TRANSPORTATION RESEARCH RECORD

No. 1467

Safety and Human Performance

Traffic and Roadway Accident Analysis and Traffic Records Research

A peer-reviewed publication of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY PRESS
WASHINGTON, D.C. 1994

Transportation Research Record 1467
ISSN 0361-1981
ISBN 0-309-06073-7
Price: $20.00

Subscriber Category
IVB safety and human performance

Printed in the United States of America

Sponsorship of Transportation Research Record 1467

GROUP 3—OPERATION, SAFETY, AND MAINTENANCE OF TRANSPORTATION FACILITIES
Chairman: Jerome W. Hall, University of New Mexico

Facilities and Operations Section
Chairman: Jack L. Kay, JHK & Associates

Committee on Methodology for Evaluating Highway Improvements
Chairman: Ezra Hauer, University of Toronto
Secretary: Warren E. Hughes, Bellomo-McGee, Inc.
Karen K. Ajluni, Brian L. Bowman, Charles Philip Brinkman, Julie Anna Cirillo, Barbara Hilger DeLucia, William D. Glauz, Alfred-Shalom Hakkert, King K. Mak, Thomas L. Maleck, Ernst Meyer, Shaw-Pin Miaou, Olga Pendleton, Bhagwant N. Persaud, Ramey O. Rogness, James M. Witkowski

Users and Vehicles Section
Chairman: Conrad L. Dudek, Texas A&M University

Committee on Traffic Records and Accident Analysis
Chairman: Paul P. Jovanis, University of California–Davis
Secretary: Jeffrey F. Paniati, Federal Highway Administration
Sherry S. Borener, David J. Bozak, Thomas E. Bryer, Myung-Soon Chang, Benjamin V. Chatfield, Ted Chira-Chavala, Barbara Hilger DeLucia, Mark Lee Edwards, C. Arthur Geurts, Alan F. Hoskin, Patricia S. Hu, John J. Lawson, Joseph C. Marsh IV, Judson S. Matthias, Ronald C. Pfefer, Carol Lederhaus Popkin, Howard S. Stein, John G. Viner, Phyllis E. Young

Task Force on Statistical Methods in Transportation
Chairman: Olga Pendleton, Texas A&M University
Secretary: Patricia S. Hu, Oak Ridge National Laboratory
Charles E. Antle, Kenneth L. Campbell, Lynda T. Carlson, Michael I. Darter, Gary A. Davis, Michael S. Griffith, Irit Hocherman, David T. Hunt, Robert L. Mason, Edward V. Milton, Frank C. Sonsini, Gary S. Spring, Delbert E. Stewart, Peter R. Stopher, Russell G. Thompson, Richard M. Weed, Robert E. White

Transportation Research Board Staff
Robert E. Spicher, Director, Technical Activities
Richard F. Pain, Transportation Safety Coordinator
Richard A. Cunard, Engineer of Traffic and Operations
Nancy A. Ackerman, Director, Reports and Editorial Services
Marianna Rigamer, Oversight Editor

Sponsorship is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 1993.

Transportation Research Record 1467

Contents

Foreword

Crash data are a central source of information for a wide variety of roadway and traffic safety activities, ranging from policy to planning and design to countermeasures to evaluation. The papers in this volume are equally far ranging in dealing with traffic records and accident analysis. Modeling of carrier accident risk is done by Lin et al. Kim et al. explore crash type and injury in Hawaii. Zegeer et al. delve into a little-studied area: commercial bus accidents and the related roadway conditions. The next three papers deal with roadway hardware and safety. Bélanger develops a method for estimating the safety of four-legged unsignalized intersections. Hauer et al. study the effects of two types of road resurfacing programs on subsequent crashes. In this study innovative statistical techniques that will be of use in other types of safety evaluations are used. The extensive data currently being collected with WIM devices are explored for uses in various safety related analyses (e.g., exposure data) by Hajek et al. McCarthy and Madanat illustrate the applicability of recently developed econometric methods in highway safety analysis. These applications can improve traffic accident model accuracy with positive benefits on safety policy and investment decisions. In the final paper Kim and Nitz discuss the use of automated software for linking records in traffic records analysis.

Time of Day Models of Motor Carrier Accident Risk

TZUOO-DING LIN, PAUL P. JOVANIS, AND CHUN-ZIN YANG

A time-dependent logistic regression model has been formulated to assess the safety of motor carrier operations. The model estimates the probability of having an accident at time interval t, subject to surviving (i.e., not having an accident) until that time. Three logistic regression models are estimated, which include time main effects (the driving time), time-independent effects (experience), time-dependent effects (time of day), and a series of time-related interactions. Driving time has the strongest direct effect on accident risk. The first 4 hr consistently have the lowest accident risk and are indistinguishable from each other. Accident risk increases significantly after the 4th hr, by approximately 50 percent or more, until the 7th hr. The 8th and 9th hr show a further increase, approximately 80 and 130 percent higher than the first 4 hr. Drivers with more than 10 years of driving experience retain a consistently low accident risk; all other categories of driving experience have a significantly higher risk. Daytime driving, particularly at the noon time (10:00 a.m. to 12:00 noon), results in a significantly lower risk of an accident. Drivers at one time of day (4:00 to 6:00 p.m.) have an accident risk about 60 percent higher than those driving during the baseline; drivers during the other three significant times of day also experience accident risks about 40 percent higher than drivers during the baseline. All three times of day involve night or dawn driving; two are associated with circadian rhythms. Rest breaks, particularly those taken before the 6th or 7th hr of driving, appear to lower accident risk significantly for many times of day.

Motor carrier safety has been an area of active study throughout the 1980s and the early 1990s. Of the factors generally considered in safety studies (i.e., driver, vehicle, roadway, and environment), particular attention has been paid to driver-related factors. One major study concluded that 65 percent of accidents may be attributable to human errors (*1*).

Driving fatigue is believed to have a particularly powerful effect on commercial vehicle drivers, representing one of the primary human factors. Fatigue significantly increases driving errors and decreases driver alertness. Two additional studies using restricted data bases have found more than 30 percent of heavy truck crashes may result from driving fatigue (*2,3*). Nevertheless, fatigue is a sufficiently vague concept in that it has not been precisely defined and measured (*4*), a fact that presents difficulties in applying fatigue concepts in accident models. Several studies have described factors associated with either physiological or psychological components of fatigue (*4–7*).

Driving hours, for one origin-to-destination trip or over several trips and multiple days, is often an important element of fatigue. Several studies have considered the appropriateness of government-regulated limits on driving hours. These studies seek to identify hours that pose higher accident risk and policy changes that could result in reduced accident risk (*8–14*). Although it may seem straightforward to account for the influence of driving hours on fatigue, there are many subtleties to be considered. Among those already studied in the literature are the effects of the following: off-duty hours immediately before a trip and multiday driving (*13,14*); heat, noise, and vibration (*15*); cargo loading and unloading (*16*); patterns of rest in sleeper berths (*17*); and alcohol and drugs (*3,18*).

This research attempts to contribute to this literature by identifying the effects of time of day on accident occurrence. Circadian rhythms, which are changes in body function following an approximate 24-hr cycle, are of particular importance. Although circadian rhythms vary somewhat from person to person, the most common pattern is one with a physiological low around 4:00 to 6:00 a.m., representing a time of particular risk to drivers. This represents a substantial societal risk as a significant amount of truck travel occurs at night.

Several relevant studies have focused on the relationship between motor carrier accidents and time of day. Harris and Mackie (*8*) concluded that the lowest levels of alertness occur for most drivers between midnight and 8:00 a.m. Several additional studies also have found elevated involvement or accident risk in this same time interval, suggesting a circadian effect (*9,17,19,20*). Interestingly, Hamelin (*10*) indicated that accident involvement rates generally increase throughout the day from a low point around 4:00 to 6:00 a.m. to a high point from midnight to 2:00 a.m. There is also a sharp peak in risk around noon. Another study of automobile drivers found that an additional period of decreased alertness occurs in the mid-afternoon (*21*). This research aims at a more explicit quantification of the effect of time of day on motor carrier accident risk.

OBJECTIVES

There is a need to develop quantitative methods to analyze the effect of time of day on accident risk. In particular, it is important to consider whether the circadian effect plays a major role in motor carrier accident risk. One objective of this study is to use time-dependent logistic regression to formulate a quantitative model that explicitly includes time of day along with other covariates. The second objective is to test the model using data from actual trucking company operations and to compare the results with those in the extant literature.

LOGISTIC REGRESSION MODEL

A general formulation for the time-dependent logistic regression model is as follows:

$$P_{it} = P(Y_{ti} = 1 \mid Y_{t'i} = 0 \text{ for } t' < t, X_i) = \frac{\exp[g(X_i, t, \beta)]}{1 + \exp[g(X_i, t, \beta)]} \quad (1)$$

Institute of Transportation Studies, Department of Civil and Environmental Engineering, University of California—Davis, Davis, Calif. 95616.

in which Y_{ti} is a response variable representing the occurrence ($Y_{ti} = 1$) or nonoccurrence ($Y_{ti} = 0$) of the event for individual i during the time interval t. X_i is an univariate or multivariate attribute vector for this individual, and $g(X_i, t, \beta)$ denotes some arbitrary function of X_i and a parameter vector β that will be estimated (*22–25*). In accident analysis, the conditional probability expressed in Equation 1 is the probability of an accident at time interval t, given survival (i.e., no accident) until that time; in other words, the model accounts for the survival effect (*14*). In this case, driving time is divided into equal-width intervals. It is not necessary to know the exact time of the accident; accuracy to the level of a specific interval (e.g., 30 min or 1 hr) is sufficient. The time interval in which the accident occurs or the time interval of successful completion of the trip is recorded.

The comparable conditional probability of surviving is defined as

$$Q_{it} = 1 - P_{it} \tag{2}$$

A convenient and simple functional form for $g(X_i, t, \beta)$ is a linear combination of the covariates

$$g(X_i, t, \beta) = \sum_{j=0}^{r} \beta_j X_{ji} \tag{3}$$

The X_{ji} ($j = 0, \ldots, r$) are the values of the r covariates for the driver i. The value of $X_{0i} = 1$ so that β_0 represents an intercept parameter in the regression model.

The full likelihood for the n drivers can be represented by the following:

$$L = \prod_{i=1}^{n} \left(\frac{P_{i\,t_i}}{Q_{i\,t_i}} \right)^{Zi} \prod_{t'_i \le t_i} Q_{i\,t'_i} \tag{4}$$

where $Z_i = 1$ for accident driver i, and $Z_i = 0$ otherwise, and t_i represents the number of time intervals for which driver i is exposed to the accident risk.

Equation 3 can be broken down into the following components:

$$g(X_i, t, \beta) = \sum_{j=0}^{r} \beta_j X_{ji} + \sum_{k=1}^{T-1} \beta_{r+k} X^{*}_{ki} + \sum_{n=1}^{s-1} \beta_{r+(T-1)+n} X_{ni}\,(t_i) \tag{5}$$

The first term of the right-hand side of Equation 5 represents time-independent covariates, the effects of which are assumed to be independent of time. The second term represents the time main effect (in this application, driving time), and X^{*}_{ki} represents the kth time interval for driving time. A trip with a length of k time intervals would be represented by a series of indicator variables with $X^{*}_{ki} = 1$. The last term represents the time-dependent covariate (in this application, time of day). The parameters $\beta_{r+(T-1)+n}$ are a series of coefficients associated with the s intervals used as categories for the time-dependent covariate (in this case, 11 categories of time of day). A similar model formulation was used elsewhere (*14*); Equation 5 represents an extension of the earlier model in that it includes time-dependent covariates.

To include the survival effect in the time-dependent logistic regression model correctly, several methods to treat time dependent covariates have been proposed. One approach (*26*) specifies a series of covariates to represent each time-dependent risk factor for each time interval. A nice feature of this method is that it suggests approaches to incorporate change in the underlying risk of an event over time and the prior history of an individual. However, the model in this general form could contain so many parameters to be estimated that it might be difficult to interpret.

Another approach (*27*) uses a parsimonious model to reduce the dimensionality of the model and to improve its interpretability. A duplication method is developed to overcome the assumption of standard logistic regression that restricts each individual to only one ultimate outcome. As an example of the method, consider a driver with an accident in the third time interval. Three records will be generated for this case. For the first two records, the values of the response variable are both 0 (non-accident); the value of the response variable will be 1 for the third record. For a driver who successfully completes a trip through the third interval, three records will also be generated; the values of the response variable for all three records are 0. The values of the vector of time-independent covariates for this individual will be the same in each of the three records, whereas the values of time varying covariates will depend on the related time interval. This approach is based on the following three important assumptions:

1. The underlying risk of the events in each time interval is assumed the same in this model (e.g., the risk in the first driving hour is the same as that in the 9th hr).
2. Closely related to the first assumption is that risk factors and outcome of interest are independent of time; that is, for a particular time of day (e.g., 8:00 to 10:00 a.m.) the accident risk for the 1st hr of driving is the same as that during the 9th hr of driving.
3. Only the current status of the risk factor is associated with the outcome of the event, prior history is considered unimportant.

