses two distinct scales to describe key elements of program effectiveness and political acceptability, making it unclear

where "varies" belongs on a uniform or combined scale. The regression results clearly support the notion that "varies" refers to a positive (low-medium-high) rather than a difference (negative-none-positive) scale in this analysis.

3. Two other semantic-value assumptions bear additional scrutiny.
   a. "Unknown" and "not applicable" belong at the midpoint of their respective ranges, rather than at one or the other of the two extremes; this is true in all cases.
   b. The difference and positive scales are separate and distinct for program effectiveness but are semantically indistinguishable in the case of political acceptability. This is an unusual finding to say the least.
4. Best fitting component weights include the following:
   a. Transit impact is the most important component of program effectiveness, followed by vehicle-miles traveled (VMT) reduction and peak trip reduction. Single-occupancy-vehicle (SOV) reduction has no weight.
   b. Employer cost is the most important component of program cost, followed by public capital cost. Employee cost has no weight.
   c. Municipal acceptance is the most important component of political acceptability, followed by public acceptance and employee acceptance. Employer acceptance has no weight.
   d. The unexpected positive correlation between program cost and political acceptability may be explained largely as a function of component weights. The relationship is significant only in the last equation, where employer costs dominate the cost variable, while employer acceptance is omitted from the political acceptability variable. Economists refer to this as a transfer payment (Small 1999). Others might call it a tax or even an unfunded mandate.

Dueker et al. (1998) evaluate 10 parking management strategies in terms of 3 scopes or applications (temporal, functional, and spatial), 2 benefits or impacts (effectiveness and efficiency), ease of administration and political feasibility. A simple regression reveals that political feasibility is negatively correlated with effectiveness, lending support to the Ferguson hypothesis. No measure of cost was provided, and ease of administration was not correlated with political feasibility among these data.

Booz-Allen & Hamilton (2000) evaluate 32 TDM strategies loosely based on OECD (1994) in terms of 9 direct travel effects, 12 indirect policy effects or implications, practical feasibility (implementation), and political acceptability. A simple regression reveals that obstacles to implementation are strongly negatively associated, direct travel effects marginally negatively associated, and indirect effects not associated with political acceptability, lending support to the Ferguson hypothesis.

It would seem that the Ferguson hypothesis holds sway over congestion control and demand management strategies in many instances. Program effectiveness is negatively associated with political acceptability according to data derived from four out of five independent studies (table 4). This is a most unfortunate result, at least from a public policy perspective.

VARIABLE CORRELATIONS

For those who prefer simpler explanations, bivariate correlations based on naïve assumptions are shown in table 5.

These correlations illustrate that

1. Arnold's independent variables are highly correlated. This is an unanticipated response to the original survey design. Although not very desirable, these associations do not seem to have affected model estimation in any negative way. In fact, the estimated equations in table B-1 are, if anything, superior to the direct correlations shown in table 5. Program effectiveness exhibits the highest direct correlation with political acceptability in Arnold's data.
2. Downs' independent variables are largely independent of one another. Program effectiveness and cost are marginally correlated with each other, in a positive direction. Program cost exhibits the highest direct correlation with political acceptability in Downs' data.
3. Zupan's variables are more like Downs' than Arnold's. Program effectiveness and cost are marginally correlated, in a negative direction.

TABLE 4 Regression Results

Author	Effectiveness	Cost	Implementation	d.f.	R^2 (percent)
	Independent variables				
Arnold	+0.51*	–0.28***	–0.57*	49	39.3
Downs	–0.32**	–0.62*	–0.35**	19	75.4
Zupan	–0.45*	+0.39*	–0.56*	18	78.3
Dueker et al	–0.84*	NA	–0.04	9	67.8
Booz-Allen Hamilton	–0.22***	NA	–0.70*	28	52.2

* significant at 0.01 level
** significant at 0.05 level
*** 1.00 < t < 2.00

d.f. = degrees of freedom
NA = not applicable
U = data are unavailable

TABLE 5 Variable Correlations

Author	Variable	Cost	Implementation	Acceptability
		Variable		
Arnold	Effectiveness	4.46	4.37	2.56
	Cost		6.12	–0.28
	Implementation			–1.54
Downs	Effectiveness	1.42	–0.13	–2.2
	Cost		0.68	–4.02
	Implementation			–2.27
Zupan	Effectiveness	–1.54	0.18	–2.08
	Cost		0.49	1.72
	Implementation			–2.59

Note: All correlations are expressed as bivariate regression tests of "t" to facilitate comparison.

This is yet another odd finding. Obstacles to implementation exhibit the highest direct correlation with political acceptability in Zupan's data.

Engineers seek technical solutions to the problems they face. Economists search for low-cost solutions. Planners search for institutional answers (Mandelbaum 1996; Marshall 1997). The independent variable most highly correlated with political acceptability (A) in this analysis is

1. Program effectiveness (E)—Arnold, the engineer
2. Program cost (C)—Downs, the economist
3. Obstacles to implementation (I)—Zupan, the planner

These simple bivariate correlations would seem to support the idea that political acceptability is in fact defined to some limited extent with particular professional perspectives kept in mind.

OUTLIERS

One advantage of the equations shown in tables B-1 to B-3 is that values can be calculated for each of the four variables (A, E, C, and I) and compared across each data set.

Arnold

Arnold's model output (table C-1) does not vary much from the original data (table A-1). The values in table C-1 are directly comparable with those in tables C-2 and C-3, however, while those in table A-1 are not. Residuals are calculated using the following formula:

$$SRSD = \left(A_P - A_O\right) / \sigma_A \tag{1}$$

where

SRSD = studentized residual
A_P = political acceptability (predicted)
A_O = political acceptability (observed)
σ_A = standard error of A_P

A negative *SRSD* implies a positive bias in favor of that measure. In Arnold's case, survey respondents reported more examples of new and reconstructed highways in Virginia than the model predicts. Toll roads and high-occupancy vehicles (HOV) lanes produced fewer examples than the model predicts, serving as a confirmation of the model.

TSM measures show greater variability in Arnold's data.

1. Intersection improvements, restricted on-street parking, and five other traffic flow enhancements are favored.
2. Intelligent transportation system improvements (still relatively new in 1992) are not favored. This may have been a short-term effect, since remedied during the intervening time period.

Paratransit services and growth management round out the list of strategies that apparently receive a limited form of preferential treatment in Virginia (Cervero 1997).

Downs

Downs finds that demand-side policies are slightly more effective than supply-side policies but more difficult to implement (table C-2). This opposes Arnold's findings.

Downs' data show the following supply-side biases.

1. For: incident management and HOV construction
2. Against: new roads and transit service improvements

Down's data show the following demand-side biases.

1. For: growth management
2. Against: commuting allowances, staggered hours, and automobile license fees

Downs is the principal author of a widely read Real Estate Research Corporation report on the costs of sprawl (Altshuler 1977). Downs presumably should favor growth management beyond any limited ability it might have to deal with traffic congestion as an urban problem, based on this previous work.

Zupan

Zupan's data exhibit the following model biases (table C-3).

1. For: tax deductible transit vouchers, transit-friendly design, carpools, staggered shifts, and preferential parking
2. Against: trip reduction ordinances, employer subsidized transit, parking pricing, and growth management

None of the biases reported here are particularly consistent, other than that growth management shows up as an outlier in all three models. This identifies growth management as an issue of greater than average political controversy. Few of these biases are statistically significant, and none is influential enough to bias parameter estimates.

In summary (table 6):

1. Arnold reveals a slight bias in favor of certain types of highway projects.
2. Zupan reveals a slight bias in favor of certain types of transit projects.
3. Most of the residuals in Downs are random artifacts of model construction.

Other comparisons are best left to readers to explore on their own. Consider the relatively straightforward issue of parking pricing.

1. Arnold has something he calls differential parking rates, a limited variation on the general theme of parking pricing.
2. Downs provides two examples, a parking tax on peak-hour arrivals and the elimination of incentives to provide free employee parking. Both of these are indirect forms of parking pricing.

TABLE 6 Outliers

Author	Strategies	Tactics	Studentized residual
Arnold	TSM (11 of 20)	Major intersection improvements	–2.02
		Removing/restricting on–street parking	–1.86
		Traffic management during reconstruction	–1.49
		Turn prohibitions	–1.1
		Minor intersection improvements	–1.09
		One–way streets	–1.05
		Integrated freeway/arterial surveillance/control system	1.03
		Reversible traffic lanes on arterials	1.29
		Traffic surveillance/control system	1.44
		Traffic management team	1.66
		Motorist information system	1.67
	Supply (4 of 7)	Reconstructing highways with improved design	–1.64
		Constructing new highways	–1.14
		Constructing HOV lanes	1.35
		Toll–based financing to expedite new facilities	1.40
	TDM (1 of 21)	Implementing/improving paratransit services	–1.34
	Growth (1 of 5)	Growth management	–1.85
Downs	Supply (4 of 8)	Rapidly removing accidents	–1.55
		Building added HOV lanes	–1.41
		Building new roads without HOV lanes	1.11
		Improved transit service, amenities	1.19
	Demand (4 of 15)	Adopting local growth limits	–1.61
		Increasing automobile license fees	1.04
		Staggered working hours	1.29
		Tax-deductable commuting allowances	1.92
Zupan	TDM (7 of 22)	Transitchek (tax-deductable vouchers)	–1.81
		Transit–friendly design	–1.26
		Carpools	–1.22
		Staggered shifts	–1.22
		Preferential parking	–1.06
		Employer-subsidized transit	1.06
		Trip reduction ordinances	1.88

HOV = high-occupancy vechicle
TDM = travel demand management
TSM = transportation system management

Note: studentized residuals may be interpreted as tests of "*t*."

3. Zupan lists two aspects of parking pricing, including parking pricing and parking ratios. Parking ratios affect parking pricing indirectly, similar to federal tax incentives.

These three authors do not necessarily agree on the individual ratings of parking-pricing approaches any more than they do on how parking pricing should be labeled as a congestion control or demand management strategy.

CONCLUSIONS

How are program effectiveness, cost, obstacles to implementation, and political acceptability defined?

1. Effectiveness: effectiveness may be defined in terms of applications or impacts and is usually represented as some combination of both. The overwhelming emphasis in the literature on effectiveness may be overdone, however,

especially if effectiveness is negatively correlated with political acceptability.

2. Cost: cost seems to be a somewhat neglected issue, at least in relation to effectiveness. Given that TSM and TDM programs often are intended to serve as low-cost solutions, this may be less of an impediment than it seems. Equity impacts may be more important than efficiency concerns in political terms and deserve greater scrutiny as a result.
3. Implementation: obstacles to implementation remain nebulous, at least in terms of definition. More work certainly could be done to identify the nature of and means to overcome those obstacles to existing implementation.
4. Acceptability: political acceptability is, if anything, the least understood aspect of the congestion control, demand management, and mobility enhancement problem.

How do professional perspectives influence these definitions?

1. Engineers: the dominance of engineering in the debate over congestion control, demand management, and mobility enhancement is clear in the definition of the problem to be solved. The emphasis on effectiveness to the practical exclusion of all else is a good indication of the engineering mindset at work. Transportation engineers must look beyond their own interests to see other aspects of the problem.
2. Economists: low-cost solutions to traffic congestion and air pollution may reduce the urgent necessity to think in economic terms, except for the considerable efficiency and equity impacts that might otherwise accrue (Garrett and Wachs 1996; Howitt and Altshuler 1999). Cost estimation seems undervalued in the current literature, with the identification of distributional effects even less in evidence. There is clearly work to be done by transportation economists in this area.
3. Planners: implementation issues are discussed slightly more often than cost issues in the literature, but the emphasis remains more anecdotal than evidentiary in terms of presentation. A practical definition of implementation issues and how these may be categorized from a strictly analytical point of view remains to be done.

How well is the decisionmaking process actually understood?

1. Policymakers: most of the transportation decisions are made by bureaucrats or politicians—hence the frequent confusion between "bureaucratic" implementation issues and "political" acceptability. A clearer separation of roles and responsibilities might help eliminate this problem, at least partially.
2. Stakeholders: policymakers are influenced by a variety of considerations in addition to their own personal opinions and moral judgments. Developers, employers, neighborhood associations, environmentalists, and a wide variety of other advocacy groups are involved in decisions that affect local communities. The perspectives of each of these groups may color perceptions of political acceptability and ultimately influence the outcome of transportation policy decisions.

Arnold (1993) is an excellent example of how to go about evaluating the political acceptability of transportation policies, programs, or projects. Downs (1992), Zupan (1992), and McBryan et al. (1996) provide useful guidance on additional aspects of program cost, obstacles to implementation, and political acceptability. Many others have contributed to the identification of congestion management strategies and measures of their effectiveness.

Confirmation of the Ferguson (1991) hypothesis leaves one important question unanswered. If the traditional public policy model holds true under most normal circumstances, either the model itself or one of its variable components must be in error in the present case. The model itself is too straightforward to be associated with any kind of specification error. The acceptability variable operates much as expected in relationship to the other variables in the model, and is therefore free from further suspicion as well. The cost variable is moderately suspect, due to its omission in some cases and its unusual performance in yet another.

Effectiveness behaves more like a cost than a benefit in most of the models tested here. One must conclude that the beneficial effects thus measured are associated with some additional indirect costs. Such hidden costs must supercede and override the direct benefits included in the measure's original, more limited definition. Further research should confirm or deny the validity of this conjecture.

REFERENCES

Altshuler, A. 1977. Book Review. *Journal of the American Institute of Planners* 43:207–9.

Arnold, E.D. Jr. 1993. Evaluation of Congestion-Reducing Measures Used in Virginia. *Transportation Research Record* 1404:4–7.

Bhatt, K. and T. Higgins. 1989. *An Assessment of Travel Demand Management Approaches at Suburban Activity Centers*. U.S. Department of Transportation, Washington, DC.

Booz-Allen & Hamilton. 2000. *Draft Travel Demand Management Strategy.* Auckland, NZ: Auckland Regional Council.

Cervero, R. 1997. *Paratransit in America: Redefining Mass Transportation.* Westport, CT: Praeger.

Comsis Corporation. 1993. *Implementing Effective Travel Demand Management Measures.* U.S. Department of Transportation, Washington, DC.

Downs, A. 1962. The Law of Peak-Hour Expressway Congestion. *Traffic Quarterly* 16:393–409.

____. 1992. *Stuck in Traffic: Coping with Peak-Hour Traffic Congestion.* Washington, DC: Brookings.

Dueker, K.J., J.G. Strathman, and M.J. Bianco. 1998. *Strategies to Attract Auto Users to Public Transportation.* TCRP Report 40. Transportation Research Board, Washington, DC.

Ferguson, E. 1991. Transportation Management Actions: What Price Traffic Congestion? *Transportation Planning* 18, no. 2:8–10.

____. 2000. *Travel Demand Management and Public Policy.* Aldershot, England: Ashgate Publishing.

Gakenheimer, R. and M. Meyer. 1979. Urban Transportation Planning in Transition: The Sources and Prospects of TSM. *Journal of the American Planning Association* 45:28–35.

Garrett, M. and M. Wachs. 1996. *Transportation Planning on Trial: The Clean Air Act and Travel Forecasting.* Thousand Oaks, CA: Sage.

Howitt, A.M. and A. Altshuler. 1999. The Politics of Controlling Auto Air Pollution. *Essays in Transportation Economics and Policy.* Washington, DC: Brookings.

Institute of Transportation Engineers (ITE). 1989. *A Toolbox for Alleviating Traffic Congestion.* Washington, DC: ITE.

____. 1997. *Trip Generation.* Washington, DC: ITE.

____. 1997. *A Toolbox for Alleviating Traffic Congestion and Enhancing Mobility.* Washington, DC: ITE.

Lee, D.B. Jr. 1978. Book Review. *Journal of the American Institute of Planners* 44:358–61.

Levinson, H.S., M. Golenberg, and K. Zografos. 1987. Transportation System Management—How Effective? Some Perspectives on Benefits and Impacts. *Transportation Research Record* 1142:22–32.

Mandelbaum, S.J. 1996. Making and Breaking Planning Tools. *Computers, Environment and Urban Systems* 20:71–84.

Marshall, T. 1997. Futures, Foresight and Forward Looks: Reflections on the Use of Prospective Thinking for Transport and Planning Strategies. *Town Planning Review* 68:31–53.

McBryan, B., J. Shadoff, D. Pike, and D. Loseff. 1996. *Transportation Demand Management: A Guide for Including TDM Strategies in Major Investment Studies and Planning for Other Transportation Projects.* Seattle: Washington State Department of Transportation.

Misch, M.R. and J.B. Margolin. 1981. *Guidelines for Using Vanpools and Carpools as a TSM Technique*. Washington, DC: Transportation Research Board.

Organisation for Economic Cooperation and Development (OECD). 1994. *Congestion Control and Demand Management.* Paris: OECD.

Orski, C.K. 1987. Can We Manage Our Way Out of Traffic Congestion? *Transportation Quarterly* 41:457–76.

Pushkarev, B.S. and J.M. Zupan. 1977. *Public Transportation and Land Use Policy.* Bloomington, IN: Indiana University Press.

Schonfeld, P.M. and H.S. Chadda. 1985. An Assessment of Urban Travel Reduction Options. *Transportation Quarterly* 39:391–406.

Shadoff, J. TDM Program Manager, Office of Urban Mobility, Washington State Department of Transportation. 1997. Personal Communication. 15 November.

Small, K.A. 1999. Project Evaluation. *Essays in Transportation Economics and Policy.* Washington, DC: Brookings.

Urban Mass Transportation Administration (UMTA). 1977. *Transportation System Management: State of the Art*. U.S. Department of Transportation, Washington, DC.

Wagner, F.A. and K. Gilbert. 1978. *Transportation System Management: An Assessment of Impacts*. U.S. Department of Transportation, Washington, DC.

Walters, A.A. 1961. The Theory and Measurement of Private and Social Cost of Highway Congestion. *Econometrica* 29:676–99.

Weiner, E. 1999. *Urban Transportation Planning in the United States: A Historical Overview.* Westport, CT: Praeger.

Zupan, J.M. 1992. Transportation Demand Management: A Cautious Look. *Transportation Research Record* 1346:1–9.