In this research, the approach (*27*) to treat the repeated measurements will be followed, but time will be treated as categorical in the model to reflect underlying risk. This relaxes the first assumption of Cupples's model. The second assumption is relaxed by including in the model interaction terms to address the potential association between driving hours and time of day.

DATA AND VARIABLE DESCRIPTION

All data are obtained from a national less-than-truckload firm. The company operates "pony express" operations from coast to coast, with no sleeper berths. The findings are thus not intended to typify the trucking industry as a whole. As the carrier takes reasonable steps to adhere to Department of Transportation service hour regulations, most drivers in the study can be considered to comply with existing limits. The data include accidents and non-accidents from the company's national over-the-road operations.

An accident is defined as "any reported event that results in damage to the truck, personal injury, or property damage." Excluded are alleged incidents (i.e., those in which someone alleges being struck by a truck, but no report was filed or verified by the carrier). The severity ranges from minor fender benders to accidents with fatalities. A non-accident is defined as "the case in which a driver successfully completes the designed trip." This is generally called "censoring" data because the accident cannot happen after the designed trip is finished.

The time-dependent logistic regression is developed using variables that include the experience of the driver with the firm, the consecutive hours of driving on the trip in question, and the time of day. The consecutive hours of driving are the actual driving time based on the designed trip of interest that restricts the maximum driving hours limits until the accident occurs or the trip is completed. The off-duty time (short breaks) and the time on-duty without driving (intermediate terminal) during this trip are then excluded.

It is possible for a driver to make several short stops either because of feeling tired or because of an intermediate terminal. These stops do not end the trip in terms of the measured time duration t_i in Equation 4; time will accumulate after the stops until an accident or completion of the trip at the destination terminal. The driver is not given the option to terminate a trip and simply stop to sleep. The truck must reach a destination at a particular time. Therefore, the driving time is either the time to an accident or the time to censoring, each of which is independent.

A problem arises because of the need to code time of day as a series of dummy variables to account for possible non-linearities. To keep the estimated number of parameters to a reasonable scale, 2-hr time periods were chosen for time of day. In this case, the first interval is midnight to 2:00 a.m.; the twelfth is 10:00 p.m. to midnight. Given that driving time data are recorded at the level of 15 min, raw data on driving time must be converted to a series of more aggregate categorical variables of 2-hr duration. Difficulties arise because drivers take rest breaks and are off duty for some time during a typical day. When these rest breaks and off-duty times occur within a 2-hr time category, it is necessary to assign the driver as either driving or not driving for that unit of time.

The rules that determine the coding of the time of day variable are as follows

- If the driver is driving for an entire time of day represented by the variable, then the driver is coded as driving during that time of day.
- If a driver's driving time crosses more than one time of day period (for example, driving from 1:45 a.m. to 2:45 a.m.), then the most proportional time of day will be coded (in this example from 2:00 a.m. to 4:00 a.m. as driving; from midnight to 2:00 a.m. as not driving).
- If a particular driving time bisects two time-of-day periods exactly, the latter time of day is coded as driving.

In this research the total number of observations used for modeling is 1924 cases, of which 694 are accidents and 1230 are non-accidents. Accidents are deliberately oversampled relative to their actual occurrence to handle the data more efficiently. Although the sampling is a type of case-control method typically used in a retrospective study, the likelihood function in Equation 4 developed for prospective studies can still be applied because the logistic regression is adopted in this research (*28*).

EMPIRICAL RESULTS

Overview of Modeling

An overview of the time-dependent logistic regression models developed in this research is shown in Figure 1. Model 1 is developed to assess the underlying hazard of driving time only. A time-independent covariate, driving experience, and a time-varying covariate, time of day, are added and estimated in Model 2. A series of models is developed to study interactions between time of day and driving time. A separate model is developed with Model 2 and interaction terms with each time of day separately with all nine cat-

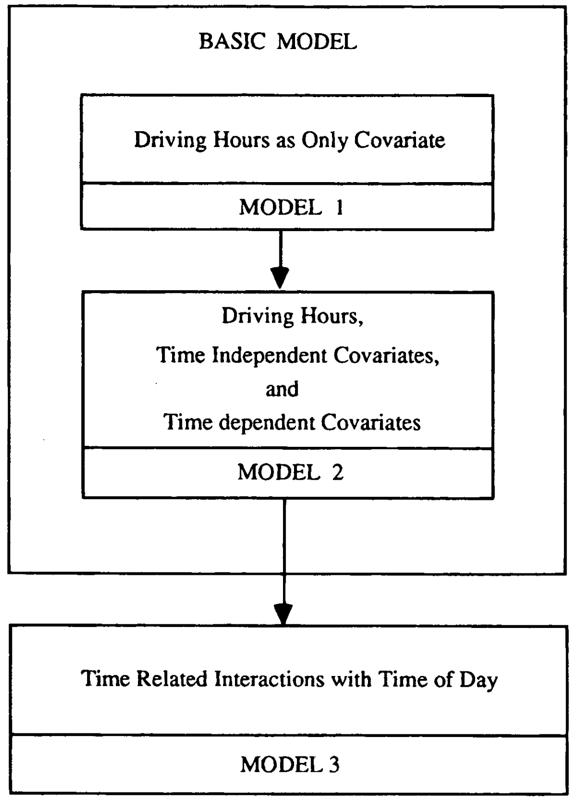


FIGURE 1 Modeling structure.

egories of driving hours. The significant variables are entered into one model and a stepwise deletion procedure used to arrive at the model shown as Model 3 (*14*). The statistical software BMDP is used to estimate coefficients and derive appropriate statistics concerning model fit.

Several tests are conducted to assess the significance of variables and models. First, a likelihood ratio test for inclusion or exclusion of a variable as a whole is used as an exploratory test of variable significance (e.g., inclusion of all categories of experience). Second, *t*-statistics are reported for each category of each variable.

The goodness of fit of a time-dependent logistic regression model to the data can be qualitatively assessed by plotting model values as a function of driving time against the product limit estimator (PLE) of the data (*23,24*). The survival function for the logistic regression is denoted as follows:

$$S(t) = \prod_{t' \leq t} Q_{it'} \tag{6}$$

and the survival function for the product limit estimator is

$$S(t) = \prod_{t' \leq t} (N_{t'} - D_{t'})/N_{t'} \tag{7}$$

where $N_{t'}$ is the number of drivers at risk at the beginning of the time interval t', and $D_{t'}$ is the number of drivers having an accident during that time interval t'. This goodness-of-fit measure has been used elsewhere (*14*).

Basic Models

A basic model that includes only driving hours is shown as Model 1 in Table 1. The model implicitly assumes that the probability of an accident is entirely determined by the driving time and is unaffected by other driver attributes. Model 1 is constructed so that there is a constant hazard within each hour and varying hazards between hours. The positive parameter in each covariate represents an increase in the log of the odds ratio or, more simply, an increase in the probability of accidents among the drivers in the specific category of the variable compared with the drivers in the corresponding baseline category. The value of the estimated coefficients represent the change in the magnitude of the chance of an accident. Accident risk is insignificantly different in the first 4 hr but rises steadily thereafter to a maximum in the 10th hr.

Model 2 shows the results of combining Model 1 with driving experience and time of day. The likelihood ratio test between Model 2 and Model 1 is significant beyond $\alpha = 0.05$, which leads to a

TABLE 1 Model Estimates and Statistics

NO	COVARIATES	MODEL 1	MODEL 2	MODEL 3
1	CONSTANT	-3.2780 *	-3.9603 *	-5.3635 *
	EXPERIENCE (year)			
2	<= 1		0.5658 *	0.5553 *
3	1 - 5		0.8210 *	0.8087 *
4	5 - 10		0.5929 *	0.5852 *
5	> 10**			
	TIME OF DAY			
6	0:00 - 2:00		0.3318 #	1.7938 *
7	2:00 - 4:00		0.0407	1.3996
8	4:00 - 6:00		0.2798	1.7563 *
9	6:00 - 8:00		0.3669 #	1.7277 *
10	8:00 - 10:00		0.2452	1.7509 *
11	10:00 - 12:00**			
12	12:00 - 14:00		0.1369	1.7179 *
13	14:00 - 16:00		0.0958	1.4638
14	16:00 - 18:00		0.4920 *	2.0918 *
15	18:00 - 20:00		0.2356	1.7032 *
16	20:00 - 22:00		0.3399 #	1.7293 *
17	22:00 - 24:00		0.0444	1.5051
	DRIVING HOURS			
18	1st HOUR (<1)	0.1404	0.1325	1.5128
19	2nd HOUR (1 - 2)**			
20	3rd HOUR (2 - 3)	0.1835	0.1903	1.5759
21	4th HOUR (3 - 4)	0.0040	0.0143	1.4655
22	5th HOUR (4 - 5)	0.4481 *	0.4673 *	1.8532 *
23	6th HOUR (5 - 6)	0.4628 *	0.4872 *	2.1375 *
24	7th HOUR (6 - 7)	0.5133 *	0.5290 *	2.1183 *
25	8th HOUR (7 - 8)	0.5392 *	0.5670 *	1.9501 *
26	9th HOUR (8 - 9)	0.8625 *	0.9119 *	2.3669 *
27	10th HOUR (> = 9)	1.8377 *	1.8200 *	3.4343 *
	INTERACTIONS			
28	(6) & (23)			-2.2060 *
29	(8) & (24)			-2.7526 *
30	(10) & (21)			-3.0946 *
31	(10) & (27)			-2.6086 *
32	(12) & (23)			-2.4369 *
33	(12) & (24)			-2.4721 *
34	(14) & (23)			-2.8250 *
35	(14) & (26)			-2.9784 *
36	(14) & (27)			-2.7428 *
37	(15) & (23)			-2.4132 *
38	(17) & (24)			-3.1159 *
39	OTHERS			-1.4307
	LOG-LIKELIHOOD VALUE	-2698.74121	-2663.0332	-2641.15161
	LIKELIHOOD RATIO TEST		71.41602 (v.s. MODEL 1)	43.76318 (v.s. MODEL 2)
	DEGREE OF FREEDOM		14	12
	CHI-SQUARE (0.95)		23.685	21.026

\# t STATISTICS SIGNIFICANT @ α=0.10

* t STATISTICS SIGNIFICANT @ α=0.05

** REFERENCED CATEGORY

rejection of the hypothesis of driving time as the only covariate. Time of day alone, without experience, failed to reject the null hypothesis of no effect as a whole.

Parameter values for driving hours in Model 2 are virtually identical to Model 1. The baseline hazard fluctuates from the 1st hr to the 4th hr with no significant difference then increases significantly until the last hr.

Drivers with experience of more than 10 years have the lowest accident risk (baseline category). The accident risk of other experience levels are all significantly different from the baseline. The highest accident risk occurs when the driving experience is between 1 and 5 years (about 2.2 times higher than for the baseline). The estimated risk increase for drivers with less than or equal to 1 year experience and those with 5 to 10 years of experience is nearly equal (about 1.7 times higher than for the baseline category).

Concerning time of day, drivers in the time between 10:00 a.m. and noon had the lowest risk, so it was defined as the baseline. The accident risk of driving during 4:00 to 6:00 p.m. is significantly higher than that of the baseline, beyond $\alpha = 0.05$. This highest accident risk may result from a combination of two effects: 4:00 to 6:00 p.m. is the evening rush hour in most major cities, increasing accident risk because of the likelihood of a collision with another vehicle; a second effect could be an association with reduced alertness because of a low circadian period for some drivers (*21*). The accident risks from midnight to 2:00 a.m., 6:00 to 8:00 a.m., and 8:00 to 10:00 p.m. are also significantly higher than during the baseline (but at $\alpha = 0.10$). Two of them involve night driving; the other involves part of the dawn period.

Inclusion of Interaction Terms

The modeling of interaction terms between time of day and driving hours is summarized as Model 3 in Table 1. The objective of testing this set of variables is to determine whether certain times of day are particularly risky (or safe) for driving hours of a particular duration. This is an examination of the effect of two time-related covariates. The addition of time-related interactions results in Model 3 having a significantly improved goodness-of-fit compared to Model 2. Figure 2 indicates little difference between the two models in a comparison of their fit to the product limit estimator of Equation 7. The fit appears good.