TABLE A-1 Arnold Data

Congestion-reducing measures	Effectiveness: Number of agencies: 0 = Minimum	1 = Average	2 = Maximum	Sum	Mean score	Cost: Number of agencies: 0 = Inexpensive	1 = Average	2 = Expensive	Sum	Mean score	Implementation: Number of agencies: 0 = Easy	1 = Average	2 = Difficult	Sum	Mean score	Number of agencies reporting use of measure
Increasing supply/adding capacity																
Reconstructing highways with improved design	3	18	37	58	1.59	2	16	39	57	1.65	4	26	26	56	1.39	64
Constructing new highways	2	11	39	52	1.71	1	5	45	51	1.86	4	23	23	50	1.38	56
Widening by adding general purpose lanes	2	18	25	45	1.51	1	20	24	45	1.51	4	24	17	45	1.29	51
Providing highway grade separations	0	6	20	26	1.77	0	3	23	26	1.88	2	8	16	26	1.54	27
Providing railroad grade separations	1	10	14	25	1.52	0	3	21	24	1.88	3	11	10	24	1.29	26
Toll-based financing to expedite new facility construction	0	5	11	16	1.69	2	6	8	16	1.38	1	5	10	16	1.56	13
Constructing HOV lanes	1	1	7	9	1.67	1	3	6	10	1.50	0	3	7	10	1.70	9
Managing existing supply/using existing capacity more efficiently																
Intersection improvements (channelization, turn lanes, signing, bus stop relocation)	2	40	26	68	1.35	10	47	8	65	0.97	23	36	5	64	0.72	75
Other signal improvements, including hardware, upgrades, retiming, removal	4	37	24	65	1.31	17	45	3	65	0.78	44	20	0	64	0.31	71
Coordinated signal systems (arterial, grid, closed loop)	0	27	35	62	1.56	11	46	2	59	0.85	32	21	5	58	0.53	66
Removing/restricting on-street parking	9	29	21	59	1.20	46	11	0	57	0.19	20	23	16	59	0.93	64
Traffic management during reconstruction or other major improvements	10	34	12	56	1.04	14	33	7	54	0.87	27	20	5	52	0.58	60
Turn prohibitions	9	33	11	53	1.04	46	7	0	53	0.13	31	17	4	52	0.48	59
Improving other traffic control devices	8	31	11	50	1.06	20	26	3	49	0.65	30	19	1	50	0.42	54
One-way streets	4	26	15	45	1.24	22	21	0	43	0.49	13	21	9	43	0.91	52
Prohibiting maintenance/repairs on major routesduring peak traffic hours	2	13	35	50	1.66	27	22	1	50	0.48	31	17	1	49	0.39	52
Providing additional lanes w/o widening (shoulders, narrower lanes)	2	16	16	34	1.41	15	14	6	35	0.74	15	15	4	34	0.68	34
Arterial access management	6	5	7	18	1.06	12	4	1	17	0.35	3	10	4	17	1.06	19
Incident detection/management system/program	0	13	1	14	1.07	4	8	1	13	0.77	4	8	1	13	0.77	14
Traffic surveillance/control system	0	4	5	9	1.56	1	6	2	9	1.11	2	4	2	8	1.00	12
Traffic management team	2	7	3	12	1.08	8	3	2	13	0.54	8	4	1	13	0.46	11
Converting existing facilities to HOV facilities	1	6	4	11	1.27	5	1	5	11	1.00	2	2	7	11	1.45	8
Goods' movement management	3	6	0	9	0.67	4	4	0	8	0.50	1	5	3	9	1.22	8
Motorist information system	1	5	2	8	1.13	2	5	2	9	1.00	5	3	1	9	0.56	7
Ramp metering	1	6	0	7	0.86	1	5	2	8	1.13	1	5	1	7	1.00	6
Reversible traffic lanes on arterials	1	6	3	10	1.20	3	6	1	10	0.80	3	4	3	10	1.00	5
Integrated freeway and arterial surveillance/control system	0	5	3	8	1.38	0	4	3	7	1.43	1	4	2	7	1.14	3

continues

TABLE A-1 Arnold Data *(continued)*

Congestion-reducing measures	Effectiveness: Number of agencies: 0 = Minimum	1 = Average	2 = Maximum	Sum	Mean score	Cost: Number of agencies: 0 = Inexpensive	1 = Average	2 = Expensive	Sum	Mean score	Implementation: Number of agencies: 0 = Easy	1 = Average	2 = Difficult	Sum	Mean score	Number of agencies reporting use of measure
Managing/reducing existing demand																
Providing public information on rideshare/transit	14	18	5	37	0.76	22	15	0	37	0.41	30	5	1	36	0.19	40
Park and ride lots	6	20	10	36	1.11	6	25	4	35	0.94	9	22	4	35	0.86	40
Implementing/improving transit fixed-route services	9	19	3	31	0.81	7	17	7	31	1.00	16	12	3	31	0.58	34
Daily flexible work hours (staggered/flextime)	4	18	9	31	1.16	25	6	0	31	0.19	18	10	3	31	0.52	33
Implementing/improving paratransit services	13	13	2	28	0.61	3	17	8	28	1.18	10	14	4	28	0.79	32
Commuter matching services	7	14	8	29	1.03	12	17	0	29	0.59	17	9	2	28	0.46	31
Subsidizing transit usage	8	16	4	28	0.86	3	16	10	29	1.24	14	12	3	29	0.62	28
Promoting nonvehicular alternatives to auto usage	11	13	3	27	0.70	8	15	3	26	0.81	10	13	4	27	0.78	26
Alternative work hours (compressed workweek)	4	17	1	22	0.86	19	3	0	22	0.14	13	5	3	21	0.52	23
Implementing express bus services	3	12	7	22	1.18	2	12	7	21	1.24	6	11	4	21	0.90	22
Implementing transportation management associations	5	11	1	17	0.76	6	12	0	18	0.67	5	13	0	18	0.72	19
Car/vanpool preferential parking	3	12	4	19	1.05	13	4	1	18	0.33	10	6	1	17	0.47	19
Communication in lieu of travel (teleconferencing)	6	10	2	18	0.78	10	8	0	18	0.44	10	6	2	18	0.56	17
Reducing or not increasing transit fares	3	9	1	13	0.85	3	6	5	14	1.14	6	6	2	14	0.71	13
Tax incentives for vanpools	3	7	1	11	0.82	3	8	0	11	0.73	4	5	2	11	0.82	12
Guaranteed ride home program	10	2	1	13	0.31	9	4	0	13	0.31	9	4	0	13	0.31	11
Communication in lieu of travel (telecommuting)	2	10	0	12	0.83	5	6	1	12	0.67	4	6	2	12	0.83	11
Government control of parking supply and location	4	5	0	9	0.56	5	2	2	9	0.67	3	5	1	9	0.78	11
Implementing/improving rail transit services	1	5	4	10	1.30	0	5	6	11	1.55	2	4	5	11	1.27	10
Reduced tolls for ridesharers	4	7	1	12	0.75	5	7	0	12	0.58	6	4	2	12	0.67	7
Differential parking rates	3	5	0	8	0.63	5	2	1	8	0.50	4	4	0	8	0.50	7
Avoiding/controlling demand growth																
Growth management by public policy/ordinance/planning	5	29	8	42	1.07	17	21	4	42	0.69	3	18	20	41	1.41	41
Auto-restricted zones	1	1	2	4	1.25	0	4	0	4	1.00	0	3	1	4	1.25	25
Designing multiuse sites to minimize traffic	3	13	2	18	0.94	5	12	1	18	0.78	5	12	1	18	0.78	18
Mandatory trip reduction for new developments	2	6	4	12	1.17	6	4	1	11	0.55	1	5	6	12	1.42	11
Road/congestion pricing	0	2	0	2	1.00	0	1	1	2	1.50	0	2	0	2	1.00	1
Averages																
Add supply	1.3	9.9	21.9	33.0	1.62	1.0	8.0	23.7	32.7	1.69	2.6	14.3	15.6	32.4	1.40	35.1
TSM	3.3	17.5	11.7	32.4	1.26	13.4	15.9	2.5	31.8	0.66	14.8	12.9	3.8	31.5	0.65	34.0
TDM	5.9	11.6	3.2	20.6	0.87	8.1	9.9	2.6	20.6	0.73	9.8	8.4	2.3	20.5	0.63	21.2
Growth control	2.2	10.2	3.2	15.6	1.06	5.6	8.4	1.4	15.4	0.73	1.8	8.0	5.6	15.4	1.25	19.2
All	**4.3**	**13.8**	**6.7**	**24.9**	**1.10**	**10.0**	**12.2**	**2.4**	**24.6**	**0.69**	**10.9**	**10.2**	**3.3**	**24.4**	**0.69**	**26.3**

HOV = high-occupancy vehicle
TDM = travel demand management
TSM = transportation system management

TABLE A-2 Downs Data

Congestion reducing policies	Effectiveness		Cost		Implementation		
	Extent	Impact	Direct to commuters	To all society	Required institution	Ease of administration	Political acceptability
Supply-side							
Coordinating signals, TV monitoring, ramp signals, electronic signs, converting streets to one-way	Narrow	Minor	None	Minor	None	Moderate	Good
Rapidly removing accidents	Variable	Great	None	Minor	None	Easy	Good
Improving highway maintenance	Broad	Moderate	None	Moderate	None	Moderate	Moderate
Upgrading city streets	Variable	Moderate	None	Moderate	None	Easy	Moderate
Increasing public transit usage by improving service, amenities	Narrow	Minor	None	Moderate	None	Hard	Moderate
Building added HOV lanes	Variable	Moderate	None	Great	Cooperative	Hard	Moderate
Building new and expanding existing off-road transit systems	Narrow	Moderate	Minor	Great	Cooperative	Hard	Poor
Building new roads without HOV lanes	Variable	Moderate	None	Great	Cooperative	Moderate	Poor
Demand-side							
Adopting local growth limits	Narrow	Minor	None	Minor	None	Easy	Good
Encouraging people to work at home	Broad	Minor	None	None	None	Moderate	Good
Changing federal work laws that discourage working at home	Broad	Minor	None	Minor	None	Moderate	Moderate
Clustering high-density housing near transit station stops	Narrow	Minor	None	Minor	Cooperative	Hard	Moderate
Encouraging formation of TMAs, promoting ride sharing	Narrow	Moderate	None	Minor	Cooperative	Hard	Moderate
Staggered work hours	Variable	Minor	None	None	Cooperative	Moderate	Moderate
Providing income tax deductability for commuting allowances	Variable	Great	None	Minor	None	Easy	Poor
Increasing gasoline taxes	Broad	Moderate	Great	Moderate	None	Easy	Poor
Increasing automobile license fees	Broad	Minor	Moderate	Minor	None	Easy	Poor
Eliminating income tax deductabilitiy of free employee parking	Broad	Great	Great	None	Cooperative	Moderate	Poor
Instituting peak-hour tolls on main roads	Broad	Great	Great	None	Regional	Moderate	Poor
Keeping densities in new growth areas above minimal levels	Broad	Moderate	None	Minor	Regional	Hard	Poor
Improving the job-housing balance	Broad	Minor	None	Moderate	Regional	Hard	Poor
Concentrating jobs in big clusters in areas of new growth	Narrow	Minor	None	Great	Regional	Hard	Poor
Parking tax on peak-hour arrivals	Broad	Great	Great	None	Regional	Hard	Poor

HOV = high-occupancy vehicle
TMA = transportation management association

TABLE A-3 Zupan Data

	Effectiveness				Cost			Acceptance				
TDM Solutions	SOV reduction	Peak trip reduction	VMT reduction	Transit impact	Employee	Employer	Public capital	Employee	Employer	Municipal	Political	Ease of implementation index
Alternative work schedules												
Staggered	None	High	None	None	Same	Higher	Same	High	Low	High	High	3
Flex-time	Negative	High	Negative	Negative	Same	Higher	Same	High	Low	High	High	4
Telecommuting	Positive	Positive	Positive	Highly neg.	Lower	Unknown	Same	Unknown	Unknown	High	Medium	Unknown
Four-day week	Medium	Medium	High	Highly neg.	Lower	Unknown	Same	Medium	Low	High	Medium	3
Alternative modes												
Carpools	High	Medium	Medium	Negative	Lower	Varies	Same	Low	Medium	High	High	3
Vanpools	Medium	Low	Low	Negative	Lower	Higher	Same	Low	Low	High	High	2
Subscription buses	Low	Low	Low	Positive	Lower	Higher	Same	Low	Low	High	Medium	2
Parking management												
Preferential parking	Low	Low	Low	None	Same	Higher	Same	Low	Low	High	High	2
Parking pricing	Medium	Medium	Medium	Low	Higher	Same	Same	Negative	Low	Negative	Negative	1
Parking ratios	Medium	Medium	Medium	Positive	Same	Lower	Same	Negative	Unknown	Negative	Negative	2
Park and rides	Medium	Medium	Medium	Positive	Varies	NA	Higher	Medium	Medium	Varies	High	4
Road pricing												
Preferential HOV lanes	Medium	Medium	Medium	Positive	Same	NA	Higher	Varies	Varies	High	Varies	4
Congestion pricing	Medium	High	Medium	Positive	Varies	NA	Lower	Low	Low	Unknown	Negative	Unknown
Transit												
Transitchek	Medium	Medium	Medium	Medium	Lower	Same	Same	High	Medium	High	High	5
Employer sponsored	Low	Low	Low	Low	Varies	Higher	Same	Medium	Low	High	High	4
Employer subsidized	Low	Low	Low	Low	Varies	Higher	Same	Medium	Low	High	High	5
Land use—zoning												
Higher densities	Medium	Medium	High	High	NA	Varies	Lower	NA	NA	Negative	Negative	2
Transit-friendly design	Medium	Medium	Medium	Medium	Same	Same	Lower	Medium	Low	Varies	Positive	4
Mixed-use development	Unknown	Unknown	Medium	Unknown	Lower	Unknown	Lower	Unknown	Unknown	Varies	Positive	4
Growth management	Unknown	Unknown	Unknown	Unknown	Same	Unknown	Lower	Unknown	Unknown	Varies	Varies	3
Institutional												
Trip reduction ordinances	High	High	High	Low/medium	Varies	Higher	NA	Low	Negative	Varies	Varies	4
Transportation management associations	Varies	Varies	Varies	Unknown	Same	Higher	NA	NA	Varies	NA	Positive	NA

HOV = high-occupancy vehicle
NA = not applicable
SOV = single-occupancy vehicle
TDM = travel demand management
VMT = vehicle-miles traveled

TABLE B–1 Arnold Model

Assumptions	Equation 1	2	3	4
Semantic differences				
Effectiveness				
Minimum–average	1	1	1.0	1.0
Average–maximum	1	1	3.2	2.7
Cost				
Inexpensive–average	1	1	1.0	0.0
Average–expensive	1	1	3.2	1.0
Implementation				
Easy–average	1	1	1.0	1.0
Average–difficult	1	1	3.2	1.7
Component weights	Not applicable			
Results				
Estimated coefficients				
Add supply				
Coefficient		0.29	0.31	0.35
Standard error		0.19	0.21	0.21
t-score		1.54	1.49	1.63
TDM				
Coefficient		–0.16	–0.14	–0.15
Standard error		0.17	0.15	0.15
t-score		–0.96	–0.93	–1
Growth control				
Coefficient		0.04	0.00	0.01
Standard error		0.14	0.13	0.13
t-score		0.28	–0.04	0.05
Effectiveness				
Coefficient	0.64*	0.46*	0.54*	0.51*
Standard error	0.14	0.19	0.19	0.19
t-score	4.49	2.42	2.88	2.73
Cost				
Coefficient	–0.09	–0.15	–0.25	–0.28
Standard error	0.16	0.17	0.18	0.20
t-score	–0.55	–0.85	–1.37	–1.38
Implementation				
Coefficient	–0.48*	–0.61*	–0.58*	–0.57*
Standard error	0.16	0.18	0.18	0.18
t-score	–3.06	–3.35	–3.28	–3.19
Goodness-of-fit				
Standard error of *Y*	0.85	0.84	0.83	0.83
R-squared	32.4%	36.7%	38.6%	39.3%
Number of observations	53	53	53	53
Degrees of freedom	50	47	47	47

* Significant at 95% confidence interval.

TDM = travel demand management

TABLE B–2 Downs Model

Assumptions	Equation 1	2	3	4
Semantic differences				
Cost				
None–minor	1	2.6	1	1.8
Minor–moderate	1	1.0	1	1.0
Moderate–great	1	1.0	1	1.0
Component weights				
Effectiveness				
Extent	1	1	1.0	1.0
Impact	1	1	1.3	1.1
Cost				
Direct to commuters	1	1	1.0	1.0
To all of society	1	1	2.2	1.9
Implementation				
Required institution	1	1	13.0	11.0
Ease of administration	1	1	1.0	1.0
Results				
Estimated coefficients				
Supply-side				
Coefficient	0.29*	0.27*	0.43*	0.37*
Standard error	0.14	0.13	0.14	0.13
t-score	2.16	2.11	3.05	2.80
Effectiveness				
Coefficient	–0.22	–0.26*	–0.33*	–0.32*
Standard error	0.14	0.13	0.12	0.12
t-score	–1.58	–2.01	–2.77	–2.79
Cost				
Coefficient	–0.59*	–0.58*	–0.64*	–0.62*
Standard error	0.14	0.13	0.13	0.13
t-score	–4.23	–4.5	–4.79	–4.95
Implementation				
Coefficient	–0.3*	–0.36*	–0.28*	–0.35*
Standard error	0.13	0.13	0.13	0.20
t-score	–2.3	–2.89	–2.2	–2.81
Goodness-of-fit				
Standard error of *Y*	0.61	0.59	0.56	0.55
R-squared	69.3%	71.1%	74.5%	75.4%
Number of observations	23	23	23	23
Degrees of freedom	19	19	19	19

* Significant at 95% confidence interval.

TABLE B–3 Zupan Model

Assumptions	Equation 1	Equation 2	Equation 3	Equation 4
Semantic values				
Effectiveness				
Varies	0	2	2	2
Implementation				
Unknown/not applicable	0	0	3	3
Acceptability				
Varies	0	2	2	2
Negative	–1	–1	1	1
Unknown/not applicable	0	0	2	2
Positive	1	1	3	3
Component weights				
Effectiveness				
SOV reduction	1	1	1	0.0
Peak trip reduction	1	1	1	1.0
VMT reduction	1	1	1	1.7
Transit impact	1	1	1	2.5
Cost				
Employee	1	1	1	0.0
Employer	1	1	1	2.2
Public capital	1	1	1	1.0
Acceptability				
Employee	1	1	1	1.0
Employer	1	1	1	0.0
Municipal	1	1	1	2.7
Public	1	1	1	1.9
Results				
Estimated coefficients				
Effectiveness				
Coefficient	–0.27	–0.34*	–0.46*	–0.45*
Standard error	0.19	0.16	0.14	0.11
t-score	–1.39	–2.12	–3.31	–4.05
Cost				
Coefficient	0.19	0.24	0.05	0.39*
Standard error	0.19	0.16	0.14	0.11
t-score	0.98	1.49	0.38	3.50
Implementation				
Coefficient	–0.41*	–0.56*	–0.66*	–0.56*
Standard error	0.19	0.16	0.14	0.11
t-score	–2.14	–3.46	–4.78	–5.09
Goodness-of-fit				
Standard error of *Y*	0.91	0.76	0.65	0.50
R-squared	29.0%	49.9%	63.9%	78.3%
Number of observations	22	22	22	22
Degrees of freedom	19	19	19	19

* Significant at 95% confidence interval.

SOV = single-occupancy vehicle
VMT = vehicle-miles traveled

TABLE C–1 Arnold Output

Congestion-reducing measures	Effectiveness	Cost	Implementation	Acceptability	SRSD
Add supply	1.70	2.03	1.65	0.36	0.00
Reconstructing highways with improved design	1.53	1.93	1.45	1.78	–1.64
Constructing new highways	1.98	2.74	1.42	1.39	–1.14
Widening by adding general purpose lanes	1.23	1.32	1.12	1.14	–0.62
Providing highway grade separations	2.11	2.75	1.96	–0.03	0.28
Providing railroad grade separations	1.25	2.71	1.18	–0.08	0.37
Constructing HOV lanes	1.98	1.59	2.41	–0.92	1.35
Toll–based financing to expedite new facilities	1.80	1.18	2.02	–0.72	1.40
TSM	0.22	–0.33	–0.21	0.31	0.00
Major intersection improvements	0.58	–0.35	–0.46	2.32	–2.02
Removing/restricting on-street parking	0.31	–0.85	0.24	1.78	–1.86
Traffic management during reconstruction	–0.27	–0.32	–0.73	1.58	–1.49
Turn prohibitions	–0.29	–0.85	–0.95	1.53	–1.1
Minor intersection improvements	0.49	–0.66	–1.4	2.12	–1.09
One-way streets	0.31	–0.85	0.10	1.19	–1.05
Improving other traffic control devices	–0.23	–0.6	–1.15	1.29	–0.71
Coordinated signal systems (arterial, grid, closed loop)	1.32	–0.71	–0.83	1.88	–0.66
Converting existing facilities to HOV facilities	0.42	1.00	1.82	–0.96	–0.18
Goods-movement management	–1.32	–0.85	0.92	–0.96	0.00
Ramp metering	–1.05	0.17	0.20	–1.06	0.44
Arterial access management	0.18	–0.61	0.45	–0.43	0.50
Providing additional lanes without widening	0.88	–0.15	–0.49	0.31	0.54
Prohibiting repairs on major routes during peak traffic	1.79	–0.77	–1.22	1.19	0.74
Incident detection/management system/program	–0.57	–0.54	–0.36	–0.67	0.88
Integrated freeway/arterial surveillance/control system	0.60	0.89	0.69	–1.21	1.03
Reversible traffic lanes on arterials	0.17	–0.44	0.42	–1.11	1.29
Traffic surveillance/control system	1.29	0.05	0.35	–0.77	1.44
Traffic management team	–0.12	–0.22	–0.99	–0.82	1.66
Motorist information system	–0.06	0.05	–0.75	–1.01	1.67
TDM	–0.74	–0.31	–0.52	–0.32	0.00
Implementing/improving paratransit services	–1.23	0.31	–0.24	0.21	–1.34
Park and ride lots	–0.01	–0.38	–0.13	0.60	–0.91
Implementing/improving transit fixed-route services	–0.88	0.07	–0.72	0.31	–0.84
Promoting nonvehicular alternatives to auto usage	–1	–0.38	–0.25	–0.08	–0.61
Subsidizing transit usage	–0.7	0.55	–0.63	0.01	–0.6
Implementing/improving rail services	0.55	1.37	1.19	–0.87	–0.3
Providing public information on rideshare/transit	–0.86	–0.85	–1.6	0.60	–0.26
Implementing express bus services	0.18	0.51	0.07	–0.28	–0.17
Government control of parking supply and location	–1.48	0.05	–0.3	–0.82	–0.12
Reducing or not increasing transit fares	–0.88	0.60	–0.38	–0.72	0.00
Alternative work hours (compressed workweek)	–0.93	–0.85	–0.77	–0.23	0.13
Implementing transportation management associations	–1.04	–0.85	–0.57	–0.43	0.16
Communication in lieu of travel (telecommuting)	–1.08	–0.51	–0.11	–0.82	0.18
Daily flexible work hours (staggered/flextime)	0.09	–0.85	–0.85	0.26	0.21
Commuter matching services	–0.13	–0.85	–0.99	0.16	0.30
Tax incentives for vanpools	–0.88	–0.85	–0.12	–0.77	0.36
Communication in lieu of travel (teleconferencing)	–0.89	–0.85	–0.75	–0.52	0.49
Differential parking rates	–1.38	–0.34	–1.02	–1.01	0.80
Reduced tolls for ridesharers	–1	–0.85	–0.44	–1.01	0.81
Car/vanpool preferential parking	–0.26	–0.62	–0.99	–0.43	0.85
Guaranteed ride home program	–1.64	–0.85	–1.41	–0.82	0.85
Growth control	–0.17	–0.25	0.71	–0.42	0.00
Growth management	–0.28	–0.46	1.53	0.65	–1.85
Trip reduction/transit requirements for new developments	0.20	–0.48	1.55	–0.82	0.21
Auto-restricted zones	0.72	–0.85	0.86	–0.13	0.30
Designing multiuse sites to minimize traffic	–0.65	–0.62	–0.38	–0.47	0.65
Road/congestion pricing	–0.84	1.18	0.00	–1.31	0.68

HOV = high-occupancy vehicle
SRSD = studentized residual
TDM = travel demand management
TSM = transportation system management

TABLE C–2 Downs Output

Congestion reducing policies	Effectiveness	Cost	Implementation	Acceptability	SRSD
Supply–side	–0.2	0.51	–0.43	0.41	0.00
Rapidly removing accidents	0.88	–0.52	–1	1.74	–1.55
Building added HOV lanes	0.08	1.14	0.40	0.41	–1.41
Coordinating signals and signs, one-way streets	–1.45	–0.52	–0.89	1.74	–0.25
Improving highway maintenance	0.80	0.31	–0.89	0.41	–0.07
Upgrading city streets	0.08	0.31	–1	0.41	0.42
New or expanded off-road transit systems	–0.65	1.93	0.40	–0.93	0.57
Building new roads without HOV lanes	0.08	1.14	0.30	–0.93	1.11
Improved transit service, amenities	–1.45	0.31	–0.78	0.41	1.19
Demand–side	0.11	–0.27	0.23	–0.22	0.00
Adopting local growth limits	–1.45	–0.52	–1	1.74	–1.61
Increasing gasoline taxes	0.80	1.97	–1	–0.93	–0.88
Encouraging people to work at home	0.01	–2.01	–0.89	1.74	–0.84
Encouraging formation of TMAs, ridesharing	–0.65	–0.52	0.40	0.41	–0.52
Parking tax on peak-hour arrivals	1.60	–0.35	1.59	–0.93	–0.35
Instituting peak-hour tolls on main roads	1.60	–0.35	1.48	–0.93	–0.28
Concentrating jobs in areas of new growth	–1.45	1.14	1.59	–0.93	–0.25
Improving the jobs-housing balance	0.01	0.31	1.59	–0.93	–0.16
Changing laws that discourage working at home	0.01	–0.52	–0.89	0.41	–0.09
Clustering high-density housing near transit	–1.45	–0.52	0.40	0.41	–0.05
Keeping densities above minimal levels	0.80	–0.52	1.59	–0.93	0.31
Eliminating incentives for free employee parking	1.60	–0.35	0.30	–0.93	0.47
Increasing automobile license fees	0.01	0.70	–1	–0.93	1.04
Staggered working hours	–0.72	–2.01	0.30	0.41	1.29
Tax deductable commuting allowances	0.88	–0.52	–1	–0.93	1.92

HOV = high-occupancy vehicles
SRSD = studentized residual
TMA = transportation management association

TABLE C-3 Zupan Output

TDM solutions	Effectiveness	Cost	Implementation	Acceptability	SRSD
Transitchek (tax deductable vouchers)	1.15	–0.46	–1.77	1.19	–1.81
Transit-friendly design	1.15	–1.15	–0.8	0.10	–1.26
Carpools	–0.59	–0.46	0.18	0.60	–1.22
Staggered shifts	–0.57	1.07	0.18	1.19	–1.22
Preferential parking	–0.64	1.07	1.15	0.60	–1.06
Vanpools	–1.22	1.07	1.15	0.60	–0.53
Four-day workweek	–0.78	–0.46	0.18	0.33	–0.52
Higher density development	2.13	–1.15	1.15	–1.82	–0.48
Subscription buses	–0.05	1.07	1.15	0.04	–0.46
Preferential HOV lanes	0.57	0.23	–0.8	0.33	–0.12
Employer sponsored transit	–0.05	1.07	–0.8	0.89	–0.01
Mixed use development	–0.47	–1.15	–0.8	0.10	0.21
Park and rides	0.57	0.23	–0.8	0.10	0.35
Telecommuting	–1.8	–0.46	0.18	0.33	0.40
Transportation management associations	–0.01	1.07	0.18	0.10	0.45
Flex-time	–1.54	1.07	–0.8	1.19	0.74
Congestion pricing	0.81	–1.15	0.18	–1.32	0.80
Parking ratios	0.57	–1.99	1.15	–2.11	0.86
Growth management	–1.26	–1.15	0.18	–0.46	0.96
Parking pricing	0.57	–0.46	2.12	–2.11	0.98
Employer subsidized transit	–0.05	1.07	–1.77	0.89	1.06
Trip reduction ordinances	1.49	1.07	–0.8	–0.76	1.88

HOV = high-occupancy vehicle
SRSD = studentized residual
TDM = travel demand management

Loglinear Models and Goodness-of-Fit Statistics for Train Waybill Data

HERBERT LEE
Duke University

KERT VIELE
University of Kentucky

ABSTRACT

Counts of carloads of train shipments are effectively described with loglinear models. This paper presents models of counts by origin, destination, and commodity type. Such models can highlight structures in the data and give useful predictions. In particular, there are definite interactions between origin and destination and between origin and commodity, and these models can capture these relationships. Model selection depends on the choice of goodness-of-fit statistic; this paper addresses several issues relating to this choice.

INTRODUCTION

Roughly 1.7 billion tons of cargo is moved by train every year within the United States. In this paper, we explore a statistical method for modeling data from train waybills. In particular, we focus on the counts of carloads of cargo by commodity type and by origin and destination. This information can be arranged into a large three-dimensional table and is thus suitable for analysis via loglinear models. In addition to describing the data, such models allow us to compare flows of freight between different areas, search the data for unusual flows, and make predictions of future flows. Choosing a good model requires the selection of a goodness-of-fit statistic, and we discuss issues involved in this process. We

---

Herbert Lee, ISDS, Duke University, Box 90251, Durham, NC 27708. Email: herbie@stat.duke.edu.

also note several challenges that this data set presents, including a lack of symmetry and a large number of zero counts.

DATA

The data we analyze are from the Carload Waybill Sample issued by the Interstate Commerce Commission for the years 1988 through 1992 (ICC 1992). The data are a stratified sample from all waybills for railroads with over 4,500 cars per year of traffic or 5 percent or more of a state's traffic. There are over 1.9 million total records, each of which has 62 fields of information. Here we focus on three fields: the origin of the shipment, the destination, and the type of commodity. Both the origin and destination are classified into 1 of 181 regions (for the continental United States) as defined by the Bureau of Economic Analysis (BEA) although some are missing or unknown. The commodities are classified by Standard Transportation Commodity Codes (STCC), as per the Association of American Railroads. Using the two-digit aggregate codes gives us 37 categories of commodities in this data set (for example, farm products, coal, printed matter, etc.). Each record in the file is a sample shipment, which may consist of multiple carloads of freight. The sample is stratified, and strata were sampled with different frequencies. Thus, to get an estimate of the total count of carloads of commodity with a particular origin and destination, we first multiply the number of carloads in a record by the inverse sampling frequency and then sum these products over all such commodity shipments from the same origin to the same destination. For example, a record of 7 carloads in a stratum that was sampled with frequency 1 in 40 would get a weighted product of 280 carloads. These sums are entered into a large three-dimensional table, which is then ready for analysis.

As an example of the heterogeneity in the data, we spotlight Chicago, Illinois, and Huntington, West Virginia. Chicago is both the origin of the most traffic, as well as the most frequent destination. Over eight million carloads originate from the Chicago region, and these shipments are spread over many different categories of commodities and are well distributed across the country. In contrast, Huntington is in the top 5 regions by origin of total freight (over 3.5 million carloads), but this freight is nearly all coal. It goes to a smaller number of destinations, and much less freight is sent to Huntington in return. In modeling this data set, we need a model flexible enough to work for both general-freight cities like Chicago and for commodity-specific cities such as Huntington.

Unlike many tables, there is no symmetry in the data since commodities (such as coal) are generally shipped along particular routes, with cities either being origins or destinations but not both. Another potential problem is the large number of zero counts. For example, few things besides coal originate from the Huntington area. However, we do note that these zeroes are not structural zeroes. While many of the zeroes are easily predictable, there is no inherent reason any entry is zero. For example, much freight now moves via intermodal transport, meaning that it could go by truck partway and then be transferred to a train at an intermediate location. Thus, the intermediate location would show as the origin with respect to the train shipment even though it is not the true origin of the commodity.

LOGLINEAR MODELS

Data consisting of counts, such as the waybills, are naturally modeled by the Poisson distribution, which takes values on the nonnegative integers. Instead of a standard regression model with an assumption of Gaussian error, we use a Poisson regression model. Such models are often called loglinear because they are a linear model for the mean after logarithms are taken. Here we model the mean of the distribution of counts from origin i to destination j of commodity k by m_{ijk}. The full loglinear model in this context is

$$\begin{aligned}&\log m_{ijk} = \log a_i + \log b_j + \log c_k + \log d_{ij} + \\ &\log e_{ik} + \log f_{jk} + \log g_{ijk}, \qquad (1)\\ &\text{or } m_{ijk} = a_i b_j c_k d_{ij} e_{ik} f_{jk} g_{ijk}\end{aligned}$$

where a_i is a main effect for origin i (and b_j and c_k are analogous), d_{ij} is an interaction effect for when origin i and destination j have cargo flows not proportional to the product of the main effects a_i and

b_j (e and f are analogous), and g_{ijk} is a three-way interaction between origin i, destination j, and commodity k. The actual counts n_{ijk} of commodity k from i to j thus follow a Poisson distribution with mean m_{ijk}:

$$P(n|m) = \prod_{i=0}^{181} \prod_{j=0}^{181} \prod_{k=1}^{37} \frac{(m_{ijk})^{n_{ijk}} e^{-m_{ijk}}}{n_{ijk}!} \qquad (2)$$

In practice, not all interaction terms may be necessary, and some may be dropped from the model. Also note that the model in (1) is overspecified (there are more free parameters than degrees of freedom), so some sort of restriction is needed. For example, $b_1 = c_1 = d_{i1} = d_{1j} = e_{i1} = e_{1k} = f_{j1} = f_{1k} = g_{i11} = g_{1j1} = g_{11k} = 1$ for all i, j, k. While loglinear models with only a few interaction terms can be fit directly, most complex models require iterative solutions, the most popular method being iterative proportional fitting (Deming and Stephan 1940). For more background on loglinear models, the reader is referred to one of the many good references on the topic (Agresti 1990; McCullagh and Nelder 1989; Bishop et al. 1975).

Loglinear models are part of the same family of models as gravity models (such as Sen and Smith 1995). Gravity models also contain a term relating the distance between the origin and destination to the rate of flow and so would have a term depending on this distance in equation (1). We have found that train cargo flow is not related to distance, and thus the additional term in the gravity model is unhelpful for our data. In contrast to focusing on modeling the effect of distance, we focus on the complex interaction effects of the covariates.

In this paper, we take the Bayesian approach. The gamma distribution serves as a conjugate prior for all parameters, and the posterior can be easily estimated via Markov chain Monte Carlo. With the full posterior, one can easily get estimates of uncertainty, in addition to simple point estimate. Either an informative prior or a noninformative (improper) prior can be used. Using a noninformative prior leads to posterior mode estimates equal to the maximum likelihood estimates. We use an essentially noninformative prior. More details on Bayesian loglinear models can be found in Gelman et al. (1995), and West (1994) discusses Bayesian loglinear models in the context of gravity models.

ASSESSING GOODNESS-OF-FIT

To compare how well different models fit, we employed cross-validation (see Stone 1974). For this data set, annual counts seemed a natural unit of validation. Thus for each year s =1,…,5, we fit each model under consideration using the other 4 years of data and used the fitted model to predict the counts for year s. These fitted counts were then used to compute a goodness-of-fit statistic $q_{r,s}$ for model r for year s. To get the overall cross-validation score, the goodness-of-fit statistics are summed across all years giving $q_r = \sum_{s=1}^{5} q_{r,s}$. The rest of this section discusses choices of goodness-of-fit statistics.

Mean square error is an appropriate goodness-of-fit statistic when the variance of observations is the same for all observations (not true for Poisson data) or when one is not interested in adjusting for differing variances, such as when one is most interested in predicting the largest table entries correctly, that is, when nominal error is more important than relative error. This may be the case for train data in that predicting 100 carloads when the truth was 200 (a nominal error of 100, relative error of 100%) is much less of a concern than predicting 100,000 carloads when the truth is 150,000 (nominal error of 50,000, relative error of 50%). Those 50,000 extra carloads could represent a much larger logistical problem than the 100 extra carloads, in which case mean square error would be a useful summary. Equivalent to mean square error is its square root, root mean square error (RMSE), which has the advantage of being on the scale of the data and thus being more interpretable.

Alternatively, one may be more interested in relative error. Statistical theory says that one should adjust for the variance in computing goodness-of-fit. The Pearson chi-squared statistic is

$$X^2 = \sum_i \sum_j \sum_k \frac{(n_{ijk} - \hat{m}_{ijk})^2}{\hat{m}_{ijk}} \qquad (3)$$

where n_{ijk} is the actual count and $\hat{m}_{ijk}$ is the predicted count. When the model holds, X^2 is asymptotically distributed as a chi-squared distribution (see for example, Agresti 1990). The denominator in (3) is the estimated variance of the prediction, and thus X^2 is a measure of relative error. However, for an application such as cargo, it does not make much sense to inflate the error when the prediction is smaller than one. For example, if the model predicts 0.1 carloads in a year, and in truth 2 carloads were observed, the contribution to X^2 would be $(2 - .1)^2/.1 = 36.1$, larger than the nominal error. When routes have hundreds of thousands of cases, a nominal error of 2 carloads is rather insignificant, and its contribution to the total error does not seem like it should be inflated. As a further complication, when this goodness-of-fit statistic is used for predictions, the model might predict a count of zero when the actual count could be nonzero. In that case, X^2 is infinite, and it is impossible to compare models. If a small number were added to each cell of the table, the comparison of models can depend on the size of the value added. To avoid these problems, we modify X^2 so that the denominator is no smaller than one:

$$\tilde{X}^2 = \sum_i \sum_j \sum_k \frac{\left(n_{ijk} - \hat{m}_{ijk}\right)^2}{\max\left\{1, \hat{m}_{ijk}\right\}}$$

The Cressie-Read power-divergence family of goodness-of-fit statistics (Read and Cressie 1988), indexed by a single power parameter λ, is a general family that includes many common measures as special cases, including the Pearson chi-square and the loglikelihood ratio statistic. Most of the members of this family have the common problem of being undefined for prediction either when some entries predicted to be zero are nonzero $(\lambda \geq 0)$, or when there are zero entries that were predicted to be nonzero $(\lambda \leq -1)$. One standard goodness-of-fit measure is in the intermediate power range and thus is directly applicable to prediction with zero entries: the Freeman-Tukey statistic, given here as parameterized in Fienberg (1979):

$$F^2 = 4 \sum_i \sum_j \sum_k \left(\sqrt{n_{ijk}} - \sqrt{\hat{m}_{ijk}}\right)^2$$

F^2, employing the variance-stabilizing transformation for a Poisson distribution, represents a compromise between the mean square error and the Pearson chi-square statistic. Note that while all members of the Cressie-Read family have the same asymptotic chi-square distribution, their distributions may be different for finite samples. In particular, when the data table is sparse (with many zeroes, as with the waybill data), there can be problems with the chi-square approximation for all of the statistics (for example, Koehler 1986).

DATA ANALYSIS

The models under serious consideration were the full model (equation 1), the model without a three-way interaction (*g* of equation 1) but including all two-way interactions, and the three models with no three-way interaction and only two two-way interactions (that is, no *g* and only two of *d, e,* and *f* in equation 1). Models with fewer terms were unable to capture the complexity of the data. Table 1 compares goodness-of-fit statistics for all models with at least two two-way interaction terms. We note that we can not use the unmodified X^2 statistic because during cross-validation, some entries predicted to be zero are instead nonzero, leading to infinite values of X^2.