Consistent with the previous model, the three categories of driving experience in Model 3 have significant positive parameters, and they are of virtually the same magnitude as in Model 2. The parameters for driving hours are similar to Model 2, but the magnitudes change because of the time-related interactions.

All the significant interactions result in the reduction of accident risk for a specific time of day over time. When interaction terms are added, four of the times of day that were indifferentiable from the baseline became significantly higher in risk from the baseline. This also happened for all three of the marginally significant times of day. On the basis of these results, there is no question that time of day and driver hours interact. The interactions thus allow differentiation of times of day of constant elevated risk from those whose risk varies with driving time.

Nevertheless, some times of day have risks no different from the baseline, specifically 2:00 to 4:00 a.m. and 2:00 to 4:00 p.m. The

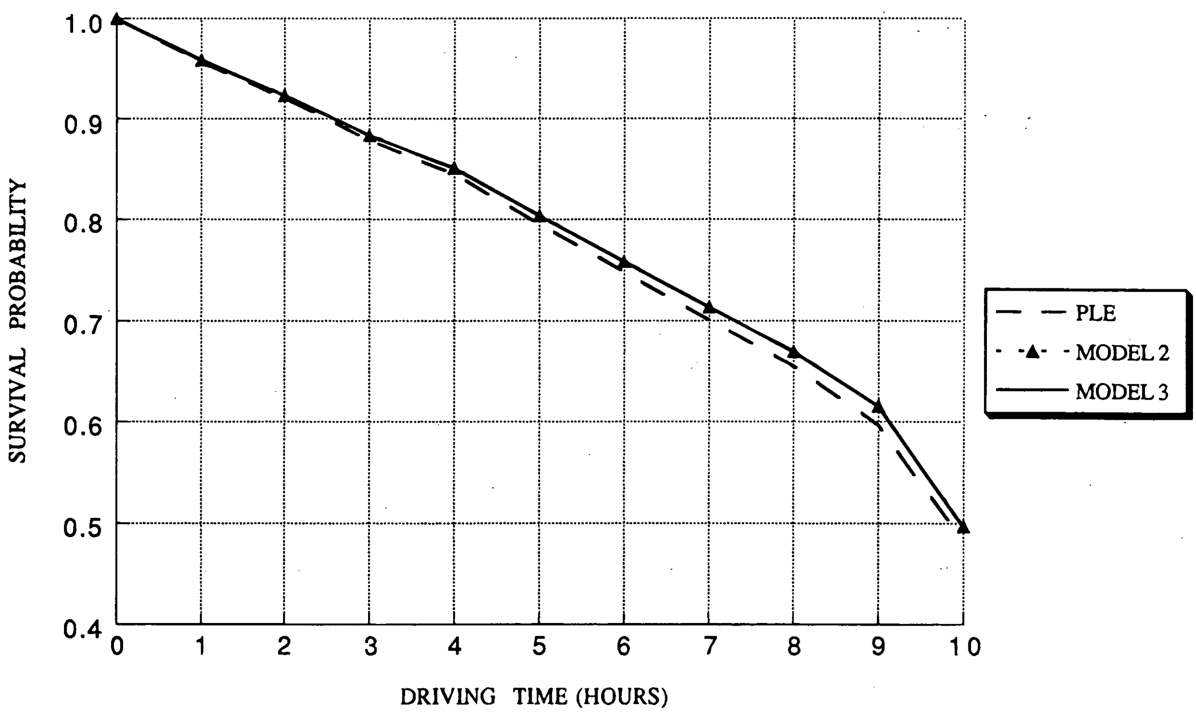


FIGURE 2 Survival curve for model goodness of fit.

TABLE 2 Survival Model Estimates and Statistics

NO	COVARIATES	COEFFICIENT	T VALUE	EXP (COEFFICIENT)
	AGE			
1	<= 39	0.1714	1.5897	1.1869
2	39 - 46**			
3	46 - 53	0.0477	0.4369	1.0488
4	> 53	0.1133	1.0273	1.1199
	EXPERIENCE (year)			
5	<= 2	0.5873	5.0676 *	1.7992
6	2 - 6	0.6428	5.4792 *	1.9018
7	6 - 10	0.5988	5.5288 *	1.8200
8	> 10**			
	DRIVING PATTERN			
9	1	0.1221	0.7881	1.1299
10	2	-0.1995	-1.1266	0.8192
11	3	0.1673	1.0869	1.1821
12	4	-0.0516	-0.3078	0.9497
13	5**			
14	6	0.0511	0.3067	1.0525
15	7	0.1418	0.9021	1.1524
16	8	0.0539	0.3421	1.0554
17	9	0.1068	0.6531	1.1127
	OFF-DUTY HOURS			
18	<= 10.5**			
19	10.5 - 13.75	-0.1311	-1.2004	0.8772
20	13.75 - 25.75	-0.1411	-1.3239	0.8684
21	> 25.75	-0.2072	-1.8946	0.8129
	TIME OF REST (hour)			
22	no rest break**			
23	<= 2	-0.0574	-0.5760	0.9442
24	2 - 4	-0.2015	-2.0307 *	0.8175
25	4 - 6	-0.3747	-3.1971 *	0.6875
26	6 - 8	-0.0383	-0.2343	0.9624
27	> 8	-0.9327	-1.2870	0.3935
	LOG-LIKELIHOOD VALUE	-5066.3731		
	GLOBAL CHI-SQUARE	81.36		
	DEGREE OF FREEDOM	22		
	P-VALUE	0		

* t STATISTICS SIGNIFICANT @ $\alpha = 0.05$

** REFERENCED CATEGORY

time period from 10:00 p.m. to midnight also has indistinguishable risk from the baseline except for a significant and negative interaction with the seventh driving hr. These times of day represent periods of particularly low risk, and they are, with one exception, independent of driving time. Other time periods with significant interactions may have lower risk for some driving times.

The prevalence among the interaction terms of significant interactions with the sixth and seventh driving hours is surprising. On the basis of the literature, there is no a priori expectation for the observation of this systematic risk reduction. Additional modeling of this data set using survival models (*29*) helps to interpret this result further.

Table 2 is the output of a survival model estimation. The model coefficients can be interpreted similarly to a linear regression model. In this case, positive coefficients imply increased risk of an accident, negative coefficients a reduced risk. Age, experience, and off-duty hours before the trip of interest are all listed as categorical variables. The time of day of multiday driving is characterized by a driving pattern number that is the output of a cluster analysis (*13*). Of particular importance to this discussion is the set of "time of rest" categorical variables, which are used to depict the taking of a rest break during a particular driving hour. Notice that rest breaks during driving hours 2 to 4 and 4 to 6 significantly lower accident risk. It appears that the interaction terms in our logistic regression model are picking up this rest break effect. The survival model is presented here strictly to clarify the interpretation of the logistic regression interaction terms. The theory of the survival model is thus not important in this context. The consistency of the effects observed is important.

SUMMARY AND RECOMMENDATIONS

A time-dependent logistic regression model has been formulated to assess the safety of motor carrier operations. The model is flexible, allowing the inclusion of time main effects, time-independent covariates, time-dependent covariates, and interaction terms. The model examines accident risk using a data set from a national less-than-truckload carrier. The model estimates the probability of having an accident at time interval *t*, subject to surviving (i.e., not having an accident) until that time interval. Individual accidents are statistically compared with a random sample of individual non-accident trips by estimating a logistic regression model with two outcomes: an accident or non-accident. Covariates tested in the model include consecutive driving time, driver experience, and time of day.

Three logistic regression models are estimated, which include main time effects (driving time), time-independent effects (driving experience), time-dependent effects (time of day), and a series of time-related interactions. Driving time has the strongest direct effect on accident risk. The first 4 hr consistently have the lowest accident risk and are indistinguishable from each other. Accident risk increases significantly after the 4th hr, by approximately 50 percent or more, until the 7th hr. The 8th and 9th hr show a further increase, approximately 80 and 130 percent higher than the first 4 hr. These results are generally consistent with those of Harris and Mackie (*8*).

Drivers with more than 10 years driving experience retain a consistently low accident risk; all other categories of driving experience have a significantly higher risk than this group.

Time of day had an effect on subsequent accident risk, but the effect was not as strong as for driving experience or driving hours. Daytime driving, particularly at noon (10:00 a.m. to 12:00 p.m.), results in a significantly lower risk of an accident. Driving from 4:00 to 6:00 p.m. has an accident risk about 60 percent higher than the baseline; drivers during the other three significant times of day also have accident risks about 40 percent higher than those during the baseline. These three involve night or dawn driving; two of them are associated with circadian rhythms.

When interactions were included, the accident risk for some times of day decrease. Particularly, most of the significant interactions fall in the sixth and seventh driving hours. Rest breaks appear to be associated generally with these risk reductions.

Time-dependent covariates play a key role in accident analysis. However, the shortage of time-varying data makes it difficult for a researcher to consider further accident analysis and solutions. As mentioned earlier, high traffic volume could be one of the reasons for the highest accident risk occurring between 4:00 and 6:00 p.m. The inclusion of road class (e.g., rural Interstate, urban local), which is a kind of time-varying risk factor, could greatly improve understanding of time-related effects. The collection of this additional time-dependent data becomes an important task in future research.

The joint study of time of day and driving time is complicated because driving time intervals could cross more than one time of day. Although some rules have been provided in this research, the approach is still rough and could result in some loss of information and bias in estimation. A more advanced approach is needed to treat the coding of time of day precisely and completely.

In this research, there is an important assumption that the prior history of an individual does not influence the outcome. Cupples et al. (*27*) used the slope of a risk factor over time to represent the effect of past history on an outcome. Time of day cannot be treated in this way because it is a categorical variable. The inclusion of prior history as a time-dependent covariate, while keeping the model parsimonious, is an important topic of future research.

ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support provided by the University of California Transportation Center.

REFERENCES

1. *Gearing Up for Safety: Motor Carrier Safety in a Competitive Environment.* Office of Technology Assessment, U.S. Congress, 1988.
2. *A Report on the Determination and Evaluation of the Role of Fatigue in Heavy Truck Accidents.* Transportation Research and Marketing, AAA Foundation for Traffic Safety, Falls Church, Va., 1985.
3. *Safety Study: Fatigue, Alcohol, Other Drugs, and Medical Factors in Fatal-to-the-Driver Heavy Truck Crashes.* Vol. 1, Report NTSB/SS-90/01, National Transportation Safety Board, Washington, D.C., 1990.
4. Haworth, N. L., T. J. Triggs, and E. M. Grey. *Driver Fatigue: Concepts, Measurement and Crash Countermeasures.* Department of Transport and Communications, Federal Office of Road Safety, Canberra, Australia, 1988.
5. McDonald, N. *Fatigue, Safety and the Truck Driver.* Taylor and Francis, London, England, 1984.
6. Brown, I. D. Fatigue. In *Medical Aspects of Fitness to Drive.* The Medical Commission on Accident Prevention, London, England, 1985.
7. Sussman, E. D., H. Bishop, B. Madnick, and R. Walter. Driver Inattention and Highway Safety. In *Transportation Research Record 1047,* TRB, National Research Council, Washington, D.C., 1985, pp. 40–48.
8. Harris, W., and R. R. Mackie. *A Study of the Relationships Among Fatigue, Hours of Service, and Safety of Operations of Truck and Bus Drivers.* Report BMCS-RD-71-2. Human Factors Research, Inc., Goleta, Calif., 1972.
9. Hackman, K. D., E. Larson, and A. Shinder. *Analysis of Accident Data and Hours of Service of Interstate Commercial Motor Vehicle Drivers.* Genasys Corp., Rockville, Md., 1978.
10. Hamelin, P. Lorry Drivers' Time Habits in Work and Their Involvement in Traffic Accidents. *Ergonomics,* Vol. 30, No. 9, 1987, pp. 1303–1333.
11. Jones, I. S., and H. S. Stein. *Effect of Driver Hours of Service on Tractor-trailer Crash Involvement.* Insurance Institute for Highway Safety, Arlington, Va., 1987.
12. Hertz, R. Hours of Service Violations Among Tractor-Trailer Drivers. *Accident Analysis and Prevention,* Vol. 23, No. 1, 1991, pp. 29–36.
13. Kaneko, T., and P. P. Jovanis. Multiday Driving Patterns and Motor Carrier Accident Risk: A Disaggregate Analysis. *Accident Analysis and Prevention,* Vol. 24, No. 5, 1992, pp. 437–456.
14. Lin, T., P. P. Jovanis, and C. Yang. Modeling the Safety of Truck Driver Service Hours Using Time-Dependent Logistic Regression. In *Transportation Research Record 1407,* TRB, National Research Council, Washington, D.C., 1993, pp. 1–10.
15. Mackie, R., J. F. O'Hanlon, and M. A. McCanley. *A Study of Heat Noise and Vibration in Relation to Driver Performance and Physiological Status.* Report DOT HS-801 313. Human Factors Research, Inc., Goleta, Calif., 1974.
16. Mackie, R., and J. Miller. *Effects of Hours of Service, Regularity of Schedules and Cargo Loading on Truck and Bus Driver Fatigue.* Report DOT HS-803 799. Human Factors Research, Inc., Goleta, Calif., 1978.
17. Hertz, R. Tractor-Trailer Driver Fatality: The Role of Nonconsecutive Rest in a Sleeper Berth. *Accident Analysis and Prevention,* Vol. 20, No. 6, 1988, pp. 431–439.
18. Sweedler, B. M., K. Quinlan, and M. Brenner. Fatigue and Drug Interaction in Fatal to the Driver Heavy Truck Crashes. *Proc., 34th Association for the Advancement of Automobile Medicine,* 1990, pp. 491–504.
19. Large Truck Survey Program. University of Michigan Transportation Research Institute, *Proc., of the National Truck Safety Symposium,* 1987.
20. Jovanis, P. P., T. Kaneko, and T. Lin. Exploratory Analysis of Motor Carrier Accident Risk and Daily Driving Patterns. In *Transportation Research Record 1322,* TRB, National Research Council, Washington, D.C., 1991, pp. 34–43.
21. Mathematics of Sleep. *Technology Review,* Feb.–March 1987.
22. Brown, C. C. On the Use of Indicator Variables for Studying the Time-Dependence of Parameters in a Response-Time Model. *Biometrics* 31, 1975, pp. 863–872.