From the table, we see that the choice of best model does depend on the choice of measure of goodness-of-fit. The full model seems best for reducing absolute error since it has the lowest RMSE (and does fairly consistently for each year of the cross-validation). If relative error is more important, the model using only two-way interactions for origin versus destination and for origin versus commodity performs best. Also of note is the

TABLE 1 Cross Validation Goodness-of-Fit Statistics for the Top Models

Model	RMSE	F^2	$\tilde{X}^2$
Full model	2029.2	6.59e+7	3.71e+9
All 2-way interactions	2118.2	9.28e+7	3.34e+9
No origin-destination interaction	4052.0	2.61e+8	3.81e+9
No origin-commodity interaction	4418.4	3.04e+8	4.03e+9
No destination-commodity interaction	2606.5	1.51e+8	2.86e+9

Note: RMSE is the root mean squared error of prediction, F^2 is the Freeman-Tukey statistic, and $\tilde{X}^2$ is the modified Pearson statistic.

model with all two-way interactions. It is a reasonable compromise model having an RMSE close to that of the full model yet also having the second lowest F^2 and X^2. Thus, this model might be chosen for its robust performance with respect to multiple goodness-of-fit statistics.

Substantively, it is interesting that the other models with only two two-way interactions do not perform as well. It seems clear that any reasonable model must include both an interaction term for origin versus destination and a term for origin versus commodity. An instructive example is that of Huntington. As mentioned earlier, Huntington primarily exports coal and only to a specific set of destinations. Yet Huntington is one of the largest areas in terms of total number of carloads. Thus, any model must be able to account for both the fact that Huntington exports a very large amount of coal but little else as well as the fact that it exports large amounts to a relatively small number of destinations, unlike general shipping hubs like Chicago. In contrast, there are no obvious examples of cities that are destinations for large amounts of particular commodities out of balance with their imports of other commodities, so the removal of the interaction term for destination versus commodity has much less impact on the fit of the model.

CONCLUSIONS

The train waybills data set is interesting both for its information on commodity flows as well as for its statistical challenges. Loglinear models provide an effective method for describing the relationship between cargo volume and origin, destination, and commodity type. The size of the data set[1] is much larger than in a standard statistical problem. While this size is beyond the capabilities of many standard statistical software packages, loglinear models can be programed directly.

Model selection raised a number of statistical issues. In contrast to many data sets used with loglinear regression, the waybills are sorted by year, providing a natural breakdown for cross validation. The choice of goodness-of-fit measures has been discussed. Dealing with the large number of zero counts during cross-validation appears to be a topic not fully addressed in the statistical literature. The analyses of this paper should be seen as a starting point for further work, both methodological and relating to the interpretations of the indicated models.

[1]Approximately 2 million records in the data file, resulting in over 119 million carloads distributed in a three-way table containing over 1.2 million cells.

ACKNOWLEDGMENTS

The authors would like to thank three anonymous referees and the Editor-in-Chief for their valuable suggestions. This work was partially supported by National Science Foundation grants DMS-9971954 and DMS-9873275.

REFERENCES

Agresti, A. 1990. *Categorical Data Analysis*. New York: Wiley.

Bishop, Y.V.V., S.E. Fienberg, and P.W. Holland. 1975. *Discrete Multivariate Analysis*. Cambridge, MA: MIT Press.

Deming, W.E. and F.F. Stephan. 1940. On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. *Annals of Mathematical Statistics* 11:427–44.

Fienberg, S.E. 1979. The Use of Chi-Squared Statistics for Categorical Data Problems. *Journal of the Royal Statistical Society B* 41:54–64.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 1995. *Bayesian Data Analysis*. London: Chapman & Hall.

Interstate Commerce Commission (ICC). 1992. Carload Waybill Statistics: Territorial Distribution, Traffic and Revenue by Commodity Class. Published on CD-ROM by the Bureau of Transportation Statistics, Washington, DC.

Koehler, K.J. 1986. Goodness-of-Fit Tests for Log-Linear Models in Sparse Contingency Tables. *Journal of the American Statistical Association* 81:483–93.

McCullagh, P. and J.A. Nelder. 1989. *Generalized Linear Models*. London: Chapman & Hall.

Read, T.R.C. and N.A.C. Cressie. 1988. *Goodness-of-Fit Statistics for Discrete Multivariate Data*. New York: Springer-Verlag.

Sen, A. and T.E. Smith. 1995. *Gravity Models of Spatial Interaction Behavior.* Berlin: Springer-Verlag.

Stone, M. 1974. Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society B* 36:111–47.

West, M. 1994. Statistical Inference for Gravity Models in Transportation Flow Forecasting. Technical Report 94–40, Duke University, Institute of Statistics and Decision Sciences, Durham, NC.

Estimation and Evaluation of Full Marginal Costs of Highway Transportation in New Jersey

KAAN OZBAY
BEKIR BARTIN
Rutgers University

JOSEPH BERECHMAN
Tel Aviv University

ABSTRACT

In this study, we present a methodology for estimating full marginal transportation costs of highway transportation in New Jersey. This methodology is specifically applied to the northern New Jersey highway network. We review the existing studies and identify the highway transportation cost categories. Cost functions are developed using New Jersey-specific data for each cost category. Along with the total cost functions, marginal costs functions are derived. These marginal cost functions are used in the application of our full marginal cost estimation methodology. Finally, the resulting marginal cost values for northern New Jersey are analyzed according to various trips distances, urbanization degrees, and highway functional types.

INTRODUCTION

There is a growing interest among transportation agencies in determining the full cost of transportation services for both short- and long-term planning purposes. The main objective behind this interest is to ensure prices paid by transportation users correctly reflect the true costs of providing transportation services. Economists argue that "getting prices right might not be the end of economic development, but getting prices wrong

Kaan Ozbay, Rutgers University, Department of Civil and Environmental Engineering, 623 Bowser Road, Piscataway, NJ 08854. Email: kaan@rci.rutgers.edu.

frequently is" (Meier 1983, 1 and 231; Timmer 1987, 39). In the case of transportation, optimal user charges should be equal to the value of the resources consumed through the use of transportation facilities. For example, for road users prices charged should consist of the damage done to the road surface (variable road maintenance costs) and the additional costs (mainly congestion costs) each user imposes on other users and the rest of society (Walters 1968; Churchill 1972). Thus, it is extremely important to accurately estimate the full cost of various modes of transportation for a given study area in order to develop effective long-term transportation pricing schemes.

This paper is mainly concerned with the estimation of the full marginal costs of highway transportation in New Jersey and the analysis of these cost models through their application to a northern New Jersey network. By "full marginal costs" we mean the full social costs of transporting an additional trip-maker over the highway network system. This information is mandatory for the development of an efficient transportation pricing scheme.

This paper has two major objectives.

1. To develop a general cost model to estimate the full costs of highway passenger transportation using New Jersey-specific data.
2. To apply this cost model to the northern New Jersey highway network to estimate the factors that affect the full cost of highway transportation in the study area. The results of this second step are amenable to policy interpretation aimed at developing efficient policies to improve the performance of the New Jersey transportation system. In the final section, we present a comparison of the costs we estimate and the user revenue collected by the government, reflecting the efficiency of the roadway pricing in New Jersey.

The full measure of highway transportation costs are usually categorized as "direct" and "indirect" costs. Direct costs, sometimes called "private" or "internal costs," include the costs auto users directly consider monetary losses, such as vehicle operating cost, car depreciation, time lost in traffic, infrastructure cost (through taxes), and so forth. Indirect costs, also called "social" or "external costs," refer to the costs auto users are not held accountable for, including those every user imposes on the rest of traffic, such as the costs of congestion, accidents, air pollution, and noise. An extensive literature review yielded the cost categories and data sources shown in table 1.

Most of the previous studies dealing with the estimation of transportation costs focus on the average cost of highway transportation (Tellis and Khisty 1995; Churchill 1972; Cipriani et al. 1998; PMSK 1993; TRB 1996). On the other hand, only a few studies deal with the estimation of marginal costs (Levinson et al. 1996, Levinson and Gillen 1998; Mayeres et al. 1996). Levinson et al. (1996) deals with both marginal and full costs of supplying transportation services. Mayeres et al. (1996) deals with the estimation of marginal external costs only. The "British Columbia Lower Mainland" study (PMSK 1993) uses societal costs such as that of roadway land value, of air and water pollution, of accidents, and of the loss of open space.

The importance of focusing on the marginal costs of service provision in a given area stems from the fact that marginal costs measure the actual increase in costs from an additional mile (or trip) traveled. Thus, marginal costs represent the additional costs the state should consider when encouraging efficient transportation use. Although traditional government cost allocation studies have gradually incorporated concepts similar to marginal costing, non-governmental costs are still largely ignored. However, the costs of congestion, pollution, and accidents are real costs to the government as well as to society. In brief, a marginal cost approach that includes practically measurable external costs tends to be more realistic.

PROPOSED MARGINAL COST ESTIMATION METHODOLOGY

In this paper we consider a common situation: the marginal cost of highway travel is higher than the average cost, reflecting the fact that an additional vehicle in traffic imposes a definite cost on all users (Mohring 1976). Figure 1 demonstrates this specific case. Due to the lack of a pricing policy that sets the price to users equal to full marginal costs (*FMC*), highway transportation infrastructures are over utilized, auto and truck users do not pay for what they consume, and the cost to society of serv-

TABLE 1 Major Cost Categories and Data Sources

Cost categories	Payer	Data sources
Vehicle operating cost ■ Auto ownership ■ Auto operations (gasoline + maintenance + insurance) ■ Tolls ■ Insurance	Private	NJDOT, Internet resources (Kelley Blue Book online), American Automobile Manufacturers Association (AAMA)
Infrastructure costs ■ Capital ■ Maintenance and improvements ■ Right-of-way	Public	NJDOT
Environmental costs ■ Air pollution ■ Noise	Public and private	Existing studies and NJDOT, U.S. Environmental Protection Agency (USEPA)
Congestion costs ■ Travel time	Private	NJDOT
Accident and safety costs ■ Bodily and property damage ■ Productivity ■ Emergency and medical services (police + ambulance + rescue)	Public and private	NJDOT

ing an additional trip is higher than the average cost at that demand level[1] (see point *A* in figure 1).

The formulation of the *FMC* involves the cost of making a trip between origin-destination (O-D) pairs in a network, which is a function of several variables, here denoted by V_j. The average cost, C_{rs}, of one trip between a specific O-D pair (*r, s*) follows:

$$C_{rs} = F\left(V_j ; q\right) \tag{1}$$

where, q denotes the demand between the O-D pair. We assume that there are q homogeneous users who make the same trip over a given time period.[2] The full total cost (*FTC*) of providing a transportation service between any O-D pair for q trips is defined as follows:

$$FTC_{rs} = qC_{rs} = qF\left(V_j ; q\right) \tag{2}$$

From (3), we obtain *FMC* for each O-D pair (*r, s*) over a given time period as follows:

$$FMC_{rs} = \frac{\partial\left(qF\left(V_j; q\right)\right)}{\partial q} = F\left(V_j; q\right) + q\,\frac{\partial F\left(V_j; q\right)}{\partial q} \tag{3}$$

This function gives the cost of adding an extra trip to the system. The first term represents the average cost, and the second term represents the additional cost of a trip. Thus, if we add one more user making an extra trip, the cost imposed by an additional trip to the rest of the traffic is $q\left(\partial F\left(V_j ; q\right) / \partial q\right)$. This cost amount is an externality, and we refer to this term as "congestion-related costs." In figure 1, the difference $C^*–C_1$, is equal to this term.

Thus, we define *FMC* of an additional trip as

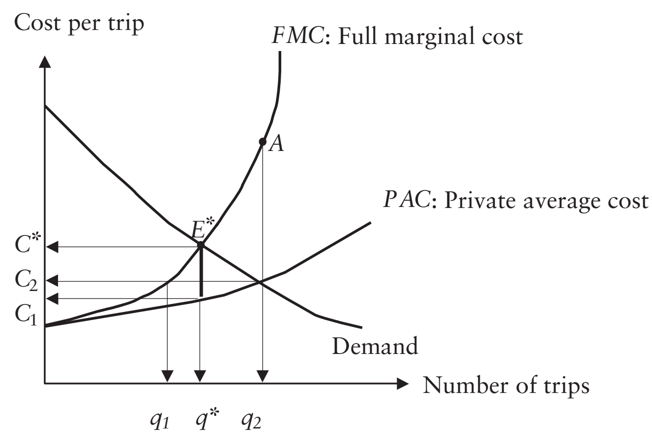
FMC ($) = private average cost ($) + congestion-related costs ($)

In terms of figure 1, computation of *FMC* is at the point of social equilibrium (E^*), where C^* is the optimal price. If the optimal cost is determined by

[1] Here we assume that highway prices are most likely equal to average cost after political considerations.

[2] The term "user" here connotes a vehicle-trip.

FIGURE 1 Hypothetical Marginal and Average Costs of Highway Transportation



q^* = Number of trips at social equilibrium
C^* = Cost per trip at social equilibrium
C_1 = Private cost per trip at social equilibrium
C_2 = Private cost per trip at private equilibrium
q_2 = Number of trips at private equilibrium

setting the tolls equal to *FMC* evaluted at $q^*, \partial FTC(q^*)/\partial q$, then the total revenue of tolls (*TR*) will be (Small 1992)

$$TR = \left(\frac{\partial FTC(q^*)}{\partial q}\right) q^* = \frac{1}{s} FTC \tag{4}$$

where *FTC* is full total cost, q is the demand and s is the degree of economies of scale. Equation (4) implies the known rule that total cost will be covered if $s \leq 1$ (where $s = AC/MC$ and *AC* is average cost and *MC* is marginal cost). As mentioned before, since marginal cost is usually higher than average cost for highly congested highways, the toll revenue compensates the full total cost of highway transportation even when no fuel tax is charged.

Estimation and Analysis of Network-Wide Full Marginal Costs

As discussed, full marginal costs is defined as the total costs accrued to society from an additional unit of travel, that is, an additional user. Although highway transportation might seem to be the production of a single output in a highway network, the reality is more complicated, in part because users make decisions within the network to minimize their own costs. They change their routes and times of travel constantly, based on network attributes, such as travel demand, number of routes between O-D pairs, capacity of each link, and so forth. Hence, if we introduce an additional demand between a selected O-D pair, not only do the travel patterns on each route connecting that O-D pair change, but the travel patterns on every route in the network will also change.

In multiple origin-destination and multiple-route networks, the practical and operation calculation of the network-wide marginal cost is complicated by the following issues.

- Do we add an extra demand unit between every O-D pair or do we pick one O-D pair and add the extra unit of demand to this pair? If so, which O-D pair do we select?
- What is the effect of this extra unit of demand on the overall network equilibrium? Does the addition of one extra unit flow (a vehicle) to a large network affect the overall equilibrium condition?

To address these issues in our proposed network-wide full marginal cost estimation methodology, we assume that the additional flow in the system does not disturb the existing flow patterns in the network.[3] We then propose adding this additional

[3] Jara-Diaz et al. (1992) has introduced the idea of network-wide marginal costs in the context of a freight network in Chile.

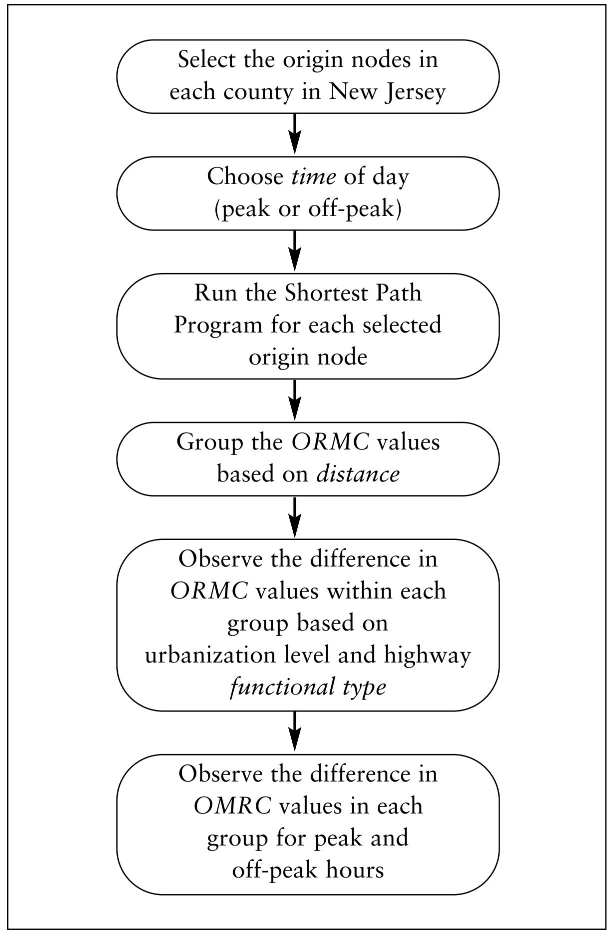
trip between a selected O-D pair onto the shortest route between this specific O-D pair and calculating the marginal cost for this trip. We call this marginal cost one-route marginal cost (*ORMC*). We are aware of the fact that the resulting value will not be the same as the true system-wide marginal cost. This value can only be obtained by performing a new traffic equilibrium assignment, which will reflect the change in flow patterns due to the addition of an extra unit of demand. However, compared to the overall demand, because the additional demand is relatively small (a single trip between a given O-D pair, we can assume that the resulting costs will be reasonable approximations of actual costs. A detailed explanation of the theoretical implications of this marginal cost estimation procedure is given in Ozbay et al. (2000).

To estimate network-wide marginal costs in the northern New Jersey network, we first determine marginal costs along the shortest routes for each individual O-D pair in the network. We then group the O-D pairs according to several quantitative and qualitative factors similar to the ones listed in Jara-Diaz et al. (1992). For each O-D pair, these factors include 1) level and variance of demand flow, 2) traffic conditions, and 3) factors induced by movements between other O-D pairs. Additionally, some physical factors relate to the O-D pair: 1) topography, 2) climate, and 3) characteristics of the corresponding right-of-way. For practical reasons, Jara-Diaz et al. (1992) does not include all of these factors in the marginal cost function. Instead the authors group the observations based on the qualitative factors. They also develop separate marginal cost functions for each category. In this analysis, we follow a similar approach. Given data availability, the factors considered in this study are

- Distance between O-D pairs
- Functional type (percentage of highway functional types on the shortest routes between O-D pairs, such as interstate, freeway, arterial, and so forth)
- Residential density of the areas where the shortest routes are located (central business district, urban, suburban, or rural)
- Time of the day (peak hours or off-peak hours).

Figure 2 depicts the process of calculating *ORMC* for several O-D pairs, grouping them according to

FIGURE 2 *ORMC* Calculation Process



the factors listed above for the northern New Jersey highway network link volumes provided by the New Jersey Department of Transportation (NJDOT).

The northern New Jersey network used in this process consists of 5,418 nodes, 1,451 of which are zonal,[4] and a total of 15,387 links. Shortest routes between zones are determined using a computer program developed in Avenue,[5] based on Pape and Moore's shortest path algorithm (Pape 1974). Every time the shortest route between an O-D pair is determined, desired link properties (such as distance, functional type of the highway, residential density, travel time, county name, and traffic volume) are extracted by the program. As for the time of the day, we use peak and off-peak loaded networks and perform the analysis for these two time intervals. Figure 2 shows the *ORMC* procedure.

[4] A zonal node means to origin-destination zones where trips originate and end.
[5] Avenue is an object-oriented programming language used to create user interface for ArcView GIS.

DEFINITION AND FORMULATION OF COST FUNCTIONS

In this study, we have reclassified highway transportation costs into three major categories: user costs, infrastructure costs, and environmental costs. We develop the total and marginal cost functions for each category using New Jersey-specific data.

User Costs

User costs are put into two major groups: 1) self-vehicle operating costs, that is, car ownership, fuel and oil consumption, regular or unexpected maintenance, and so forth and 2) user interaction costs, that is, accident- and congestion-related costs.

User interaction costs are difficult to calculate for the following reasons:

- The key unit values, value of time (*VOT*) and value of life-injury, are mostly based on the judgment of highway users
- Accident- and congestion-related costs are interrelated and affect all auto users. For instance, at first glance an accident seems to incur costs only for the parties involved. However, the resulting delay causes congestion, making for low-speed operating conditions and time loss for other users. Figure 3 shows the various categories of user costs.

The following two subsections present the formulation of marginal user-cost functions, self-vehicle operating costs and user interaction costs.

Self-Vehicle Operating Costs

Self-vehicle operating costs include vehicle depreciation, fuel, oil, tire-wear, insurance, parking fees, tolls, and regular and unexpected maintenance. The general form of an operating cost function follows.

$$C_{opr} = f\left(C_d, C_g, C_o, C_t, C_m, C_I, C_{pt}\right) \qquad (5a)$$

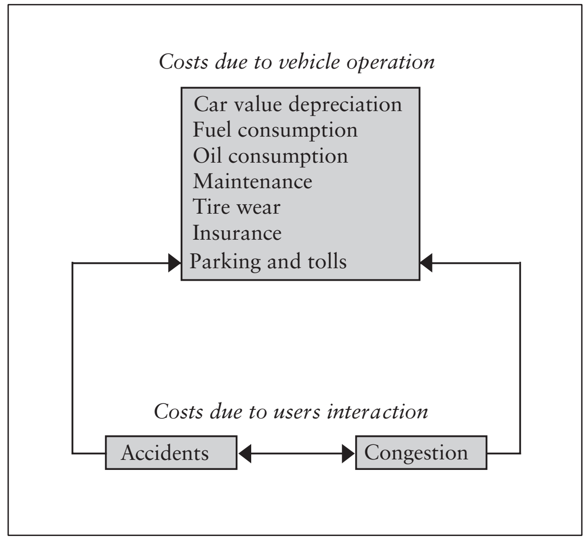
where

C_{opr} is vehicle operating cost over many years (dollars/vehicle)

C_d is depreciation cost for a vehicle over many years

C_g is gas cost (dollars/mile)

FIGURE 3 User Cost Categories



C_o is oil cost (dollars/mile)

C_t is tire cost (dollars/mile)

C_m is maintenance cost (dollars/mile)

C_I is insurance cost (dollars/year)

C_{pt} is parking fees and tolls (dollars/mile).

Depreciation is caused by wear and tear on the vehicle over time and by the change in demand and taste of users. Hence, depreciation cost is assumed to be related to the vehicle's mileage and age. Maintenance, fuel, oil, and tire-wear costs and parking fees and tolls depend mainly on the distance traveled.[6] We used Kelley Blue Book (2000) to estimate our vehicle depreciation cost function. The Honda Civic is taken as the representative car model since it has been the best-selling economy car in the United States for several consecutive years (LTA 1998). The statistical results of this analysis are given in table 2.

Data on insurance costs, parking fees, and tolls are from *Cost of Owning and Operating Automobiles, Vans, and Light Trucks* (USDOT FHWA 1991). Maintenance, oil, and fuel and tire-wear costs are taken from American Automobile Manufacturers Association (AAMA) (1996), in which the cost values are given as national averages and defined on a per mile basis. Table 3 provides a summary of these unit costs.

[6] Here we disregard the effects of traffic, volume, temperature, and altitude on fuel and maintenance costs.

TABLE 2 Results of Regression of Depreciation Cost

Sample size: 217	Coefficients	Value	Standard error	*t*-value	Pr(>|*t*|)	
Intercept	α_0	6,240.3569859	115.0817835	54.2254108	0.00	
m/a	α_1	0.1035028	0.0034358	30.1247472	0.00	
a	α_2	677.7279739	11.9850174	56.5479340	0.00	R_2=0.94

$$C_d = \alpha_0 + \alpha_1(m/a) + \alpha_2 a$$

TABLE 3 Operating Costs

Operating expenses	Costs*
Gas and oil	0.061 (dollars/mile)
Maintenance	0.029 (dollars/mile)
Tires	0.0145 (dollars/mile)
Insurance	1,350 (dollars/year)
Parking and tolls	0.0182 (dollars/mile)

* 2000 dollars (without tax)

Sources: AAMA (1996), USDOT (1991)

The vehicle operating cost function is developed by combining the values given in tables 2 and 3.

$$C_{opr} = 6{,}240.36 + 0.104\frac{m}{a} + 2{,}027.73a + 0.1227m \qquad (5b)$$

where C_{opr} is vehicle operating cost (dollars/vehicle over many years)
m is the vehicle mileage (miles)
a is the vehicle age (years).

The regression analysis results given in table 2 indicate a depreciation cost possibly higher than would be expected. However, since our depreciation cost function uses the trade in values, the resulting depreciation cost reflects real world values.

Marginal vehicle operating cost is estimated in terms of distance traveled, and this assumption cancels out insurance cost in the marginal cost formula since these are usually defined in terms of vehicle age. Marginal vehicle operating cost (MC_{opr}) per mile is estimated as

$$MC_{opr} = 0.1227 + \left(\frac{0.104}{a}\right) \qquad (6)$$

It is clear from equation (6) that marginal vehicle operating cost per distance decreases as the vehicle gets older. Intuitively, the longer the vehicle is utilized, the lower the marginal cost of running it becomes. This is mostly due to decreasing marginal deprecation cost over time (see table 4). In our analysis, we used an average vehicle age of 8.5 years, reflecting the national average in the United States, as reported in AAMA (1996).

User Interaction Costs: Congestion and Accident Costs

Congestion costs are defined as time loss and discomfort for drivers. The magnitude of these costs is directly related to the time lost and to user characteristics.

- Time loss is determined through the use of a travel time function and trip characteristics, such as distance between O-D pairs, traffic volume, and highway capacity. Once the trip characteristics are known, the travel time function is used to calculate the time lost in the traffic between each O-D pair.
- User characteristics, on the other hand, are expressed through the dollar value each user places on a specific unit of his or her time. However, user characteristics are not homogeneous and not easy to identify. Thus, in this study we use an average "Value of Time" (*VOT*). Small (1992) suggests that *VOT* should be taken as 50% of the gross wage rate. This value can vary among different states and cities. Thus, we decided to use a range of 40 to 170% of the pre-tax hourly wage rate as the *VOT* in our analyses. The New Jersey Department of Labor reports the hourly wage rate in 2000 as \$19 per hour (NJDOL 2000). Thus, our *VOT* ranges from \$7.6 to \$32.3.

In this study, we employed the commonly accepted travel time function, the Bureau of Public Road's (BPR) volume-capacity function. Using the

TABLE 4 Contribution Percentages of Operating Cost Categories

Year	Depreciation (percent)	Fuel, gas, oil, tires, parking, and tolls (percent)	Insurance (percent)
1	74	14	12
2	60	22	18
3	52	26	22
4	47	29	24
5	43	31	26
6	40	33	28
7	37	34	29
8	36	35	30
9	34	36	30
10	33	36	31
11	32	37	31
12	31	37	32
13	30	38	32

Note: An annual mileage of 13,000 miles is assumed.

BPR function, total congestion costs on link *a* to *b*, with a traffic volume of *Q*, is calculated as follows.

$$C_{cong} = QT_{ab}(VOT) = QT_0\left(1+0.15\left(\frac{Q}{C}\right)^4\right)(VOT) \quad (7)$$

The marginal congestion cost function is then the first order derivative in terms of traffic volume, *Q*.

$$MC_{cong} = \frac{\partial C_{cong}}{\partial Q} = Tab(VOT) + Q(VOT)\frac{\partial Tab}{\partial Q} \quad (8)$$

where

MC_{cong} is the marginal cost of congestion (dollars/hour)

VOT is the average value of time (dollars/hour)

T_0 is free-flow travel time between points *a* and *b* (hours)

T_{ab} is the travel time required to travel between points *a* and *b* (hours)

Q is the average volume on the link connecting points *a* and *b* (vehicles/hour)

C is highway capacity (vehicles/hour).

The first term in the right-hand side of equation (8) represents the time cost experienced by one user, and the second term stands for the cost imposed on the rest of the users on that link.

The second type of user interaction costs, accident costs, can be classified into two major groups:

- Foregone production or consumption by individuals or both, easily converted into monetary units
- Life-injury damages by users, not easily converted into monetary units.

The available New Jersey data contain a detailed accident summary for 1995, including the pedestrians affected, grouped by incident types and by county in New Jersey.

In order to estimate the cost of accidents over a given period of time, we need to know the accident occurrence rate (number of accidents over time) and the unit cost of an accident. If we develop a function to estimate the number of accidents occurring over a period of time, accident costs can also be measured by multiplying the number of accidents by their unit cost values. Clearly, costs vary, accident by accident. However, similar accidents have costs that fall more or less in the same range. Thus, we classified accidents as 1) fatal, 2) injury, or 3) property damage.[7]

There are also various geometric design features of a roadway that affect the possibility of an accident, such as the number of lanes, horizontal and vertical alignment, superelevation, sight obstructions, and so forth. However, it is not an easy task to include every variable in the accident occurrence rate function. Thus, the accident occurrence rate is assumed to be correlated only with highway functional type, average traffic volume, and the length of the highway. For this purpose, highways are grouped into three categories according to their functional properties. These are interstate, freeway-expressway, and arterial-collector-local.[8]

The generalized form of the total accident cost function is given as follows:

$$C_{acc} = \sum_{i=1}^{3} C_f p_{fi} + C_h p_{hi} + C_d p_{di} \quad (9)$$

[7] Vehicle fire and cargo spills are disregarded since their occurrence rates are relatively negligible.

[8] The classification of highways is based on the available accident data.

where

C_{acc} is the total accident cost per year (dollars/year)

C_f is the unit cost of a fatal accident per crash (dollars)

C_d is the unit cost of a property damage accident per crash (dollars)

C_h is the unit cost of an injury accident per crash (dollars)

p_{fi} is the number of fatal accidents per year for highway type *i*

p_{hi} is the number of personal injury accidents per year for highway type *i*

p_{di} is the number of property damage accidents per year for highway type *i*.

It should be noted that accident cost as given in equation (9) does not include the costs of congestion caused by accidents. In order to utilize the equation (9), we have to develop p_{fi}, p_{hi}, and p_{di} functions using the available accident data. As mentioned above, the number of accidents is assumed to be correlated with roadway length (M), as a measure of network properties, and average traffic volume (Q), as an output measure. The general form of the accident occurrence rate (the number of accidents over a given time period) function is given as follows:

$$p = \alpha_1 M^{\alpha_2} Q^{\alpha_3} \tag{10}$$

where $\alpha_1, \alpha_2, \alpha_3$ are the estimated coefficients of equation (10).

Nine regression analyses were run to estimate accident occurence rate as a function of average traffic volume and the roadway length for each highway category. Hence, we obtained 9 different functions. The results of the regression analyses are given in table 5. We have decided to exclude fatality accident occurrence rate functions for freeway-expressway and interstate highway functional types from our analyses since the coefficents in these functions are not statistically significant (see table 5). We suggest that when more data on accidents become available, these occurrence rate functions be reestimated.

TABLE 5 Accident Occurrence Rate Regression Analyses Results

	Property damage		**Fatality**		**Injury**	
Arterial-Local-Collector	Coefficient	*t*-value	Coefficient	*t*-value	Coefficient	*t*-value
Intercept	4.7E-7	–3.9189995	4.15E-5	–2.8057275	5.95E-9	6.1740294
Q	2.1937	5.1888142	0.7357	1.7986337	2.5084	7.1900767
M	0.4592	1.3204427	0.8945	2.6580631	0.7366	2.5664179
R^2	0.65		0.42		0.80	
Freeway and Expressway	Coefficient	*t*-value	Coefficient	*t*-value	Coefficient	*t*-value
Intercept	0.0139	–1.7571000	7.7523	0.7731267	0.1851	0.9552417
Q	0.9508	3.9950085	–0.1435	–0.5537863	0.6837	3.9563777
M	0.2317	1.2375911	0.6553	3.2139405	0.501	3.6822804
R^2	0.63			0.43		0.76
Interstate	Coefficient	*t*-value	Coefficient	*t*-value	Coefficient	*t*-value
Intercept	3.23E-5	–1.5835942	7.79E-5	–0.9158328	3.114E-10	0.9158328
Q	1.0924	1.8072909	0.8066	0.8808309	2.0963	0.8808309
M	0.9043	4.7104977	0.3316	0.6027988	0.9766	0.6027988
R^2	0.72			0.11		0.81

C_f, C_h, and C_d values cover both direct and indirect costs caused by an incident, and their values are taken from Miller and Moffet (1993). These values are given in table 6.[9]

The total accident cost function is developed for each highway type, based on the results obtained from the regression analysis and the unit cost values for each accident type, as shown in table 6. The marginal accident cost functions are determined simply by taking the first order derivative of the total accident cost function with respect to volume, Q.

arterial-local-collector

$$MC_{accc} = (0.007)M^{0.4592}Q^{1.1937} + (125.58)M^{0.8945}Q^{-0.2643} + (0.0022)M^{0.7366}Q^{1.5084} \tag{11}$$

freeway-expressway

$$MC_{acc} = (89.81)M^{0.2317}Q^{-0.05} + (18{,}257.87)M^{0.501}Q^{-0.3163} \tag{12}$$

interstate

$$MC_{acc} = (0.2394)M^{0.9043}Q^{0.0924} + \left(9.42 \cdot 10^{-5}\right)M^{0.9766}Q^{1.0963} \tag{13}$$

where
M is roadway length (miles) and
Q is average traffic volume (vehicles/day).

It is reasonable to think that high volumes of traffic reduce vehicle-flow speed and, thus, fatal accidents. However, Vickrey (1968) argues that the rate of total accidents increases with increasing average daily traffic. As seen in table 5, our estimated accident occurance rate functions support Vickrey's (1968) hypothesis.

Infrastructure Costs

Infrastructure costs include all long-term expenditures, such as facility construction, material, labor, administration, and right of way costs. Also included are interest over the lifetime of the facility, regular maintenance expenditures for keeping the facility in a state of good repair, and occasional capital expenditures for traffic-flow improvement.

Highway investment and its costs can be best described by defining input prices, output, and network properties (Levinson et al. 1996). Input includes the cost of all phases of construction, such as roadway design, land acquisition, labor, construction material, and equipment. Network properties represents the physical capabilities of the constructed highway facility, which includes the number of lanes, lane width, pavement durability, the number of intersections, ramps, overpasses, and so forth. In addition, environmental factors are important elements in highway construction. Highway location, demographics of the district, soil properties, geometry of land, weather conditions, and other factors have an effect.

In computing marginal infrastructure cost, new construction and land-acquisition costs cancel out since these costs are not a function of traffic volume, Q. Thus, maintenance and improvement are the only cost category that remains in our marginal infrastructure cost function. We attempt to express the maintenance cost in terms of input and output. Input in this context includes all components of maintenance work, such as equipment usage, earthwork, grading, material, labor, and so forth. Output implies the traffic volume on the roadway. The data employed include all types of maintenance and improvement works completed between 1991 and 1998 in New Jersey. Given the database, we decided to divide the maintenance and improvement works into three categories:

1. Major reconstruction with/without roadway widening

TABLE 6 Accident Costs by Type

Accident type	Value per crash (dollars)
Fatality	4,113,956
Injury	144,291
Property damage	6,783

Note: The unit cost values are converted to year 2000 dollars, assuming an average 3.5% per year inflation rate.

Source: Miller and Moffet (1993)

[9] Property damage cost is assumed to comprise only the damage to the vehicle. Any type of injury falls in the personal injury category.

2. Roadway widening with/without resurfacing
3. Resurfacing with/without minor roadway widening

The estimated cost function for each category follows.

$$C_{M1} = (7.877)tm^{0.445}tev^{0.373}Q^{0.256} \tag{14}$$
$$R^2 = 0.81$$

$$C_{M2} = (19{,}930.37)s^{0.364}ns^{0.116} \tag{15}$$
$$R^2 = 0.61$$

$$C_{M3} = tm^{0.7556} + Q^{1.2931} \tag{16}$$

where
s is total volume of surface course used, ft^3
tm is total material used, ft^3 (surface course, base course, and sub-base)
ns is number of overhead signs installed,
tev is total earthwork, ft^3 (excavated soil or removed material, embankment, and so forth)
Q is traffic volume (vehicles/day).

Functions (14), (15), and (16) provide close estimates of the cost of each type of roadway maintenance project. However, the functions are to be used on a "per project" basis. In other words, to utilize these cost functions we need to know how often each type of maintenance work is undertaken, given the traffic conditions and pavement characteristics. There are valid methods of estimating resurfacing cycles (Small et al. 1989); however, we know of no practical methods for estimating the cycle of rest of the maintenance work categories. The first two maintenance work categories are required when the roadway is not capable of carrying the increased traffic. Its analysis requires a transportation demand model, which is out of the scope of this paper. Therefore, we have decided to exclude the first two cost categories from our marginal cost analysis.

Thus, the marginal cost function for resurfacing is calculated as follows:

$$MC_{inf} = 1.2931Q^{0.2931}\sum_{i=1}^{T}\frac{r}{1-e^{-rn_i}} \tag{17}$$

where
MC_{inf} is marginal maintenance cost in year 2000 (dollars/trip)
Q is traffic volume (vehicles/day)
T is number of resurfacing cycles throughout the lifetime of a pavement (25 years assumed)
n_i is time interval between each resurfacing dates and the year 2000 (years)
r is interest rate.

It should be noted that equation (17) considers all the resurfacing works done over the lifetime of the roadway. Ozbay et al. (2000) explains the method used to estimate the number of resurfacing cycles over the lifetime of a highway.

Environmental Costs

The environment costs caused by highway transportation regarded here include air pollution costs and noise costs. Development of specific cost functions for these environmental cost categories is beyond the scope of this study. Hence, we have adopted these cost functions from the germane literature.

Air Pollution Costs

Air pollution is defined as the change in ambient gas percentages and particulates resulting from human activities. Highway transportation accounts for a large portion of all air polluting activities through motor vehicle emissions. The contribution of highway transportation results either from the direct emission of these pollutants or from chemical reactions of these emitted pollutants with each other or with materials already existing in the atmosphere, such as PM_{10} (particulate matter 10 microns or smaller).

We consider the major pollutants emitted from motor vehicles to be volatile organic compounds (VOC), carbon monoxide (CO), and nitrogen oxides (NO_x). These pollutants have several adverse health effects on living organisms, land, crops, water, and air, among others. For a detailed description of these health effects, see Lynam and Pfeifer 1991.

In this study, we adopt an emission function to estimate the quantity of pollutant generated by motor vehicles. We put some factors aside, such as topographical and climatic conditions of the region, vehicle properties, vehicle speed, acceleration and deceleration, and fuel type. Next, unit cost values of each pollutant are calculated based on the methods presented in the literature. Unit cost calculations will be based on pollutant emission amounts in New Jersey, as reported by the Environmental Protection Agency (USEPA 1995).