23. Abbott, R. D. Logistic Regression in Survival Analysis. *American Journal of Epidemiology* Vol. 121, No. 3, 1985, pp. 465–471.
24. Efron, B. Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve. *Journal of the American Statistical Association*, Vol. 83, No. 402, 1988, pp. 414–425.
25. Hosmer, D. W., and S. Lemeshow. *Applied Logistic Regression.* John Wiley & Sons, Inc., N. Y., 1989.
26. Wu, M., and J. Ware. On the Use of Repeated Measurements in Regression Analysis with Dichotomous Response. *Biometrics*, Vol. 35, 1979, pp. 513–521.
27. Cupples, L. A., R. B. D'agostino, K. Anderson, and W. B. Kannel. Comparison of Baseline Repeated Measure Covariate Techniques in the Framingham Heart Study. *Statistics in Medicine*, Vol. 7, 1988, pp. 205–218.
28. Prentice, R. L., and R. Pyke. Logistic Disease Incidence Models and Case-Control Studies. *Biometrika* 66, 1979, pp. 403–411.
29. Yang, C., P. P. Jovanis, and T. Lin. *Assessing Motor Carrier Driving Risk Using Cox's Semi-Parametric Model with Multiple Stop Effects.* ITS-Davis Research Report 92-10, University of California, Davis, 1992.

Publication of this paper sponsored by Task Force on Statistical Methods in Transportation.

Analyzing the Relationship Between Crash Types and Injuries in Motor Vehicle Collisions in Hawaii

KARL KIM, LAWRENCE NITZ, JAMES RICHARDSON, AND LEI LI

A statistical model was developed to explain the relationship between types of crashes and injuries sustained in motor vehicle accidents. By using techniques of categorical data analysis and comprehensive data on crashes in Hawaii during 1990, a model was built to relate the type of crash (e.g., rollover, head-on, sideswipe, rear-end, etc.) to a KABCO injury scale. An "odds multiplier" was developed that enabled comparison according to crash type of the odds of particular levels of injury relative to noninjury. The effects of seat belt use on injury level also were examined, and interactions among belt use, crash type, and injury level were considered. Differences between crash types and the effectiveness of seat belts are discussed along with how log-linear analysis, logit modeling, and estimation of "odds multipliers" may contribute to traffic safety research. Some implications of the findings for appropriate interventions and future research are presented in a concluding section.

Over the years, there have been concerted efforts directed at reducing mortality and morbidity associated with traffic collisions. In spite of these efforts, automobile crashes still claim thousands of lives in America and cost billions of dollars in medical treatments and lost wages. The relationships between collisions and injuries are complicated by the presence of multiple factors—environmental, behavioral, vehicular, and others. It is sometimes useful to consider a causal chain of factors relating background driver characteristics to risk-taking behaviors, then to interactions among the driver, vehicle, and environment, thereby producing explanations of how accidents occur and what are some of the related health or economic outcomes. This study focuses narrowly on the relationship between crash type and injury level as an initial step toward developing a more complicated model of the structure of automobile crashes and injury outcomes.

The relationship between crash type and injury is one that deserves further inspection for several reasons. Obviously, all collisions are not the same—at least not in terms of the chance of injury or fatality. Moreover, although some studies have focused on specific types of crashes (e.g., rear-end collisions or rollovers), it is also important to categorize types of crashes, not only to better understand the different causes of injury but also to understand how different crash types suggest different types of intervention. For example, although head-on collisions suggest the need for more physical improvements such as roadway barriers, rear-end collisions point to the need for vehicle warning systems or more driver education on safe following distances. Broadside collisions may indicate a need for signalized intersections. Rollovers may suggest the need for stronger vehicle standards or environmental changes. Seat belt use is included in the model because whether someone is wearing a seat belt is part of the physical circumstances of the crash. In the attempt to predict injury severity, it is logical to include seatbelt use as, in effect, part of the crash type. Seat belt use is expected to have a differential effect across crash types, (e.g., greater for rollovers than for sideswipes.)

The present study is part of a larger effort funded by NHTSA, U.S. Department of Transportation, known as the Crash Outcome Data Evaluation System (CODES) project. Seven sites were selected, including the University of Hawaii, to build data bases linking crash reports, emergency medical records, medical claims data, and other health-related information. Grantees were also required to conduct various analyses of the effectiveness of seat belt and motorcycle helmet use.

There are several reasons Hawaii provides an excellent site for the analysis described in this study. Hawaii is an island, isolated from the U.S. mainland, with a centralized system of government made up of four county governments and the state government. The state maintains computerized records of all major traffic crashes. With year-round favorable weather, a variety of different urban, rural, and suburban roadway and highway conditions, and a resident population of more than one million, conditions are good for conducting traffic safety research. Following a brief discussion of data and methodology the results of the modeling efforts, the estimation of model effects, and some concluding remarks are presented.

SOURCES OF DATA AND METHODOLOGY

In this section the Motor Vehicle Accident (MVA) file and the methodology for modeling the relationship between crash type and injury levels are described. The log-linear modeling approach is described as well as a method of computing an odds multiplier, which is a useful means both of summarizing the results of the categorical data analysis and of comparing crash type and injury categories. The relationship between seat belt use and injury and an estimation procedure for examining that interaction with the other two variables in our model are specified.

The data used are collected by police officers dispatched to the scene of a collision. The data are collected under less than ideal conditions but represent the best and most comprehensive data available on crashes in Hawaii. The information is entered by hand on to forms that are then sent to the Department of Transportation, State of Hawaii. Data from the forms are keypunched into a computer system, edited, and used for various analyses. This analysis is based

K. Kim, Department of Urban and Regional Planning; L. Nitz, Department of Political Science, University of Hawaii at Manoa, 2424 Maile Way, Honolulu, Hawaii 96822. J. Richardson, Department of Management and Industrial Relations, University of Hawaii at Manoa, 2404 Maile Way, Honolulu, Hawaii 96822. L. Li, School of Public Health, University of Hawaii at Manoa, EWC Box 1248, 1777 East West Rd., Honolulu, Hawaii 96848.

on all available data for 1990. It focuses attention on three variables: (a) crash type, (b) injury level, and (c) seat belt use.

The crash-type variable was derived by examining each collision and the maneuvers of every involved vehicle. It is important to note that a distinction was made between vehicles that struck other vehicles (e.g., "rear-enders" or "broadsiders") and those vehicles that were struck (e.g. "rear-ended" or "broadsided").

The injury variable is based on a standard KABCO scale in which the major categories are "killed," "incapacitating injury," "nonincapacitating injury," "possible injury," and "uninjured." One question that remains unanswered is the reliability of police reporting of injuries. This question is one that the CODES project will address, because the linked data will facilitate the comparison of police, hospital, and insurance company reporting of injuries.

Seat belt use is a dichotomous response variable. It is a particularly interesting variable because Hawaii has one of the highest reported rates of seat belt use in the U.S. (*1*). Even so, the reported use rate of 97 percent in the MVA data exceeds the observed rate of 85 percent. For many accidents, seat belt use is self-reported, because the driver has gotten out of the car by the time the police officer arrives. Hawaii's mandatory seat belt law gives such drivers a clear incentive to claim they were using a seat belt. The use rate drops considerably in the more severely injured categories, to below 50 percent for killed. The reported use rates for severely injured or killed drivers are likely to be more accurate, either because the person is still in the car when the officer arrives or because the injuries and damage make it apparent that seat belts were not used. Even though greater overreporting of seat belt use in the less injured categories will increase the apparent effectiveness of seat belts, the effect is small compared with the total differences in use rates across the injury categories.

The analysis was restricted to drivers of automobiles to ensure that exposure to injury was comparable, because there are differences between front and rear seat, driver and passenger environments. Single and multiple vehicle crashes are included. Frequencies and percentage distributions for the three variables are shown in Table 1.

TABLE 1 Frequency Distribution of Variables

Variable	Frequency	Percent
Crash Type		
Head-on	927	3.1
Rear-ended	5,961	20
Rear-ender	7,287	24.4
Broadsided	3,130	10.5
Broadsider	3,134	10.5
Sideswipe	8,945	30
Roll-over	445	1.5
Seatbelt Use		
Nonbelted	768	2.6
Belted	29,061	97.4
Severity of Injury		
K Fatal	29	0.1
A Incapacitating	234	0.8
B Nonincapacitating	2,364	7.9
C Possible	3,086	10.3
O No Injury	24,116	80.8

Source: Motor Vehicle Accident File, State of Hawaii

As seen in Table 1, the data used are categorical; that is, they represent various categories of crash type or injury. There is a variety of different approaches for handling categorical data analysis, including log-linear models, additive models, and partitioning of chi square. For further discussion of the differences between these approaches, see the work by Feinberg (*2*). Only one variable, injury scale, can be construed as ordinal, thereby greatly limiting the use of ordinary regression or ANOVA-type procedures. Of the different approaches, the log-linear approach was selected because of the added convenience of being able to investigate the underlying interrelated causal structure, work with odds ratios, and estimate the odds of being in a particular injury category given crash type or belt use. At the same time, the log-linear approach allowed estimation of parameters and corresponding tests of significance to discern the size and relative nature of effects in the model. Although SAS/STAT CATMOD and BMDP were used to derive these results, there are other packages that provide adequate support of log-linear modeling procedures.

There are different strategies for categorical data analysis and model building. First the relationships between different variables and injury level were examined. Characteristics of drivers, vehicles, and environments were examined, with the conclusion that crash type was a significant factor in determining injury. A model preserving categories of crash type was developed to see the differential injury outcomes. The objective was to build the simplest model that would adequately fit the data. Golub and Recker (*3*) use the method of log-linear modeling to analyze truck-involved freeway accidents. They consider the relationship between crash types and each of different variables, including collision factors, crash locations, weather, and others. The log-linear approach strives to fit cell frequencies with an additive model, incorporating main effects and interactions between variables. This is a three variable model with variables C (crash type), S (seat belt use), and I (injury level). Typically, an X^2 or G^2, log-likelihood ratio, goodness-of-fit statistic is used to determine the acceptance or rejection of the model. The best-fitting model includes all three main effects and all three possible two-way interactions

$$\log_e (m_{ijk}) = u + u_{I(i)} + u_{C(j)} + u_{S(k)} + u_{IC(ij)} + u_{IS(ik)} + u_{CS(jk)} \quad (1)$$

where m_{ijk} is the expected cell frequency and u is the parameter to be estimated. The overall grand mean of the cell frequencies is u. Each of the subscripted u parameters represents a deviation from the grand mean due to that effect. For example, $u_{C(j)}$ is the crash type effect, with a separate parameter estimate for each crash-type category. In standard hierarchical notation, this model is denoted by [IC] [IS] [CS], where all of the lower order terms are implicitly included [see Feinberg (*2*)]. From the best-fitting log-linear model, the parameter estimates, u, are obtained as well as their statistical significance.

Next, a log-linear modeling approach was selected in which the response variable (I) is expressed as a log odds (logit) because it allows, for example, the comparison of the odds of fatality among those involved in head-on crashes to the odds of fatality among those involved in other types of crashes. With the logit model, the parameters provide a measure of the magnitude and direction of effects of the independent variables on the response variable. From the log-linear model (1), using injury category O, no injury, as the baseline, the logit model for injury level is

$$\begin{aligned}\log_e(m_{ijk}/m_{Ojk}) &= [u + u_{I(i)} + u_{C(j)} + u_{S(k)} + u_{IC(ij)} + u_{IS(ik)} + u_{CS(jk)}] \\ &\quad - [u + u_{I(O)} + u_{C(j)} + u_{S(k)} + u_{IC(Oj)} + u_{IS(Ok)} + u_{CS(jk)}] \\ &= [u_{I(i)} - u_{I(O)}] + [u_{IC(ij)} - u_{IC(Oj)}] + [u_{IS(ik)} - u_{IS(Ok)}] \\ &= w_i + w_{C(j)} + w_{S(k)} \qquad (2)\end{aligned}$$

where the w is the parameter to be estimated. Note that w is calculated from the u estimated for the log-linear model (*2*). For more discussion of the relationship between log-linear and logit models, see work published elsewhere (*2,4–7*).

FINDINGS

The parameter estimates and the standard errors scores are given in Table 2. The fitted model produced a G^2 (log-likelihood ratio estimate) of 32.4, with 24 df, and a probability of .12, signifying model acceptance. There are a number of different results shown in the table. The row of main effects shows the relative distribution of injury type. The effects of crash type and seat belt use are also listed. The table includes standard errors for each of the effects. Most of the effects are statistically significant (indicated by *) at the .05 level. The estimates provide a measure of the magnitude and direction of the effects. For example, being involved in a head-on or rollover crash (positive values) increases the likelihood of being fatally injured, and being rear-ended (negative values) reduces the chance of being a traffic fatality.

To better interpret the findings, Table 3 was prepared. This table enables comparison of the odds of injury against no injury (baseline) according to different crash types and the use or nonuse of seat belts. The presentation of the odds multiplier is a useful way of examining these data. To compute the odds multipliers, exponentiate both sides of the logit model,

$$(m_{ijk}/m_{Ojk}) = \exp[w_i]\exp[w_{C(j)}]\exp[w_{S(k)}] \qquad (3)$$

The first factor is the baseline odds of being in injury category i relative to no injury. The next two factors are the odds multipliers for crash type and seatbelt use.

The baseline odds (the first row in Table 3) represent the odds of injury to no injury across injury type. The first column represents the odds of no injury to no injury, which is obviously equal to one. It is interesting to note that the odds are greatest for nonincapaci-

TABLE 2 Parameter Estimates of Log-Linear Model of Crash Types, Seat Belt Use, and Severity of Injury (Model: [IC] [IS] [CS], G^2 = 32.4, df = 24, p = .12)

Severity of Injury		No Injury [0]	Possible Injury [C]	Non-Incapacit. [B]	Incapacit. Injury [A]	Fatal [K]
Injury Main Effect:	u	2.281	0.56*	0.9*	-0.99*	-2.752*
	s.e..	0.055	0.069	0.06	0.088	0.178
Seatbelt x Injury						
Nonbelted:	u	-0.68*	-0.54*	-0.131*	0.287*	1.063*
	s.e.	0.048	0.062	0.053	0.083	0.141
Belted:	u	0.68*	0.54*	0.131*	-0.287*	-1.063*
	s.e.	0.048	0.062	0.053	0.083	0.141
Crash type x Injury						
Head-on:	u	-0.914*	-0.634*	-0.23*	0.599*	1.18*
	s.e.	0.087	0.103	0.096	0.138	0.291
Rear-ended:	u	0.496*	1.074*	0.375*	-0.548*	-1.398*
	s.e.	0.18	0.182	0.182	0.231	0.701
Rear-ender:	u	0.683*	0.258*	0.024	-0.263	-0.702
	s.e.	0.113	0.117	0.118	0.162	0.43
Broad-sided:	u	0.223	0.112	-0.065	-0.118	-0.152
	s.e..	0.116	0.121	0.122	0.176	0.429
Broadsider:	u	0.408*	-0.14	-0.082	-0.117	-0.069
	s.e..	0.116	0.124	0.124	0.181	0.429
Sideswipe:	u	0.724*	-0.175	-0.214*	-0.219	-0117
	s.e.	0.116	0.124	0.124	0.181	0.429
Roll-over:	u	-1.62*	-0.497*	0.191	0.666*	1.259 *
	s.e..	0.101	0.115	0.1	0.154	0.293

*Indicates significance at p ≤ .05.

TABLE 3 Odds Multipliers of Seat Belt Use and Crash Types on Severity of Injury ("No Injury" is Reference Category)

Severity of Injury	No Injury [0]	Possible Injury [C]	Nonincap. Injury [B]	Incap. Injury [A]	Fatal [K]
Baseline odds	1	0.179	0.251	0.038	0.007
Odds Multipliers					
Seatbelt					
Nonbelted	1	1.15	1.73	2.63	5.71
Belted	1	0.87	0.58	0.38	0.18
Crash type					
Head-on	1	1.32	1.98	4.54	8.12
Rear-ended	1	1.78	0.89	0.35	0.15
Rear-ender	1	0.65	0.52	0.39	0.25
Broadsided	1	0.9	0.75	0.71	0.69
Broadsider	1	0.58	0.61	0.59	0.62
Sideswipe	1	0.41	0.39	0.39	0.43
Roll-over	1	3.07	6.12	9.84	17.8

Source: Motor Vehicle Accident File, State of Hawaii
Note: The odds of a particular category relative to no injury are calculated by multiplying the odds multiplier and the baseline odds.

tating injury and smallest for fatal injury. Other findings can be grouped by (a) relationship between crash type and injury [CI], (b) relationship between seat belt use and injury [SI], and (c) interaction effects among crash type, seat belt use, and injury [CSI].

Crash Type and Injury

Although the baseline odds of being killed in a car crash in Hawaii are small, the odds of being killed increase greatly for those involved in head-ons and rollover crashes. The odds of incapacitating and nonincapacitating injury also increase significantly for those involved in head-ons and rollovers. The odds of receiving an incapacitating or fatal injury for rear-end, broadside, and sideswipe crashes are considerably lower than the odds for head-ons and rollovers. If rear-enders are compared with rear-ended, the odds of possible and nonincapacitating injuries are greater for those rear-ended and the odds of incapacitating injury and fatal injury are greater for the rear-enders. Being broadsided appears to increase overall the odds of injury across category, compared with the broadsiders. Sideswipes had the lowest odds of injury across all categories. Another interesting finding is that, although being rear-ended has the lowest odds of being killed (even lower than sideswipes), being rear-ended has the highest odds of possible injury (with the exception of rollovers).

Seat Belt and Injury

Also in Table 3 are estimates of the odds of injury to noninjury for seat belt users and nonusers. According to these results, the odds of being killed for nonusers of seat belts are 32 times greater than the odds for seat belt users. The odds of an incapacitating injury are nearly seven times greater for nonusers. In fact, the odds across all injury categories (K, A, B, C) are greater for nonusers than for users of seat belts. These findings also support other research that has found that seat belts are effective in preventing serious or fatal injuries.

Other Interaction Effects

To test for the interaction effects the log-linear model for crash type and seat belt use, [CS] as well as the saturated model [CSI], was also run. Significant interaction effects were found between rollover crashes and seat belt use, namely, that nonusers were more likely to be involved in rollovers than were users of seat-belts. It was also found that seat belt users were more likely to be rear-ended or broadsided than nonusers. These interaction effects were attributed to behavioral factors that were not controlled for in this study and will be the subject of future research. When the fully saturated model was run, no significant higher order interaction effects were found, demonstrating that the developed model is acceptable. Moreover, the absence of significant three-way interactions suggests that the benefits of seat belt use, as described earlier, exist in all categories of crash type.

CONCLUSIONS

Two types of conclusions were reached. The first pertains to the findings, and the second to the methodology. The findings suggest that injury levels are related to crash type and that rollovers and head-on collisions produce the most severe injuries. The lowest odds for fatality are for those who have been rear-ended. There are, moreover, different levels of injury associated with rear-end, broadside, and sideswipe collisions. The odds of serious injury are greater for broadside collisions than for either sideswipes or rear-end collisions. Being a rear-ender increases the odds of incapacitating or fatal injuries compared with being rear-ended. Seat belt use was shown to have strong positive effects in reducing the risk of injury and fatality, although the effects were greatest in the serious and fatal injury categories.

Log-linear analysis provides a powerful tool with which to examine the relationships among crash types, seat belt use, and injury levels. The strategy in this study was to use log-linear analysis to uncover underlying relationships, to convert log-linear equations into logit functions to estimate parameters and model effects, and finally to convert the logit model results into odds multipliers to yield comparisons among various categories of crash type, seat belt use, and injury level. The method is particularly useful in circumstances (typical of epidemiological studies) in which the actual number of cross-classified events (e.g., fatally injured drivers involved in rollover crashes) may be small or the research questions involve categorical data.

More work is needed on both fronts. The relationship between injury and crash type warrants further consideration of behavioral and human factors. Future modeling efforts will concentrate on building structural equations relating background driver characteristics to risk-taking behaviors to crash types and injuries. As part of the Hawaii CODES project, economic data will also be examined to link the costs of medical treatment with various crash types. For modeling efforts, more widespread use of categorical data analysis techniques is predicted, and other researchers are invited to apply similar techniques to estimate the effects of crash type, seat belt use, alcohol involvement, driver education, and other factors on injury level. We would like to test this method on a larger data set—perhaps using national data or at least using data from a larger population base.

REFERENCES

1. *SAS, SAS/STAT—Users Guide*. SAS Institute, Inc., Cary, N.C., 1988.
2. Feinberg, S. *The Analysis of Cross-Classified Categorical Data*. The MIT Press, Cambridge, Mass., 1980.
3. Golub, T., and W. Recker. An Analysis of Truck-Involved Freeway Accidents Using Log-Linear Modeling. *Journal of Safety Research*, Vol. 18, 1987, pp. 121–136.
4. Jobson, J. D. *Applied Multivariate Analysis, Volume II: Categorical and Multivariate Methods*. Springer-Verlag, N. Y., 1992.
5. Agresti, A. *Categorical Data Analysis*. John Wiley & Sons, Inc., N. Y., 1990.
6. Selvin, S. *Statistical Analysis of Epidemiological Data*. Oxford University Press, New York, 1991.
7. Kim, K. Effects of Enforcement on Seat Belt Use in Hawaii. In *Transportation Research Record 1325*, 1991, pp. 51–56.

Publication of this paper sponsored by Task Force on Statistical Methods in Transportation.

Commercial Bus Accident Characteristics and Roadway Treatments

CHARLES V. ZEGEER, HERMAN F. HUANG, JANE C. STUTTS, ERIC RODGMAN, AND JOSEPH E. HUMMER

Traffic accidents involving buses result in about 35,000 injuries in the United States each year. This study describes bus and motor vehicle accident characteristics and recommends roadway-related countermeasures. Analyses were carried out on a primary study file of 8,897 commercial bus crashes in five states—Illinois, Maine, Michigan, Minnesota, and Utah—for 1985 through 1989. A subset file with urban crashes in four states and the entire Illinois motor vehicle accident file with all vehicle types were also analyzed. The overall number of crashes was highest in winter, perhaps partly because of snow and ice. Older buses were overrepresented in injury and fatal crashes in comparison with newer buses. Neither bus driver age nor gender was related to accident involvement. Bus crashes at traffic lights were more likely to cause injuries and fatalities than those at stop signs. In Illinois, the most common bus accident types were rear end with one vehicle stopped, sideswipe same direction, and turning. Rear-end and angle accidents were most likely to cause injuries and fatalities. A number of measures may be used to improve bus safety. Roadway improvements on bus routes include wider travel lanes, paved shoulders or bus pull-off lanes, wider intersection turning radii, separate turn lanes, restriction of on-street parking, proper use and placement of signs and lane markings, and separate left-turn phasing. General roadway improvements in suburban and rural areas that can also reduce bus crashes include flatter roadside slopes and improved design of guardrail and roadway alignment. Future research needs related to bus transit safety are also discussed.

Traffic crashes and injuries related to buses represent a safety problem on U.S. highways. For example, in 1990 an estimated 64,000 of the 627,000 registered buses nationwide were involved in crashes, or 10.2 percent. By comparison, only 5.8 percent of other types of vehicles were involved in crashes in that same year. The crash rate for buses was 11.17/million vehicle miles, compared with 5.51 for passenger cars (*1*). The higher involvement rate of buses results from greater exposure to potential accident situations including stop-and-start operation (and perhaps more encounters with other vehicles associated with congested urban roadways).