Basically, we need to find the emission rate (grams/mile) for the pollutants *VOC, CO,* NO_x, and PM_{10}. Multiplying the emission rate with the total miles traveled in the considered network would give us the total amount emitted for each pollutant in New Jersey. The marginal cost function is developed simply by multiplying the unit cost values of each pollutant (dollars/gram) by the increase in the amount of pollutant emitted due to a unit increase in the traffic volume.

The proposed emission function is based on fuel consumption. It is assumed that the amount of pollutant released during motor vehicle operation is proportional with the amount of fuel consumed. The fuel-consumption function depends on the vehicle and is in quadratic form as follows (Ardekani et al. 1992).

$$F = 0.0723 - 0.00312V + 5.403 \cdot 10^{-5} V^2 \qquad (18)$$

where

F is fuel consumption at cruising speed (gallons/mile) and

V is average speed (miles/hour).

The emission rates of each pollutant (grams/gallon) are 69.9 grams for CO, 13.6 grams per NO_x, and 16.2 grams for VOC (SYNCHRO). Since we do not have a direct PM_{10} emission function, we utilize an emission rate specific to New Jersey (0.0825 grams/mile). See Ozbay et al. 2000 for the specific reasons why the unit grams/mile is chosen.

The unit cost of each pollutant is given in table 7. The fuel-consumption function given in equation (18) is a function of speed, V, and thus a function of traffic volume, Q. Total air pollution cost for a link of one mile, with a traffic volume Q (vehicles/hour) is calculated as follows.

$$C_{air} = Q(0.01094 + 0.2155F) \qquad (19)$$

$$MC_{air} = 0.01094 + 0.2155(F + Q\frac{\partial F}{\partial Q}) \qquad (20)$$

where

C_{air} and MC_{air} are measured in dollars per mile per hour, and F is calculated by equation (18).

TABLE 7 Cost of Each Pollutant Type

	VOC	NO_x	CO	PM_{10}
Unit morbidity cost per ton	$1,676	$3,039	N/A	$6,542
Unit mortality cost per ton	$2,779	$7,320	$15.21	$126,074
Total unit cost per ton	**$4,455**	**$10,349**	**$15.21**	**$132,616**

VOC = volatile organic compounds
NO_x = nitrogen oxides
CO = carbon monoxide
PM_{10} = particulate matter 10 microns or smaller
N/A = not applicable

Sources: Morbidity costs are directly taken from Small and Kazimi (1995). However, mortality costs are calculated specifically for New Jersey by following the same method used in Small and Kazimi (1995). See Ozbay et al. (2000) section 3.4.1 for a detailed explanation.

Noise Costs[10]

There are several methods used to define noise in a numerical range so that any noise source can be examined as it is heard by the human ear. In general, it is accepted that a sound becomes annoying after 50 dB(A) (A-weighted decibles). Any sound level above this limit definitely imposes a cost on society.

Social costs of noise are generally estimated by calculating the depreciation in the value of residential units alongside highways. The closer a house is to a highway, the higher these costs are. In this study, we use the Noise Depreciation Sensitivity Index (*NDSI*) as given in Nelson (1982). *NDSI* is defined as the ratio of the percentage reduction in the house value and the change in the noise level. Nelson (1982) suggests a value of 0.40% for the *NDSI*.

The house value depreciation function is defined as follows:

$$ND = N_b\left(L_{eq} - L_{max}\right)DW_{avg} \qquad (21)$$

$$L_{eq} = 10\log Q - 10\log r + 20\log V + 20 \qquad (22)$$

[10] "The same factors, which have brought us air pollution in crisis proportions, namely increasing population, urbanization, industrialization, technological change, and the usual relegation of environmental considerations to a position of secondary importance relative to economic ones, have also brought us the noise phenomenon" (Anthrop 1973).

where

ND is depreciation due to noise (dollars)

L_{eq} is defined as the Equivalent Sound Level (dB(A)). See Galloway et al. (1969)[11]

Q is traffic flow (vehicles/hour)

r is distance to the highway (feet)[12]

V is average speed of the traffic (miles per hour)

N_h is number of houses affected (number of houses per mile^2), calculated by multiplying the average residential density (RD, number of houses per mile^2) around a highway by the distance to that highway in feet (r) and the length of the relevant highway section in miles (d).[13]

$$N_h = 2(RD)rd \tag{23}$$

L_{max} is maximum acceptable noise level (50 dB(A) in this study)

D is percentage discount in value per an increase in the ambient noise level (0.4%)

W_{avg} is average house value (dollars), given in table 8.

Based on equations (21), (22), and (23), the noise cost function is developed as follows.

$$C_{noise} = 2\int_{r_1=50}^{r_2=r_{max}} \left(L_{eq} - 50\right) D W_{avg} \frac{RD}{5280} dr \tag{24}$$

where C_{noise} is the noise cost around a one-mile long roadway segment over so many years. Marginal cost is the first order derivative of equation (24) with respect to Q.

$$MC_{noise} = \frac{\partial C_{noise}}{\partial Q} = \frac{(RD)(r_2 - r_1)W_{avg}}{2{,}640}\left(\frac{10}{Q1n10} + \frac{20(\partial V \setminus \partial Q)}{V1n10}\right) \tag{25}$$

In this formulation, the total noise generated around a road segment is taken into account. Representing the maximum distance to highway, r_2 can be calculated by equating L_{eq} (equation (22)) to 50 dB(A), where the traffic noise is above dB(A).

[11] This function is only valid for the vehicle flows above 1,000 vehicles/hour.

[12] Minimum distance to a highway is assumed to be 50 feet.

[13] The multiplication by 2 in equation (23) is used to calculate the number of housing units on each side of the roadway.

TABLE 8 Housing Value in New Jersey

Value range	Dollars
Lower value quartile	158,410
Median value	228,940
Upper value quartile	317,385

RESULTS

For one-route marginal cost ($ORMC$) estimations, we selected one origin in each county in northern New Jersey. $ORMC$ values are calculated for the shortest routes between these selected origin-destination (O-D) pairs. In this process, we employed the marginal cost functions developed for each cost category presented. The generalized cost formula used in $ORMC$ calculations follows.[14]

$$ORMC_{r,s} = \sum_{i=1}^{k} FMC^i = \sum_{i=1}^{k} MC^i{}_{opr}d + MC^i{}_{cong} + MC^i{}_{acc} + MC^i{}_{inf} + MC^i{}_{air} + MC^i{}_{noise} \tag{26}$$

where

FMC is full marginal cost (dollars/mile)

MC_{opr} is marginal vehicle operating cost (dollars/trip)

MC_{cong} is marginal congestion cost (dollars/trip)

MC_{acc} is marginal accident cost (dollars/trip)

MC_{inf} is marginal infrastructure cost (dollars/trip)

MC_{air} is marginal air pollution cost (dollars/trip)

MC_{noise} is marginal noise cost (dollars/trip)

(r, s) is O-D pair

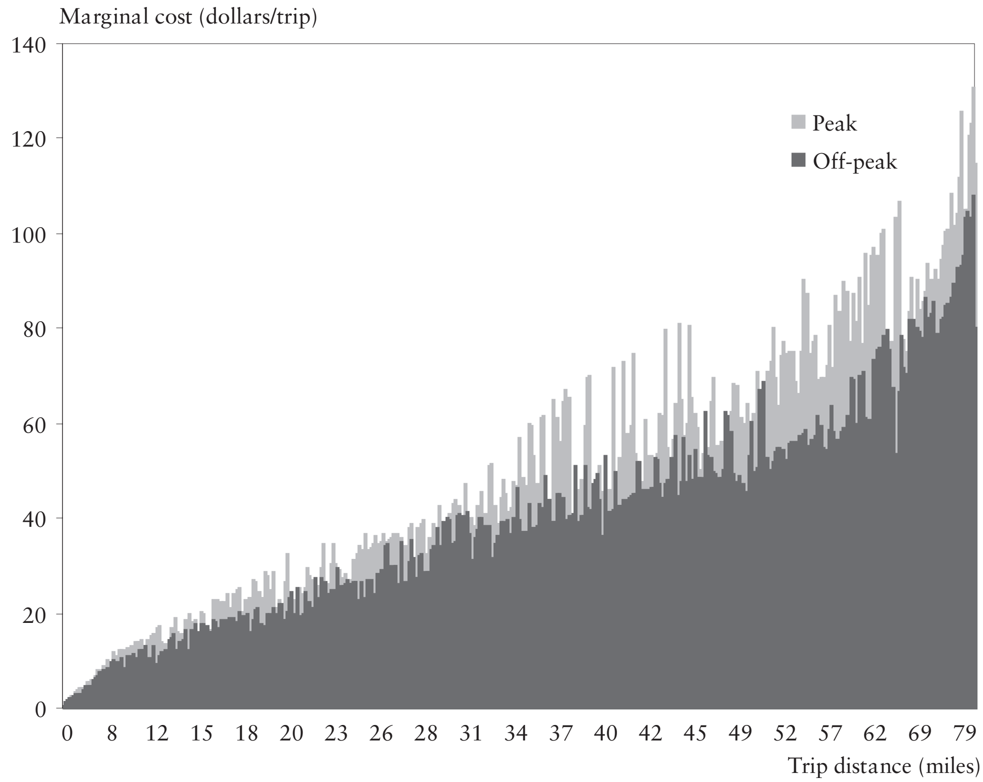
k is number of links between O-D pairs on the shortest route

d is trip distance (miles).

In total, we have 18,850 $ORMC$ values with their corresponding attributes. As mentioned, $ORMC$ values have a cost range based on the value of time (VOT) assumptions. We assumed a VOT range of 40 to 170% of the average hourly wage in New

[14] The units of noise and air pollution costs are given as dollars/trip here. However, it should be noted that in our analyses for each O-D pair in the network, respective units have been utilized according to trip characteristics.

FIGURE 4 *ORMC* **Distribution with Respect to Trip Distance for Peak and Off-Peak Hours (*VOT* = $7.6)**



Jersey. This enables us to better estimate full marginal cost under various time values. The analysis is also repeated for off-peak periods to observe the difference in the marginal cost values.

In figure 4, *ORMC* values are plotted with respect to trip distance for both peak and off-peak hours, assuming a *VOT* of $7.6/hour. As expected, peak-hour values are greater than off-peak-hour values, and the difference becomes significant as trip distance increases. Thus, the addition of longer trips due to urban sprawl can be expected to have increasingly higher impacts in terms of full marginal costs.

Figure 5 shows *ORMC* distribution with respect to trip distance when *VOT* is equal to $32.3, an assumed upper bound. It is clear that the difference in *ORMC* values for peak and off-peak hours are greater than those of figure 4. This result can be supported by the fact that congestion cost is more sensitive to *VOT* assumptions during peak hours than to *VOT* values at off-peak hours. Moreover, congestion costs appear to be the major driving component of overall costs. Thus, it is important to emphasize the effects of congestion-reduction measures in terms of overall costs.

Table 9 gives the *ORMC* functions with respect to trip distance (d) and time period for each *VOT* assumptions, estimated using the data points shown in figures 4 and 5. These functions can be used as a quick reference to the magnitude of *ORMC* values for given trip distances.

In order to observe the effect of highway functional type and the degree of urbanization on *ORMC* values, we need to hold trip distances as constant. We assume that for the same trip distance, the difference in *ORMC* values is attributed solely to highway functional type (interstate-freeway-expressway, principal arterial, minor arterial, and local-collector) and the degree of urbanization.

First, we analyze the effect of highway functional type on the *ORMC* value for a given trip distance. The analysis shows that the change in *ORMC* values with respect to highway functional type does not have a general pattern, irrespective of trip distances. Thus, we examine this relationship for different trip distance ranges. For relatively short distances (that is, 0 to 10 miles) the routes with a higher percentage of local-collector highways tend to have smaller *ORMC* values.

Figures 6 and 7 depict the effect of the percentage of local-collector highways of the shortest routes on *ORMC* values during peak and off-peak hours for a trip distance of two miles. During peak and off-peak hours, as the local-collector highway

FIGURE 5 ***ORMC*** **Distribution with Respect to Trip Distance for Peak and Off-Peak Hours (*VOT* = $32.3)**

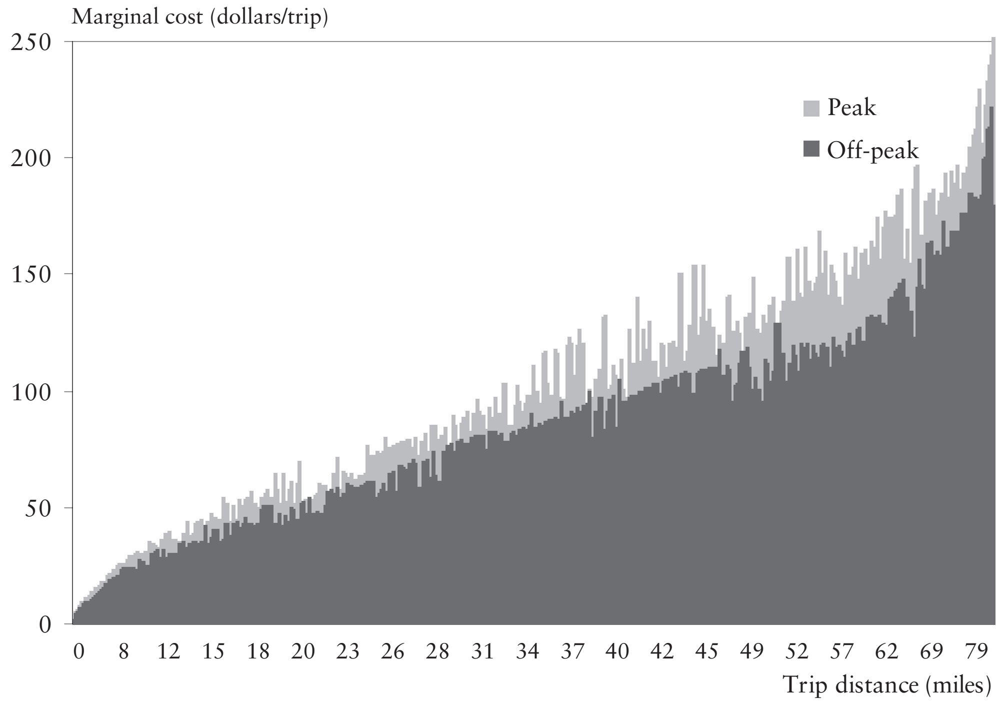


TABLE 9 Data Fitting for *ORMC* Values Given for Peak and Off-Peak Hours with Different VOT Values

	VOT = $7.6		*VOT* = $32.3	
	Equation	R^2	Equation	R^2
Peak	$y = 0.7568d$	0.857	$y = 1.846d$	0.923
Off-peak	$y = 0.6404d$	0.855	$y = 1.5421d$	0.907

Note: y is *ORMC* value (dollars/trip)
d is trip distance (miles)

type percentage increases, the *ORMC* value decreases. The same patterns obtained in figures 6 and 7 hold for trip distances up to 10 miles.

Figures 8 and 9 depict the variation of *ORMC* with respect to the percentage of minor arterial highway functional type for a trip distance of two miles. Unlike with local-collector highways, as the percentage of minor arterial highway of a route increases, the *ORMC* value increases as well. However, *ORMC* distribution with respect to the percentage of minor arterial road, as shown in figures 8 and 9, holds for trip distances up to three miles. For trip lengths between 3 and 10 miles, *ORMC* values tend to decrease as minor arterial percentage increases.

Since short trips do not generally use interstate-freeways-expressways, the effects of this highway functional category on *ORMC* distribution cannot be accurately analyzed for short-trip distances. As for principal arterials, sufficient information can be gathered for trip distances longer than three miles. Figure 10 shows that within the same trip distance range (3 to 10 miles), *ORMC* value increases with increasing principal arterial percentage for a given route.

ORMC distribution patterns change within the 0 to10 mile range because as trip distance increases, the percentages of each highway functional type changes as well. Up to three miles, the road types used are mainly local-collectors and minor arterials. It is obvious that local roads are more convenient than minor arterials for shorter trips. Above three miles, the utilization of principal arterials becomes significant, and *ORMC* value increases due to the increased congestion along these routes. Finally, beyond 10 miles, minor arterial and local-collector type of highways are not utilized as significantly as are interstate-freeways-expressways and principal arterials.

FIGURE 6 *ORMC* Distribution with Respect to Highway Functional Type Percentage During Peak Hours for Trip Distance of Two Miles (*VOT* = $7.6): Marginal Cost vs. Local-Collector Highway

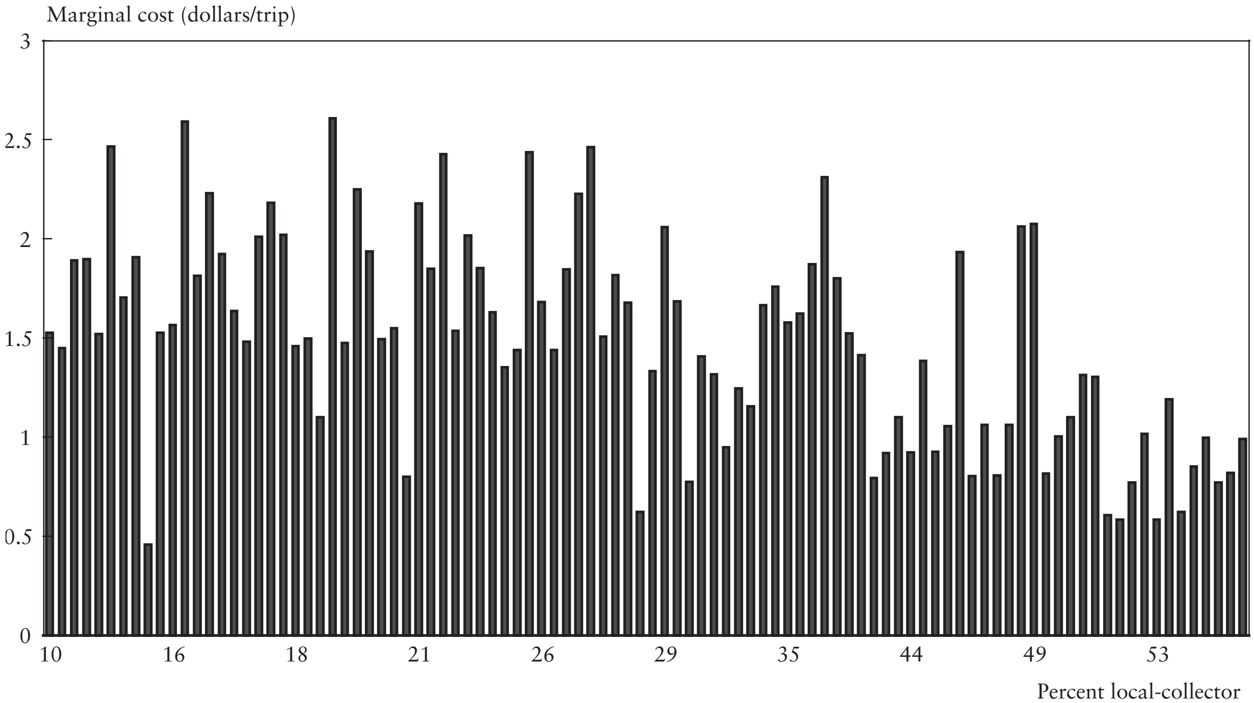


FIGURE 7 *ORMC* Distribution with Respect to Highway Functional Type Percentage During Off-Peak Hours for a Trip Distance of Two Miles (*VOT* = $7.6): Marginal Cost vs. Local-Collector Highway