Bus crashes take a substantial toll of injuries and deaths. In 1990, an estimated 32,000 bus occupants sustained minor or moderate injuries in highway crashes. Another 3,000 sustained serious injury (*1*), including 32 deaths. The number of occupant injuries or fatalities per 100 crash involvements was 54.7 for buses, compared with only 29.5 for passenger cars (*2*). This higher rate may be attributed to the lack of passenger restraints on buses and to the large number of occupants on buses. In addition, bus crashes are associated with approximately 100 deaths to nonoccupants (largely pedestrians and bicyclists) and 200 deaths to occupants of other vehicles per year, according to the Fatal Accident Reporting System (*2*).

This study quantifies the characteristics and causes of crashes involving commercial buses (i.e., all types of full-sized buses except for school buses) and their resulting injuries. The analysis considers bus crashes with other motor vehicles and bus run-off-road crashes (e.g., rollovers or striking poles, trees, and other fixed objects).

In conclusion, recommendations are made for highway improvements to reduce the number and severity of bus-related highway crashes. Recommended modifications to bus design features and bus driver training are found elsewhere and are directed primarily to highway designers and engineers (*3*).

LITERATURE REVIEW

It is convenient to distinguish between collision and noncollision bus accidents. Collisions include crashes with other motor vehicles, bicycles, pedestrians, and fixed objects. Examples of noncollisions are passenger falls while boarding, alighting, or riding buses. This section discusses previous literature on collision accidents only.

Dixon et al. (*4*), examine the injury-producing mechanisms for five types of collision accidents: head-on, rear-end, sideswipe, side impact, and rollover. Head-on collisions are those that involve impact at the front of the vehicle, causing the bus to decelerate (and where the direction of deceleration is toward the rear of the bus). Rear-end collisions usually involve another vehicle running into the back of the bus, causing the bus to accelerate. Contact with the bus described as a "glancing blow to the side" is a sideswipe collision. A side impact is characterized by lateral acceleration.

Of the five accident types, rollover accidents are the most likely to result in severe passenger injury or death. A lack of occupant restraints (i.e., seat belts) results in uncontrolled body movement, and during rollover passengers fall against internal bus fittings and other passengers. Partial or complete passenger ejection through windows, doors, or openings in the passenger compartment created by the collision may result in severe injuries. Injuries may also occur because of the collapse of a roof or wall into the passenger compartment (*4*). For all motor vehicle accidents, occupants who are ejected are four times more likely to suffer a serious or fatal injury than occupants who are not ejected (*1*).

Jovanis et al. (*5*) conducted one of the few studies that analyzed data bases related to bus crashes on a large scale. That study analyzed accident report data from PACE, a suburban bus transit agency in metropolitan Chicago, for 1982 through 1984. Eighty-nine percent of the 1,800 bus accidents involved collisions with another vehicle or object; 11 percent involved noncollision passen-

C. V. Zegeer, H. F. Huang, J. C. Stutts, E. Rodgmon, Highway Safety Research Center, University of North Carolina, 134½ E. Franklin St., Campus Box 3430, Chapel Hill, N.C. 27599. J. E. Hummer, Department of Civil Engineering, North Carolina State University, P.O. Box 7908, Raleigh, N.C. 27695.

ger injuries while boarding, alighting, or moving about the bus. The authors suggest that buses pose the greatest risk to automobile occupants when the buses are stationary, such as when stopped behind other vehicles or when processing passengers. Bus accidents did not appear to be more prevalent during times of darkness. The number of accidents dropped during night hours, reflecting lower service frequency and lower levels of automobile traffic. The gender and age of bus drivers did not contribute to accidents. However, the number of years of experience was found to be a contributing factor. Drivers with 3 to 6 years of experience at PACE were significantly overrepresented compared with those with less or more experience. The accident frequency by time of day generally followed congestion patterns. The number of accidents along a route was virtually linear with its mileage and negatively correlated with vehicle headway and speed (*5*).

Other researchers have analyzed bus accident data on a large scale in Great Britain; Delhi, India; and Victoria, Australia (*6–9*). In Great Britain, 43 percent of bus passenger injuries occurred as the result of collisions; 57 percent was the result of falls and other incidents under normal conditions (*6*). Another study in Great Britain found that 20 percent of the casualties were pedestrians (*7*). More than 50 percent of bus injury accidents occurring in Delhi involved pedestrians, cyclists, and motorcyclists (*8*). Nearly 14 percent of the bus injury accidents in Victoria involved pedestrians.

Additional research is needed to fill the gaps that exist in the literature. For example, the characteristics of bus accidents have not been adequately compared with the characteristics of accidents involving other vehicle types. Little information is available on the role of roadway, driver, or environmental features in bus accidents. Furthermore, additional information is needed on the types of roadway treatments that could potentially reduce the number of bus crashes on streets and highways. This study addresses some of the gaps.

DATA SOURCES

The data base chosen for analysis in this study was the Highway Safety Information System (HSIS). This data base consists of computerized information related to motor vehicle crashes, traffic volume data, and roadway characteristics from Michigan, Minnesota, Maine, Illinois, and Utah. The HSIS data were obtained from the respective states by the University of North Carolina's Highway Safety Research Center through funding from FHWA. The HSIS states were chosen on the basis of the availability of good quality data, the capabilities for merging various data files, and other factors. Although these states may not be representative of the entire United States, they have data on accidents in urban and rural areas, on different types of roads, under a variety of climatic and geographic conditions, and on roadways with various design features. Thus, the information obtained from these states was useful in achieving the primary study objective.

The HSIS files contained information on 8,897 bus crashes that occurred from January 1, 1985, through December 31, 1989. For each accident, the available information included when the crash occurred (i.e., time, day of week, month), environmental conditions (light and weather conditions), vehicle information (e.g., age of bus), driver information (age, gender, injury), accident type (i.e., single vehicle, sideswipe, turning accident, etc.), and crash severity. These variables were analyzed to gain a better understanding of factors related to bus crashes.

ANALYSIS METHODS

Several methods were used to analyze the HSIS bus accident data. The most common analysis technique was a simple comparison among the levels of a particular variable. For example, the numbers of bus-involved accidents reported by day of the week were computed and compared. These simple comparisons were useful when the variable was not related to exposure. The study team also made many comparisons between the levels of a variable on the basis of the percentage of severe accidents (i.e., accidents involving one or more fatalities or injuries). These comparisons show the levels of the variable that deserve particular attention.

For some types of analyses, there was a need to compare crash factors for buses with those of other vehicle types. For example, to determine the types of crashes in which buses were overrepresented (e.g., sideswipe crashes), it was necessary to compare the distribution of crash types for buses with that for cars and pickups, trucks, and school buses. For these analyses, all 620,000 vehicle crash involvements (including 1,500 bus-involved crashes) from the Illinois accident file for 1988 and 1989 were used. Bus crashes were then compared with crashes of other vehicle types for accident types and crash severity.

Some analyses were performed for all bus-involved accidents in the five-state sample. However, these analyses include accidents involving intercity buses on rural highways, which have different characteristics from accidents involving intracity transit buses. Therefore, most detailed analyses were performed for bus-involved accidents on urban surface streets, using the best available definition for those factors in each state.

Although the computer accident files contained a large sample of bus crashes and dozens of variables of interest for each accident, a limitation of the study was that no bus "exposure" data were available. In particular, statewide bus mileage data were not available for computing overall bus accident rates (e.g., in terms of bus accidents per million vehicle miles of travel) or for computing accident rates by driver characteristic, age of bus, and so forth. The lack of suitable exposure data has also been a problem in safety analysis of trucks (e.g., by truck size and configuration) and other vehicle types for research purposes.

A substitute or surrogate measure of exposure can be used, the "innocent victim technique." For this study and for driver categories, driver age, driver sex, and bus model year variables were analyzed with the innocent victim technique. This technique adjusts for the exposure of driver or vehicle-related groups using only accident data. The technique relies on the assumption that a group's (e.g., bus group's) exposure is related to the number of times the group's members are involved in crashes in which they are not the "at-fault" or striking vehicle, and thus are "innocent victims."

The best way to understand the innocent victim technique is to think through an example application. Suppose an analyst wants to know whether younger drivers are overrepresented in intersection-related crashes. The analyst computes that 20 percent of drivers involved in intersection-related crashes were less than 25 years old. The analyst then computes that 15 percent of the innocent victims of crashes at intersections were less than 25 years old. The ratio of the 2 percentages provides an indication of overrepresentation. Because the percentage of younger drivers involved in all intersection-related accidents is higher than the percentage of younger innocent victims, the analyst concludes that younger drivers are overrepresented. The innocent victim technique has been used by researchers in accident studies for more than 20 years. Readers interested in a review of the theory and applications of the technique are referred to Bowman and Hummer (*10*).

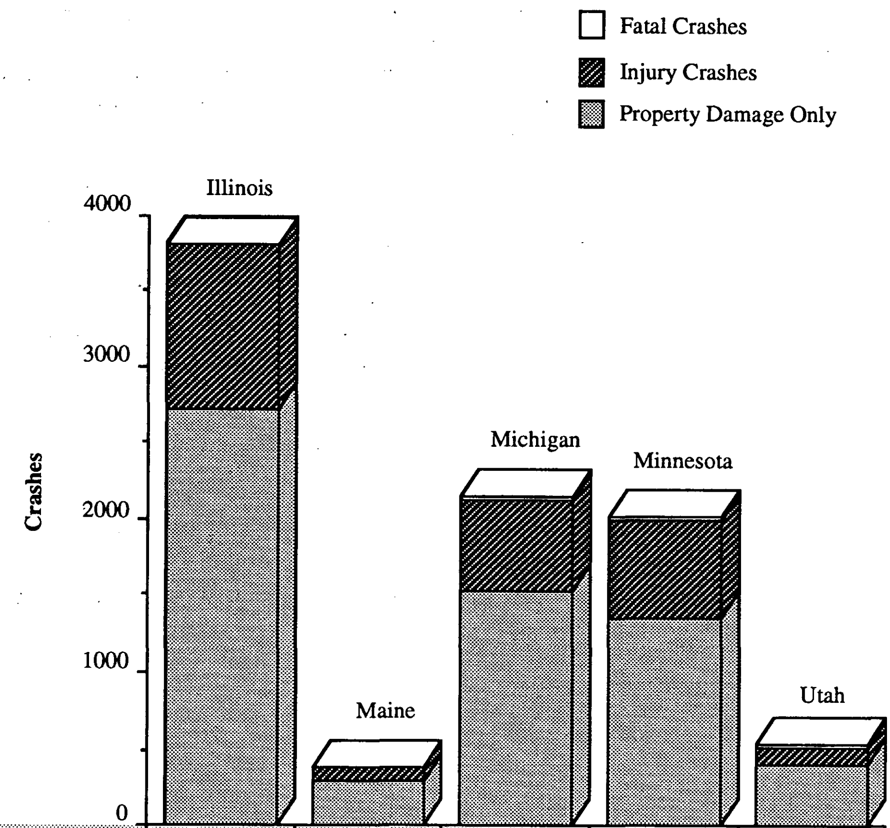
The innocent victim technique was used with data from Michigan and Illinois for bus crashes. To be effective, the technique requires relatively large samples, and the states provided the largest samples of bus-related accidents among the five states. Innocent victims were defined using the best available variables and accident types in the two states. In Michigan, bus innocent victims were identified when the bus driver had "no hazardous action" coded and the other driver had some type of hazardous action coded. This is a very strong definition of an innocent victim. In Illinois, bus innocent victims were identified when a bus was struck in a rear-end collision.

RESULTS OF COLLISION ACCIDENT ANALYSIS

General

A total of 8,897 crashes involving commercial buses was identified from the HSIS files. These bus crashes included 3,825 (43.0 percent) from Illinois (mostly from Chicago), 2,160 from Michigan (24.3 percent), 2,014 from Minnesota (22.6 percent), 526 from Utah (5.9 percent), and 372 from Maine (4.2 percent). A greater number of accidents in a state does not mean that buses are less safe in that state, since no measure of exposure, such as number of buses registered or bus miles driven, is available to normalize these data. Figure 1 shows the distribution of bus crashes by severity of the crash for each of the five states included in the files. Overall, 0.7 percent (65) of the crashes resulted in fatal injury, 28.5 percent (2,537) in nonfatal injury, and 70.8 percent (6,295) in property damage alone. For fatal and injury crashes combined, Minnesota was highest with 32.7 percent and Maine lowest with 22.6 percent. For this study, a serious bus crash was defined as one that resulted in at least one injury or fatality.