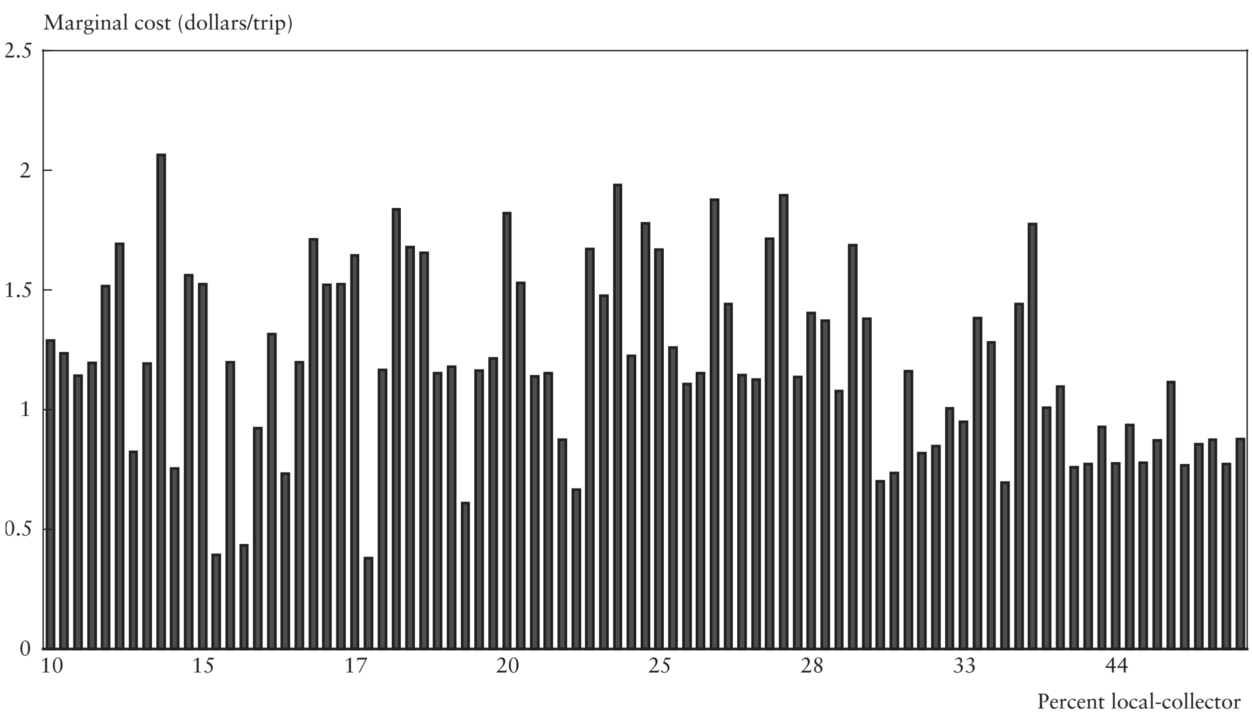


FIGURE 8 *ORMC* Distribution with Respect to Highway Functional Type Percentage During Peak Hours for a Trip Distance of Two Miles (*VOT* = \$7.6): Marginal Cost vs. Minor Arterial Highway

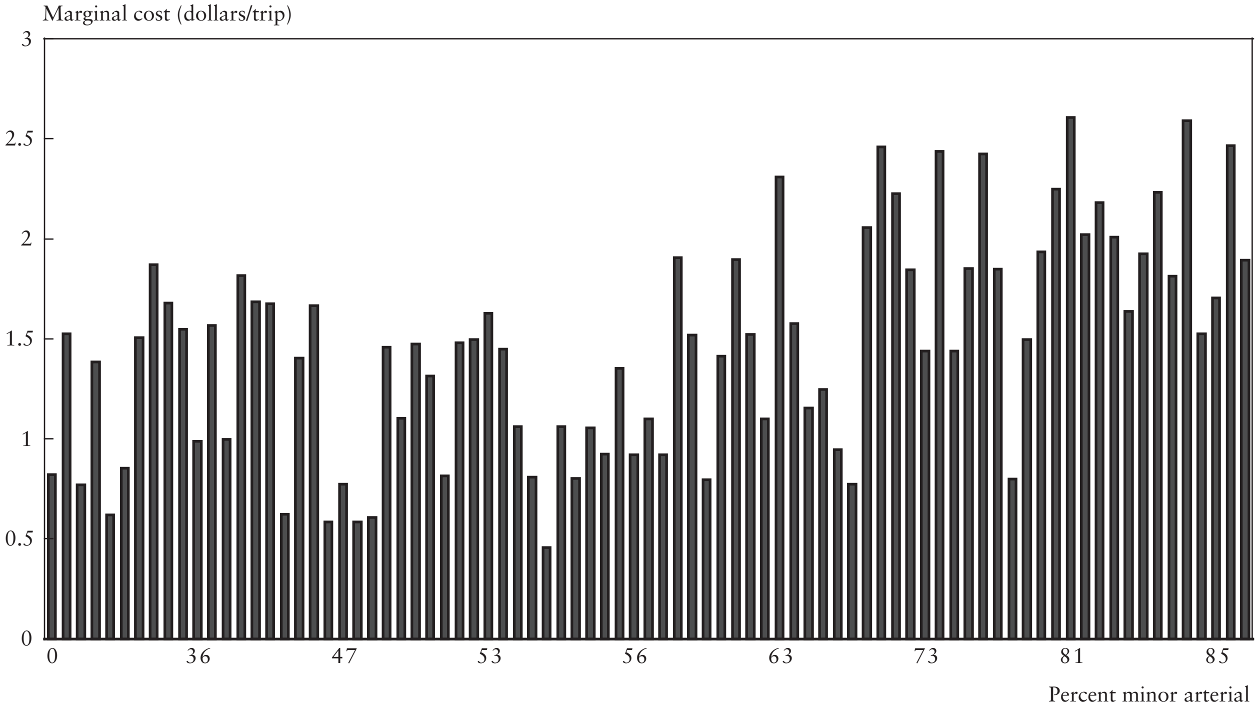


FIGURE 9 *ORMC* Distribution with Respect to Highway Functional Type Percentage During Off-Peak Hours for a Trip Distance of Two Miles (*VOT* = \$7.6): Marginal Cost vs. Minor Arterial Highway

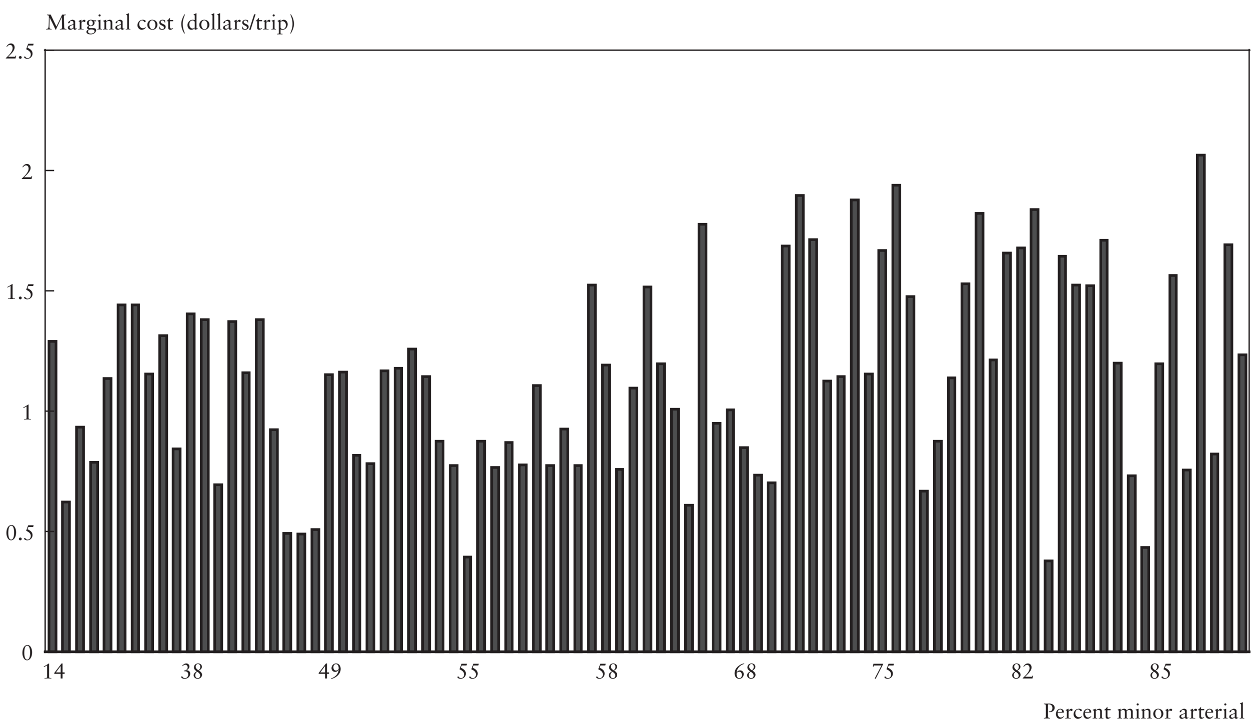


FIGURE 10 *ORMC* Distribution with Respect to Highway Functional Type Percentage During Peak Hours for a Trip Distance of Seven Miles (*VOT* = $7.6): Marginal Cost vs. Principal Arterial

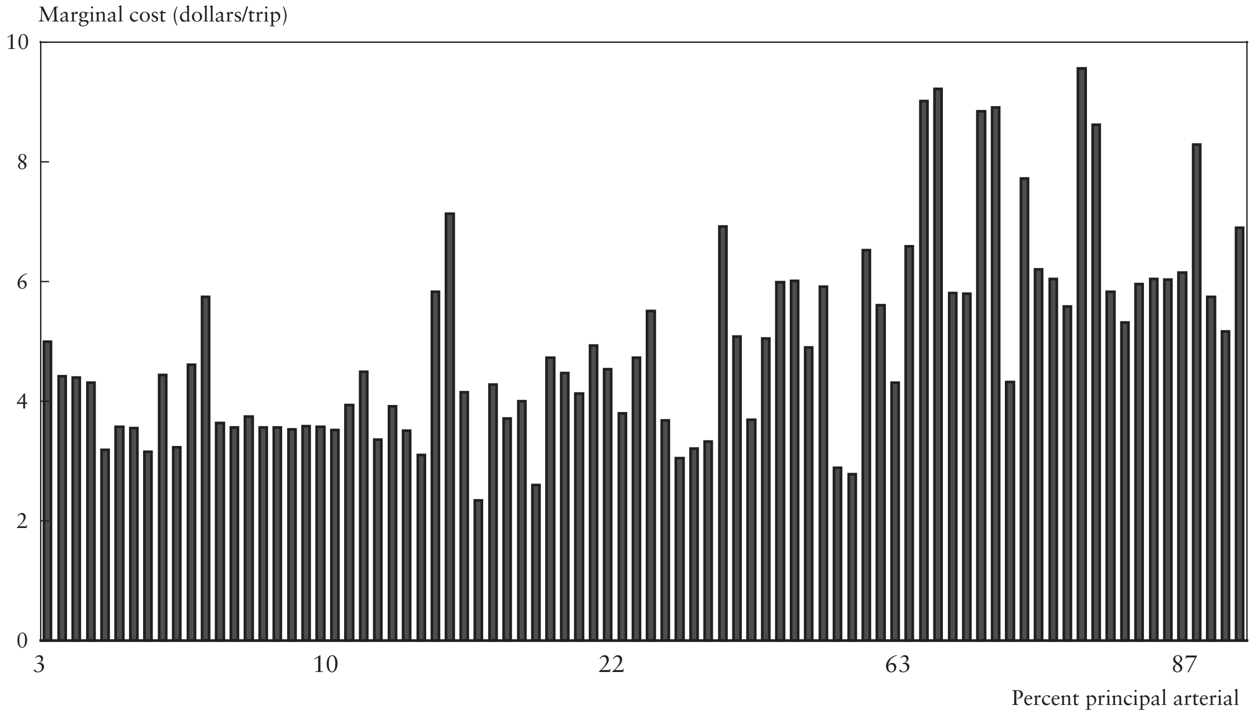


FIGURE 11 *ORMC* Distribution with Respect to Highway Functional Type Percentage During Peak Hours for a Trip Distance of 25 Miles (*VOT* = $7.6): Marginal Cost vs. Interstate-Freeway-Expressway

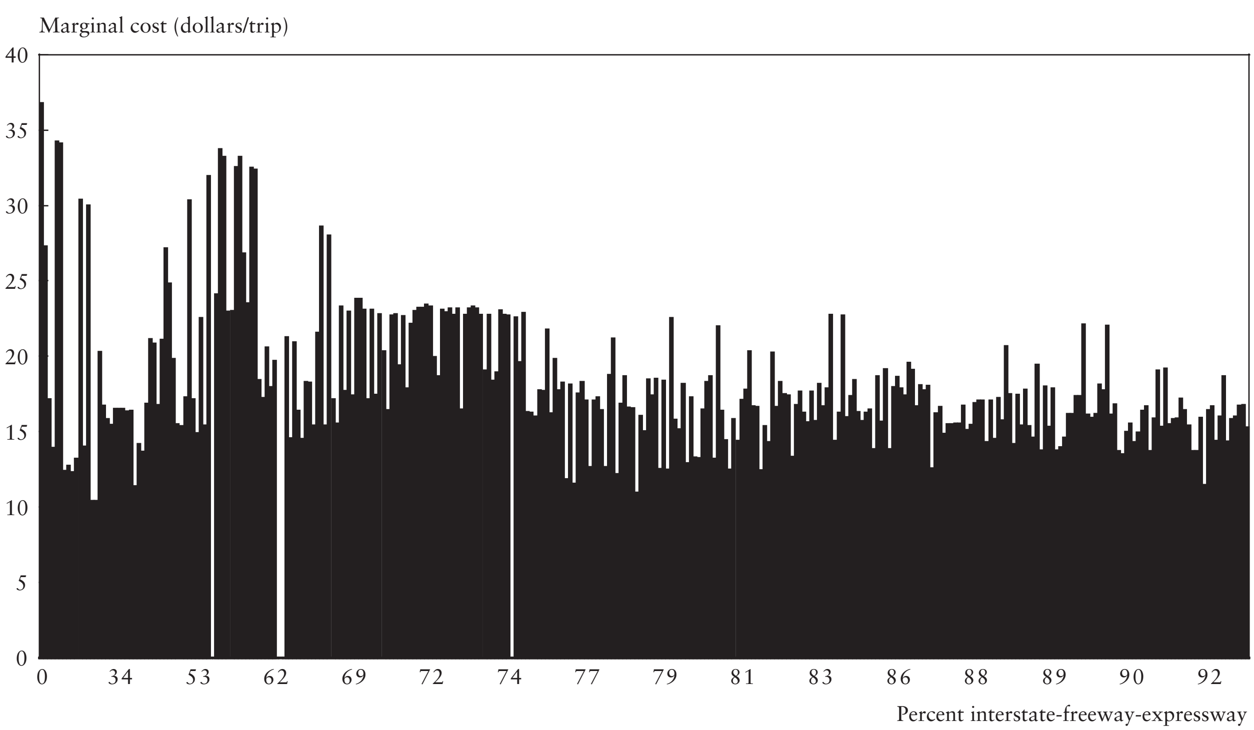
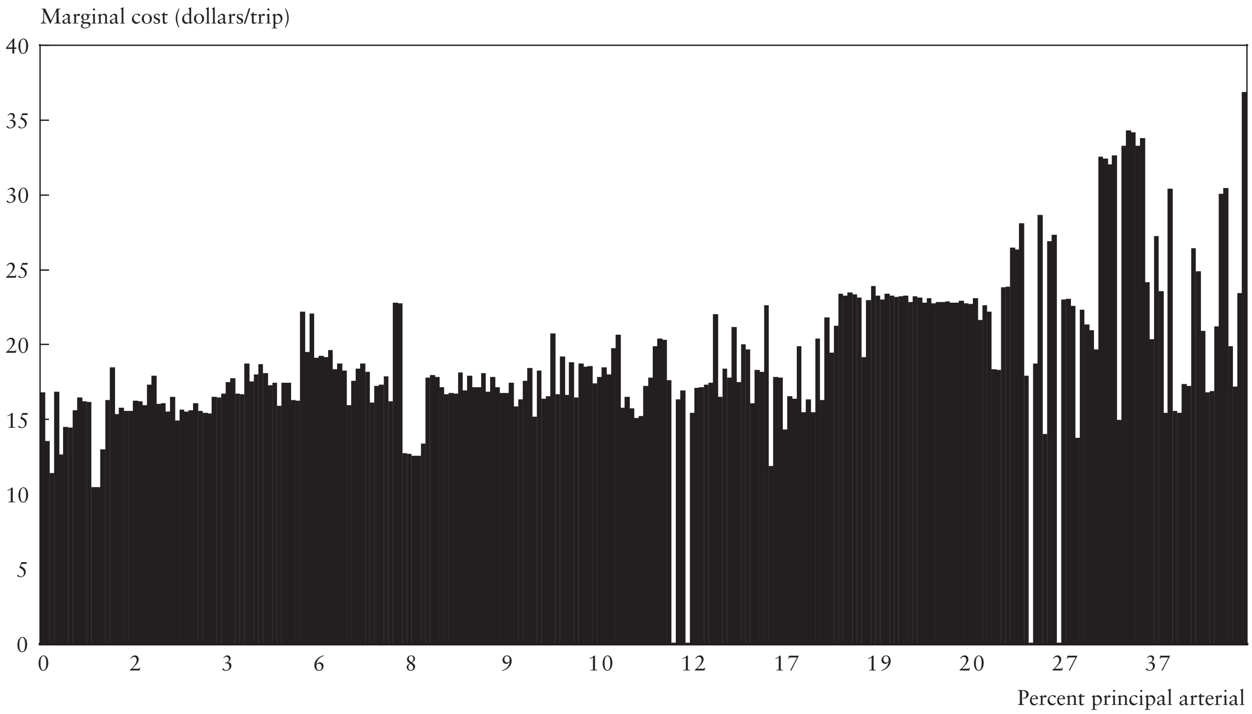


FIGURE 12 ***ORMC*** **Distribution with Respect to Highway Functional Type Percentage During Peak Hours for a Trip Distance of 25 Miles (*VOT* = $7.6): Marginal Cost vs. Principal Arterial Highway**



Next, we analyze *ORMC* distribution with respect to the percentage of highway functional type for longer trips distances. In this section, we only present the analysis performed for a trip distance of 25 miles. However, it should be noted that similar patterns are observed for all trip distances longer than 10 miles.

Figure 11 depicts the *ORMC* distribution with respect to interstate-freeway-expressway percentages for peak periods. *ORMC* values tend to decrease as interstate-freeway-expressway percentage increases. The same pattern holds during off-peak periods as well. Figure 12 depicts *ORMC* distribution with respect to percentage of the principal arterial type. It is seen that the pattern in figure 10 is valid for the 25-mile trip range. As the trip distance exceeds approximately 50 miles, the interstate-freeway-expressway functional type comprises most of the route distance. This fact restricts the analyses of *ORMC* distribution with respect to principal arterial as well as to the percentage of interstate-freeway-expressway.