Of the total 8,897 crashes, 5,283, or 59.4 percent, were identified as urban crashes. These urban bus crashes had a severity pattern very similar to the overall sample, with 0.5 percent resulting in fatal and 28.3 percent in nonfatal injuries. Most of the urban bus crashes in the HSIS file occurred in Illinois (63.6 percent), with smaller percentages in Minnesota (16.1 percent), Michigan (15.7 percent), and Utah (4.6 percent). The available HSIS data for Maine did not permit the identification of urban crashes.



	Illinois	Maine	Michigan	Minnesota	Utah
Fatal Crashes	23 (0.6%)	1 (0.3%)	18 (0.8%)	17 (0.8%)	6 (1.1%)
Injury Crashes	1,081 (28.3%)	83 (22.3%)	601 (27.8%)	642 (31.9%)	130 (24.7%)
Prop. Damage Only Crashes	2,721 (71.1%)	288 (77.4%)	1,541 (71.3%)	1,355 (67.3%)	390 (74.1%)
Total	3,825	372	2,160	2,014	526

FIGURE 1 Distribution by state of bus crashes in HSIS file.

The analysis of Illinois data comparing bus crashes with other crash types showed that commercial bus crashes represented only slightly more than 0.2 percent of all crashes in Illinois in 1988 to 1989. In comparison, cars and pickup trucks were involved in 87.2 percent of crashes, and large trucks in 6.2 percent. Compared with car and pickup truck crashes, bus accidents are about equally likely to result in a fatality or injury (Figure 2). Truck crashes and events involving other vehicle types were found to have the highest fatality rates.

Temporal

Approximately equal numbers of crashes involving commercial buses were reported for each of the 5 years in the HSIS file. The total number of reported crashes was lowest in 1987 (1,643) and highest in 1985 (1,837) and 1989 (1,838). The number of injury crashes ranged from 499 in 1987 to 568 in 1985. Overall, 29.2 percent of the crashes resulted in injury, with some evidence of a decline in this percentage over the 5-year period.

Injury and overall crashes were lowest in July and August, likely reflecting the reduced number of bus trips and reduced ridership typical during this time. Although the overall number of crashes is greatest in January and February, April and May have the highest percentages of crashes involving injury.

As expected, the percentage of urban crashes on weekends is lower than on weekdays. The distribution of injury crashes is similar to that of total crashes. A higher percentage of crashes occurs on Friday than on other weekdays (significant at the 0.05 level using the chi square test). However, crashes on Friday are less likely to result in injury than on some other days. Traffic volumes may be higher on Fridays, resulting in slower travel speeds, which in turn mitigate accident severity. The greatest percentage of injury crashes occurs on Tuesday.

Frequency of urban bus crashes by time of day generally followed expected bus travel patterns (Figure 3). Crashes were most common during the afternoon rush hours, from 3 p.m. to 6 p.m. (28.3 percent of the total). Another 56 percent of crashes occurred during the morning commute and midday hours, from 6 a.m. to 3 p.m. Although considerably fewer crashes occurred during the evening and night, these tended to be more severe. Nearly 40 percent of bus crashes occurring from 9 p.m. to 3 a.m. resulted in injury.

Environmental Factors

Light Condition

An analysis of bus crashes by light condition on urban streets was based on data from Illinois, Minnesota, and Utah. Accidents were more common during daylight hours (80.3 percent). Lower percentages of crashes occurred after dark on lighted streets (12.3 percent), during dawn or dusk (4.9 percent), or in darkness with no street lights (2.5 percent). These percentages for urban crashes agreed closely with the total sample (rural and urban areas) of bus crashes. The 2-year sample of Illinois data revealed that 78.7 percent of commercial bus accidents occurred in daylight, compared with 68.8 percent of car and pickup accidents and 92.9 percent of school bus accidents.

Urban bus accidents occurring at night on lighted streets had a higher percentage of injury plus fatal accidents (33.8 percent) than did those during daylight (28.3 percent), dawn or dusk (26.1 percent), or dark without lights (25.2 percent). These differences were significant at the 0.05 level. The higher severity of crashes at night on lighted roadways could be the result of the greater use of lighting on high-speed arterial routes, compared with lower-speed collector or local streets.

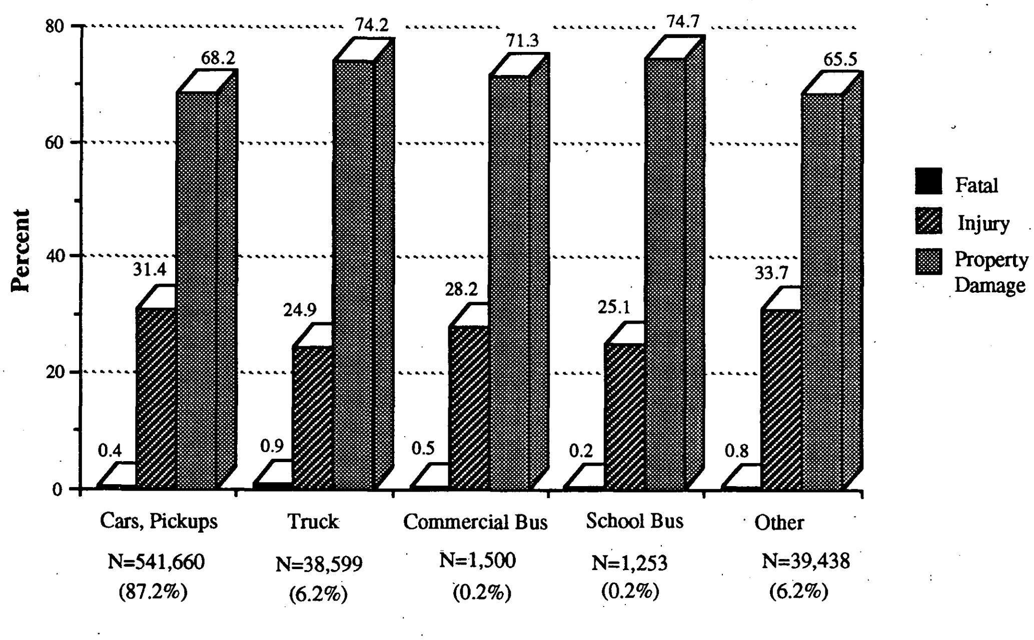


FIGURE 2 Distribution by vehicle type and crash severity of 1988–1989 Illinois crashes. (Note: "Other" includes vans, farm equipment, motorcycles, and vehicles coded as "Other").

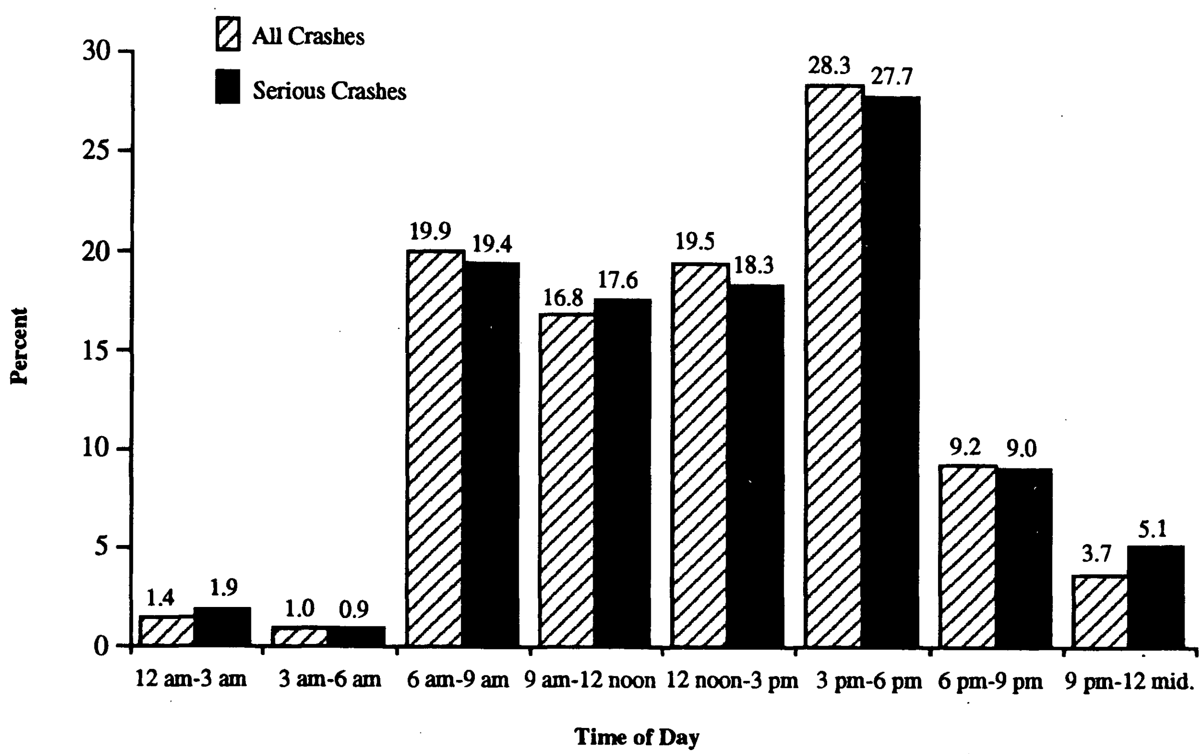


FIGURE 3 Distribution by time of day of urban bus crashes.

Weather and Road Conditions

Of the total bus accidents (urban and rural areas), 64.8 percent occurred on dry pavement compared with 20.7 percent on wet pavement and 13.9 percent on snow and ice. On urban streets, accident percentages were slightly higher on dry pavement (66.2 percent) and wet pavement (22.9 percent), but were lower on ice and snow (10.6 percent). This lower percentage of urban crashes on ice and snow could be related to better snow removal or lower speeds, or both, in urban areas than in rural areas.

Bus accidents in the total sample tended to be more severe on wet roads than on other pavement conditions, with 32.2 percent of wet-road crashes resulting in injury or fatality. This compared with 28.9 percent injury or fatal crashes on dry roads and 26.4 percent on snowy or icy roads. Wet roads are more associated with longer braking distances than are dry roads, which can result in higher-speed impacts (all else being equal). The lower severity on snowy or icy roads could be the result of added driver caution, including reduced travel speeds.

Vehicle Factors

The primary vehicle factor available for analysis from the HSIS crash file was the model year. Model year was analyzed with simple comparisons and with the innocent victim technique. Buses built in 1975 through 1979 were involved in a higher percentage of reported accidents (31.8 percent) in the four states with available data than any other model years. This finding is most likely because of a larger number of these vehicles in service (greater exposure) than other model years. Older buses were also overrepresented in injury and fatal crashes. The injury or fatality rate was almost 6 percentage points lower for buses built after 1984 than for buses built before 1975, and the chi-square statistic for this was significant at the 0.05 level.

Because direct vehicle exposure data were not available, the innocent victim technique was used to account for the relative exposure of the different model years. The analysis revealed that older buses were significantly overinvolved in reported accidents in Illinois ($p = 0.01$); in Michigan, the relationship was marginally significant ($p = 0.10$). A possible explanation for the better performance of newer buses is that changes in bus design through the years, such as better visibility from the driver's seat, power steering, and improved brakes, have had a positive impact.

Driver Factors

Several driver-related factors were analyzed. Driver age was investigated through simple comparisons. In the full five-state sample, drivers near the age of 40 years were involved in many more reported crashes than were other age groups. More than 30 percent of all reported bus crashes involved a bus driver aged 36 to 45. Of course, this finding may be because of a greater number of bus drivers aged 36 to 45, or the large number of miles driven by this age group, or both.

Drivers near the age of 40 experienced more serious crashes than they did all other crashes. In contrast, the proportion of serious crashes for drivers under 35 and over 65 years old was lower than that for all crashes. This finding, which was statistically significant at the 0.005 level, may be because of the route and schedule tendencies of the different driver groups. Younger and older drivers may drive less demanding routes or schedules with fewer passengers, or both. On the other hand, the innocent victim analysis showed that driver age was not related to accident involvement.

Another driver age-related variable analyzed was driver experience, which was recorded only in Utah. No statistically significant differences were found among groups with different amounts of driving experience in terms of involvement in injury and fatal accidents. Note that this variable was driving experience, and not bus driving experience. None of the five states recorded that.

The gender of the bus driver proved to be unrelated to accident involvement. Male bus drivers were involved in almost 80 percent of the crashes in the four states where data were available (Maine did not report driver gender). However, the innocent victim technique showed that there was no strong relationship between driver gender and accident involvement. In addition, there was no statistically significant relationship between driver gender and accident severity ($p > 0.10$). This finding was corroborated with the innocent victim technique.