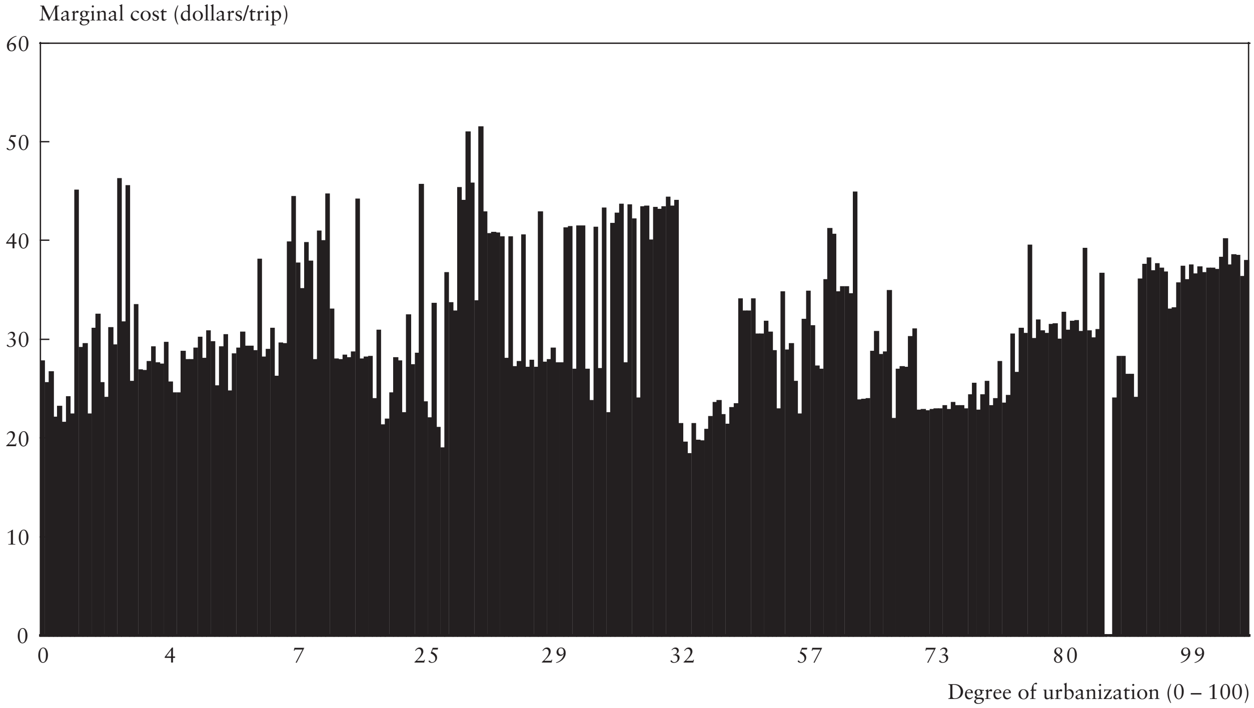
Finally, we attempt to correlate the variation in *ORMC* values and degree of urbanization using the data generated. Figure 13 shows the *ORMC* variation with respect to percentage of urbanization over a given trip distance. Similar analyses are done for all the trip distance ranges both for peak and off-peak periods. However, *ORMC* variations with respect to degree of urbanization do not follow a typical pattern. Thus, we can conclude that the degree of urbanization around highways does not necessarily imply an increased congestion level.

EVALUATION OF THE CURRENT PRICING POLICY

In the ongoing efforts to reduce congestion through the use of congestion tolls, knowing the full marginal cost of highway transportation can be vitally important. Leaving aside the practical difficulties and political complexities of this concept, we evaluate the efficiency of the current practice of collecting highway user fees in New Jersey relative to the results obtained above.

As stated in section 2, highway marginal cost pricing requires that every user be held responsible for the cost he or she imposes on the rest of the traffic with his or her additional trips. Hence, in theory, user fee per trip should be equal to the external cost of a trip (Small 1992). Therefore, if we compare the value of the actual user fees per trip currently imposed in New Jersey with our estimate external-

FIGURE 13 *ORMC* Distribution with Respect to Degree of Urbanization During Peak Hours for a Trip Distance of 40 Miles (*VOT* = $7.6): Marginal Cost vs. Degree of Urbanization



ities through the *FMC* methodology, we can measure the effectiveness of highway pricing policies in New Jersey.

Although average congestion cost and vehicle operating costs are fully experienced by the users, infrastructure and maintenance costs are paid through fuel and vehicle registration and other taxes. Hence, we need to determine if the user fees collected by the government are sufficient to cover the "external" costs of highway transportation, such as increased travel time, pollution, and accidents. It is known that a certain portion of congestion and accident costs are external, meaning that that portion is directly imposed on the rest of the traffic by an additional trip. In our analysis, we have calculated congestion externalities. As for accident externalities, we have adopted a ratio of marginal to average accident cost of 1.52 in our analyses (Newberry 1988). Finally, we consider air pollution and noise costs as external costs to the rest of the traffic and society.

However, the detailed analysis of this task is not straightforward because trips have several quantitative and qualitative measures that cannot be grouped together easily. Consider, for example, the difference between a 50-mile trip and a 3-mile trip, or 2 trips with the same distance but on different highway types. Due to these differences, there is not a unique value for *FMC* per trip. For our analysis here, we have used the average of all *ORMC* values within a trip distance ranging from 10 to 15 miles[15] and then weighted the averages for peak and off-peak hours. The average *FMC* values by each cost category are presented in table 10. It should be noted that the contributions of each cost category to *FMC* as shown in table 10 are not unique for all trip distance ranges; however, we believe that table 10 provides a good idea of each cost category's contributions.

Using the air pollution, noise costs, congestion externalities, and a ratio of marginal to average accident cost of 1.52 for accident externalities, we calculate the external cost of making a trip within a distance range of 10 to 15 miles as $1.252.

We now need to find out if the cost imposed by the government is equal to our FMC estimates. FHWA reports that an amount of $2,703,741,000

[15] The "Summary of Travel Trends" (USDOT FHWA 1997) reports an annual national average vehicle trip length of 9.06 miles. A value specific to New Jersey is not available. Thus, we have chosen a range of 9 to 15 mile trip length.

TABLE 10 Full Marginal Cost by Categories for a Trip Distance Range of 9 to 15 Miles

Operating cost (dollars)	Congestion cost (dollars)	Congestion externality	Accident* cost (dollars)	Infrastructure cost (dollars)	Air pollution cost (dollars)	Noise cost (dollars)
1.389	3.786	0.635	1.009	0.062	0.114	0.158

* Accident externality = marginal accident cost. Average accident cost = 1.009 – (1.009/1.52) = 0.345

for New Jersey was collected through federal and state fuel and vehicle tax, state and local tolls in 1998 as highway user revenues (USDOT FHWA 1999). Dividing this amount by the annual total number of trips taken in New Jersey in 1998 (6.31 billion), we get an estimate of the cost of a trip in New Jersey as $0.428 (NJDOT 2000).[16] This is the average amount that the government charges each user per trip. Comparing this amount with our FMC, we observe that it is less than what we regard as necessary to compensate for the full marginal cost per trip.

What, then, should be the correct amount of increase in user fees imposed on users to compensate for the marginal roadway pricing in New Jersey? Let us assume that $1.252 is the user fee per trip that the state government targets. Let us also assume that the state government decides to collect the deficit in user fees only through a state fuel tax. The annual user revenue that should be collected becomes $1.252 × 6.4386 billion trips (see footnote 16), equal to $8,061,127,200. Assuming the federal vehicle and fuel tax revenues ($962,433,000) and state and local tolls revenues ($619,862,000) remain the same, the dollar amount the state needs to collect is now $8,061,127,200 – ($962,433,000 + $619,862,000), or $6,478,832,200. This is the amount that needs to be raised by state vehicle and fuel taxes. FHWA reported that vehicle tax collected in 1999 was $631,506,000. Hence, $6,317,609,656 minus $631,506,000 equal to $5,847,326,200 would be the total amount that the state government needs to collect by state fuel tax only (USDOT FHWA 1999). Dividing this amount by the taxable amount of fuel consumed in New Jersey in 1999 (4,688,147,000 gallons) would be equivalent to the new additional state fuel tax, which comes out to $1.247 per gallon

TABLE 11 Fuel Prices and Percent Taxes in European Countries

Country	Percent taxes[a]	Tax	Price per gallon[b]
United Kingdom	76.8	3.295	$4.29
Netherlands	68.4	2.708	$3.96
France	72.7	2.661	$3.66
Italy	67.7	2.464	$3.64
Germany	70.7	2.418	$3.42
USA	24.1	0.419	$1.74
USA (Recommended)	47.3	1.563[c]	$3.303

[a]Gas tax as a percentage of retail price of gallon of gas
[b]Retail price per gallon of premium leaded as of September 2000
[c]Tax amount includes federal tax plus our recommended state fuel tax, $1.247, instead of the current state fuel tax of $0.1038 per gallon

Source: International Energy Source, National Energy Information Center, Quoted in *The Detroit News* 9/20/2000.

(USDOT FHWA 1999). This additional amount is far more than the current state fuel tax of $0.1038 per gallon.[17]

Kulash (2001) states, "There are valid social and economic reasons why road users should pay for the full range of costs that they impose on the public, but they pose a social and economic shock as well." Thus, although the collection of this revenue through the gas tax is not an impossible task, given the fact that compared to European countries this amount in the United States is considerably less, it does not appear to be an easy policy to sell to the American people, given the historical realities of this country. Table 11 presents the fuel tax charged in different countries in Europe as a percentage of the fuel price. As seen, the current fuel tax in the

[16] The number of trips reported in 2000 is reduced to 1998 values, assuming a 2% increase per year in the total number of trips.

[17] This value is the weighted average of all fuel tax rates based on the taxable amount in 1999.

United States is far less than in European countries. Even our estimated fuel tax percentage is less than the values in effect in European countries.

As we mentioned, the objective of marginal roadway pricing is to reduce congestion by charging users the additional amount that they impose on others. The concept has valid economical reasons on how to achieve optimal pricing. However, from our results it is clear that the concept presents serious practical difficulties regarding its political consequences. Though our calculation methodology is straightforward and based on averages, it demonstrates the extent of the difficulty of this problem. Here, our first scenario was to increase only the fuel tax. There are other means to collect the same of amount of revenue to compensate for the full marginal cost per trip, such as tolls. The introduction of advanced technology such as automatic vehicle identification systems can serve well for this specific purpose of imposing trip-based charges. On the other hand, it is clear that increased gas taxes and tolls will reduce the demand for highways, which will reduce the external costs such as congestion, pollution, and others. This will reduce the *FMC* per trip and the amount of taxes and tolls needed to a more acceptable level. However, this kind of dynamic analysis of the change of demand as a result of pricing is beyond the scope of this paper.

CONCLUSIONS

In this study, a new methodology for estimating network-wide full marginal costs is presented. This methodology is applied to determine the full marginal cost of highway transportation in northern New Jersey. The variation in marginal cost value due to trip distance, degree of urbanization, and highway functional type are analyzed. Each set of observations is made for different VOT assumptions and time periods (peak and off-peak hours). Our main conclusions follow.

1. The difference in the marginal cost value for peak and off-peak hours becomes more significant with longer trip distances due to the increase in congestion costs.
2. It is estimated that marginal costs decline as a percentage of trip distance performed on freeway and expressway-type facilities increased.
3. It is observed that along the routes that have a higher percentage of principal arterials, marginal costs tend to increase.
4. Urbanization around the highways has no significant effect on marginal costs.
5. We also used our full marginal cost findings to evaluate the current pricing policy. It is observed that the government's highway user revenue is far below the amount required to meet the marginal roadway-pricing criterion.

These results can be used by policymakers to assess the effectiveness of the overall transportation system. For example, the finding that longer trips have considerably higher costs, mainly due to congestion, can be used to develop proper congestion toll schemes. In general, the evaluation of a decision of whether or not to invest in a new facility can be facilitated by comparing marginal social benefits with the germane marginal social costs, given a specific location.

It should be noted, however, that the results presented in this study are specific to a New Jersey area. Furthermore, the marginal cost values reported here are sensitive to other assumptions not included in this study. For example, the travel time function used to calculate congestion costs could affect marginal cost values significantly. In this study, the Bureau of Public Road's (BPR) travel time function is utilized. The variation in the cost values can be observed using different travel time functions.

ACKNOWLEDGMENTS

This work has been supported by funding provided by New Jersey Department of Transportation. The authors would like to thank Dr. Abbas Hirya for his support, constructive comments, and suggestions throughout this study. The authors would also like to thank two anonymous reviewers for their excellent and constructive comments and suggestions.

REFERENCES

American Automobile Manufacturers Association (AAMA). 1996. *Motor Vehicle 1996 Facts and Figures*. Washington, DC.

Anthrop, D.P. 1973. *Noise Pollution.* Lexington, MA: D.C. Heath and Company.

Ardekani, S., E. Hauer, and B. Jamei. 1992. *Traffic Impact Models in Traffic Flow Theory.* Oak Ridge National Laboratory Report, available at www.ornl.gov.

Churchill, A. 1972. *Road User Charges in Central America.* World Bank Staff Occasional Papers, Johns Hopkins University, Baltimore, MD.

Cipriani, R., M.J. Porter, N. Conroy, L. Johnson, and K. Semple. 1998. The Full Costs of Transportation in the Central Puget Sound Region in 1995. TRB Preprint: 980670. Transportation Research Board 77th Annual Meeting, Washington, DC.

Galloway, W.J., E.W. Clark, S.J. Kerrick, B. Beranek, and N.V. Nuys. 1969. *Highway Noise: Measurement, Simulation, and Mixed Reactions.* National Cooperative Highway Research Program Report 78. National Cooperative Highway Research Program, Washington, DC.

Jara-Diaz, S.R., P.P. Donoso, and J.A. Araneda. 1992. Estimation of Marginal Transport Costs: The Flow Aggregation Function Approach. *Journal of Transport Economics and Policy* 26, no. 1:35–48.

Kelley Blue Book. 2000. New and Used Car Database. Available at www.kbb.com.

Kulash, D.J. 2001. Transportation User Fees in the United States. *Transportation Quarterly* 55, no. 3:33–49.

Land Transport Authority (LTA). 1998. Available at www.gov.sg/lta.

Levinson, D.M. and D. Gillen. 1998. The Full Cost of Intercity Highway Transportation. TRB Preprint: 980263. Transportation Research Board 77th Annual Meeting, Washington, DC.

Levinson, D., D. Gillen, A. Kanafani, and J.M. Mathieu. 1996. *The Full Cost of Intercity Transportation—A Comparison of High-Speed Rail, Air, and Highway Transportation in California.* UCB-IS-RR-96-3, Berkeley, California, Institute of Transportation Studies, University of California at Berkeley.

Lynam, D.R. and G.D. Pfeifer. 1991. Human Health Effects and Highway-Related Pollutants. *Highway Pollution.* Amsterdam: Elsevier.

Mayeres, I., S. Ochelen, and S. Proost. 1996. The Marginal External Costs of Urban Transport. *Transportation Research* 1, no. 2:111–30.

Meier, G.M. 1983. *Pricing Policy for Development Management.* Baltimore: John Hopkins University Press.

Miller, P. and J. Moffet. 1993. *The Price of Mobility: Uncovering the Hidden Costs of Transportation.* New York, NY: Natural Resources Defense Council.

Mohring, H. 1976. *Transportation Economics.* Cambridge: Ballinger Company.

Nelson, J.P. 1982. Highway Noise and Property Values: A Survey of Recent Evidence. *Journal of Transport Economics and Policy,* pp. 117–38.

Newberry, D.M. 1988. Road User Charges in Britain. *Economic Journal* 98.

New Jersey Department of Labor (NJDOL). 2000. Occupational Employment Statistics Wage Survey. Available at ww.njpin.com/OneStopCareerCenter/LaborMarketInformation//lmi23/.

New Jersey Department of Transportation (NJDOT). 2000. New Jersey Transportation Fact Book. Available at http://www.state.nj.us/transportation/publicat/Facts/.

Ozbay, K., B.O. Bartin, and J. Berechman. 2000. *Full Costs of Highway Transportation in New Jersey.* NJDOT Ongoing Project Report.

Pape, U. 1974. Implementation and Efficiency of Moore Algorithms for the Shortest Route Problem. *Mathematical Programming* 7, 2:212–22.

Peat Marwick Stevenson & Kellog (PMSK). 1993. *The Cost of Transporting People in the British Columbia Mainland.* Transport 2021 Technical Report 11. Ministry of Transportation and Highways, Victoria, BC, Canada.

Small, K. 1992. *Urban Transportation Economics.* Newark, NJ: Harwood Academic Publishers GmbH.

Small, K. and C. Kazimi. 1995. On Costs of Air Pollution from Motor Vehicles. *Journal of Transport Economics and Policy* 29:7–32.

Small, K., C. Winston, and C.A. Evans. 1989. Road Work: A New Highway Pricing and Investment Policy. The Brookings Institution, Washington, DC.

SYNCHRO Traffic Signal Simulation Software User Manual. Version 5. Albany, CA: Trafficware.

Tellis, R. and J.C. Khisty. 1995. Assessment of the Actual Cost of Automobile Transport. *7th WCTR Proceedings.*

Timmer, P. 1987. *Institutional Development: Incentives to Performance.* Baltimore: Johns Hopkins University Press.

Transportation Research Board (TRB). 1996. *Estimating Marginal Costs of Freight Transportation.* Transportation Research Board, Special Report 246.

U.S. Environmental Protection Agency (USEPA). 1995. National Air Pollution Emission Trends. Available at http://www/epa.gov/ttn/chief/trends/trend98.

U.S. Department of Transportation (USDOT), Federal Highway Administration (FHWA). 1991. *Cost of Owning and Operating Automobiles, Vans, and Light Trucks.* Washington, DC.

____. 1997. Summary of Travel Trends: 1995 Nationwide Personal Transportation Survey, prepared by Oak Ridge National Laboratory.

____. 1999. *Highway Statistics Summary.* Washington, DC.

Vickrey, W. 1968. Automobile Accidents, Tort Law, Externalities, and Insurance: An Economists Critique. *Law and Contemporary Problems* 33:464–87.

Walters, A.A. 1968. *The Economics of Road User Charges.* Baltimore: John Hopkins University Press.

JOURNAL OF TRANSPORTATION AND STATISTICS

Guidelines for Manuscript Submission

Send all submission materials to

Marsha Fenn, Managing Editor
Journal of Transportation and Statistics
Bureau of Transportation Statistics
US Department of Transportation
400 7th Street SW, Room 3430
Washington, DC 20590
Email journal@bts.gov

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.

The cover page should include the title of the manuscript, author name(s), and the telephone number and surface and email addresses of the contact author. The second page should contain a 100-word abstract briefly describing the contents of the paper, the research method, and the results.

Graphic elements (figures and tables) must be called out in the text. Graphic elements must be in black ink although exceptions may be made in certain cases.

References follow the style outlined in the *Chicago Manual of Style*. All non-original material must be sourced.

Accepted papers must be submitted electronically in addition to a hard copy. Make sure the hard copy corresponds to the electronic version. Authors will be required to return article page proofs to the Managing Editor within 48 hours of their receipt.

Acceptable software for text and equations is limited to Word, WordPerfect, and LaTeX. Data behind all figures, maps, and charts must be provided in Excel, Word, or DeltaGraph. Data in 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 GIS programs. Please save your GIS files using the following extensions: .eps or .pdf. If this is not possible, maps may be provided in ArcView3.x project files (.apr), ArcGIS 8.x map files (.mxd), or Arc/Info metafiles (.map). Files using the .apr or .mxd extensions can be supplied on a CD-ROM.

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 4 Number 1
April 2001
ISSN 1094-8848

CONTENTS