The bus driver condition reported on the accident form proved to be minor in explaining accidents. Ninety-seven percent of all bus-involved accidents in Illinois, Maine, and Minnesota (where driver condition was reported) had a "normal" bus driver condition recorded. The driver condition recorded for most of the remaining cases was "other" or "unknown." The bus driver was reported to have been drinking alcoholic beverages in only 14 of 5,861 accidents (less than one-fourth of 1 percent). In the 2-year Illinois sample of accidents, a driver was reported to have been drinking in about 3 percent of car and pickup truck accidents as compared with less than 1 percent for drivers in commercial bus-involved accidents.

Roadway Factors

The full bus crash data base showed that there was no traffic control present in about 46 percent of the cases. In other cases, a traffic signal (34.3 percent) or a stop sign (12.4 percent) was present. Bus crashes at traffic signals were more likely to cause injuries and fatalities than bus crashes at stop signs. This difference was significant at the 0.01 level.

Road alignment data for urban streets were collected in Michigan, Minnesota, and Utah. Most collisions (about 95 percent) took place on straight roads. Injuries and fatalities appear to be more likely in accidents on straight roads than on curved roads (29.9 percent versus 20.2 percent), but this finding is based on a small sample of curved roads and should be interpreted with caution.

The 2-year sample of Illinois accidents (comparing buses with other vehicles) revealed that about 55 percent of commercial bus accidents occurred at nonintersections, and the remaining 45 percent occurred at various types of intersections. Relatively similar percentages of car and pickup crashes and school bus accidents happened at intersections. However, only one-third of truck accidents occurred at intersections. Situations that may result in bus accidents at intersections include the following:

- Buses stopping to pick up passengers from stops located at intersections (while the general traffic stream is moving on a normal green phase) and
- Buses entering or leaving curb loading areas (which may not be anticipated by some drivers).

Accident Type

Each of the five HSIS states coded accident type differently. Extra attention was paid to analyzing accident type because this variable reveals patterns of accidents and helps suggest possible countermeasures related particularly to roadway design and bus driver operation.

Figure 4 provides a general accident-type breakdown for all bus-involved accidents in Illinois. Rear-end accidents with one vehicle stopped (probably most often the bus), sideswipe same-direction accidents, and turning accidents were the most common in the sample. Pedestrian and pedalcycle (bicycles, tricycles, etc.) accidents were uncommon, but when they occurred they usually resulted in an injury or fatality. Rear-end accidents, angle accidents, and other accidents (mostly single-vehicle, fixed-object accidents) also had high percentages of injuries and fatalities. Other states showed basically similar patterns.

Results from the 2-year Illinois sample comparing commercial buses with other vehicles helped clarify the general pattern. Commercial bus-involved accidents are more often "sideswipe same-direction" accidents and are less often "rear-end, both moving" accidents, compared with accidents involving other vehicles.

Single-vehicle bus accidents on urban streets (including fixed-object, overturn, and animal accidents, but not including pedalcycle and pedestrian accidents) were not common and resulted in injuries or fatalities less often than other accident types. Only 139 such accidents were reported on urban streets in four states (Illinois, Michigan, Minnesota, and Utah) during the sampled years. Only 27 of those accidents involved an injury, and there were no fatalities. Single-vehicle accidents on urban streets tend to occur more often than multivehicle accidents at night, in the snow and ice, and during right turns.

In multivehicle accidents, buses were more likely to be struck than to strike another vehicle. The 2-year Illinois sample comparing commercial buses to other vehicles showed that

- Commercial buses were struck by automobiles 1,474 times but struck automobiles 1,051 times.
- Commercial buses were struck by trucks 180 times but struck trucks 77 times.
- Commercial buses were struck by other vehicles (not trucks or automobiles) 100 times but struck other vehicles 61 times.

The comparison between buses and trucks, both large vehicles, is revealing. Overall, commercial buses did the striking 1,204 times and were struck 1,769 times; trucks did the striking 40,826 times and were struck 28,885 times. Thus, buses were less likely to be the offending vehicle in bus crashes; trucks were more likely to be the offending vehicle in truck crashes. School buses had a similar accident pattern to commercial buses.

A breakdown of multivehicle bus-involved accidents on urban streets in Illinois revealed some interesting trends. Almost 12 percent of all 3,075 multivehicle accidents in this sample were reported as sideswipe same-direction accidents when the bus was going straight. These accidents may have been the result of buses pulling into and out of curb loading areas. In 84 percent of the angle accidents, the bus was reported to be going straight. For rear-end accidents in which one vehicle was stopped, the bus was coded more often as stopped in traffic rather than picking up passengers, going straight, or stopped for traffic control. The bus was coded turning in about half of the turning accidents. Only 6 percent of the 3,075 multivehicle accidents involved a right-turning bus.

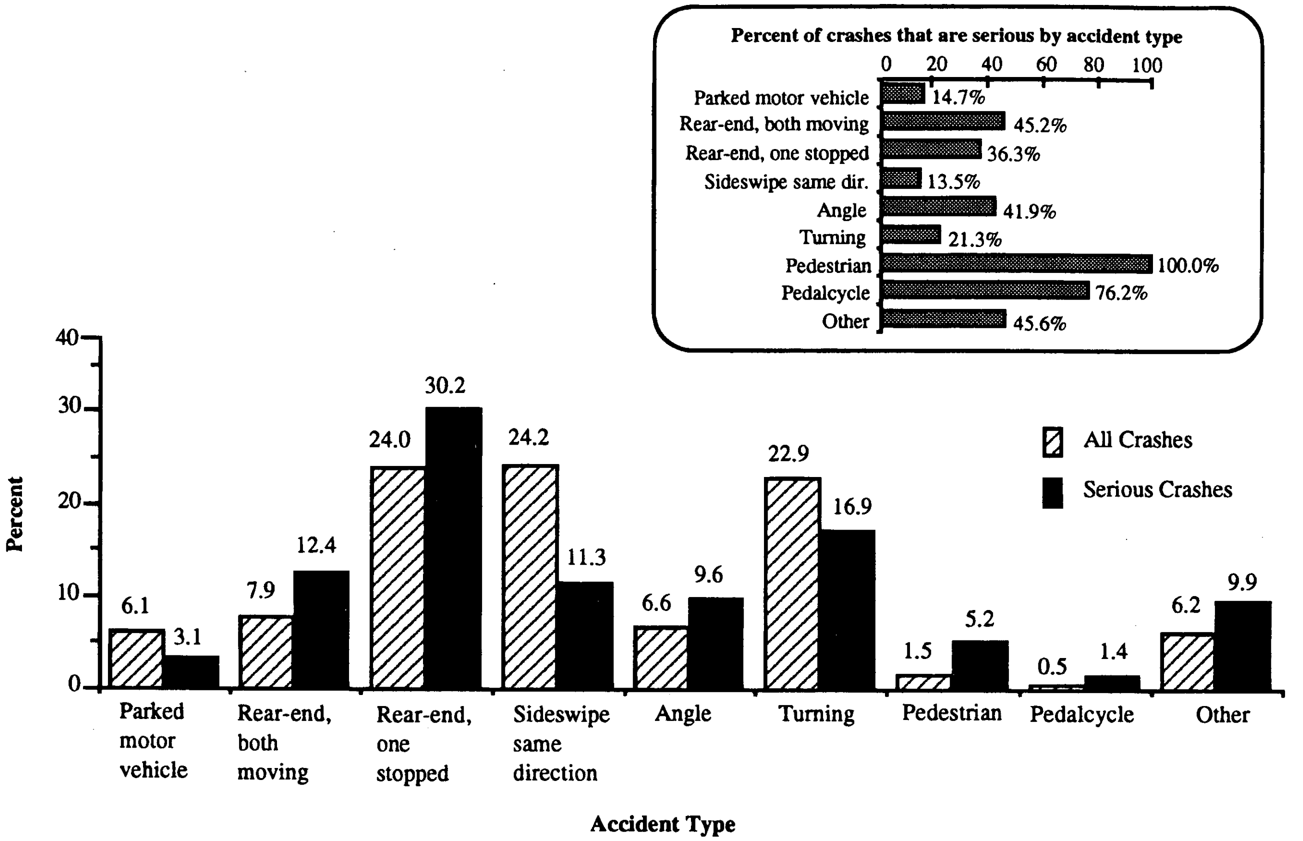


FIGURE 4 Distribution by accident type of bus crashes in Illinois.

Accident type and time of day were significantly related at the 0.005 level. Angle accidents were overrepresented at nighttime, rear-end accidents with one vehicle stopped were overrepresented during morning peak hours, parked vehicle accidents were overrepresented during early afternoon, and sideswipe same-direction accidents were somewhat overrepresented during afternoon peak hours.

Of the 8,897 commercial bus crashes in the HSIS files, pedestrians were involved in 189 (2.1 percent). Nearly all (98.4 percent) of these pedestrian accidents resulted in injuries or fatalities. In fact, 13 accidents (6.9 percent) were fatal. The 2-year Illinois data file showed that 1.2 percent of all commercial bus crashes involved pedestrians, compared with 0.3 to 0.5 percent of other vehicle types. Many of these bus-pedestrian crashes may occur when individuals running to or waiting at a bus stop are struck by an approaching bus or when individuals exiting are struck by a departing bus.

CONCLUSIONS AND RECOMMENDATIONS

This study was carried out to examine the characteristics of crashes involving transit buses (defined in this study as all buses involved in a reported motor vehicle crash except school buses) and to make recommendations for reducing the incidence of bus crashes and related personal injuries. The study included a detailed review of the available literature and an analysis of 8,897 bus accidents in Illinois, Maine, Michigan, Minnesota, and Utah. These crashes became the primary study file. In addition, separate analyses were carried out on a smaller sample of 5,283 crashes (59 percent of the original study sample) identified as occurring on urban streets. The study also examined the characteristics of noncollision bus-related injuries, such as falls while boarding or alighting the bus (*3*).

The analysis was primarily descriptive, involving cross tabulations of selected variables of interest and testing of differences in the resulting distributions. In addition, application of the innocent victim technique allowed some control over exposure differences that might otherwise confound results. Using the Illinois data only, a comparative analysis was conducted comparing bus crashes with other motor vehicle (passenger car, truck, etc.) crashes.

In terms of crash severity, less than 1 percent (0.7 percent) of bus crashes in the overall five-state file resulted in fatal injury; 28.5 percent resulted in nonfatal injury, and the remaining 70.8 percent involved property damage only. The pattern for urban crashes only was similar, with 0.5 percent fatal and 28.3 percent nonfatal injury.

Commercial bus accidents represented less than one-fourth of 1 percent of all motor vehicle crashes occurring in Illinois during 1988 to 1989. Also from the Illinois data, bus accidents and car and pickup accidents were all about equally likely to result in a fatality; however, truck accidents were twice as likely to result in a fatality as accidents involving other vehicle types.

The number of bus crashes is lowest in July and August and highest in January and February. However, winter crashes tend to be less severe, so that the greatest percentage of injury crashes actually occurs in May.

Although the analyses of the 8,897 bus crashes in this study were not in-depth case study investigations, the analyses of many crash factors allow educated judgments of probable causes and develop potential countermeasures corresponding to each probable cause. On the bases of results of the analyses of bus crash factors, the bus safety literature, and decades of highway safety research and experiences on causes and treatments for various crash types, a number of general measures are recommended to reduce the likelihood of bus crashes and resulting passenger injuries. Measures relating to roadway design include the following:

1. Wider intersection turning radii—The analysis showed that rear-end crashes represent one of the most common bus crash types, particularly at intersections. One means of reducing the incidence of rear-end crashes to the bus at intersections is to provide wider intersection turning radii. Because of the length of transit buses, problems may occur when buses turn right at intersections or driveways with a very tight turning radius. This will require the bus to swing wide and often encroach on the oncoming lane of the side street to the right of the bus, which can increase the risk of an accident with an oncoming vehicle from the side street. In addition, with a tight turning radius, the bus must slow down considerably when making such a right turn, and a rear-end crash to the back of the bus can result. By designing or reconstructing the curb radius to be wider, the bus can then make an easier turn without slowing to a near stop and without swinging across the center line as it makes its right turn. This can reduce the risk of rear-end and other crashes involving the bus.

2. Wider lanes on bus routes—Another primary transit bus accident type involves sideswipe collisions between buses and other motor vehicles. Because of the wider vehicle dimensions on buses, it is important that lane widths be adequate to minimize the chance for sideswipe accidents involving vehicles in adjacent lanes. With narrower lanes, the potential for sideswipe accidents is increased, particularly when a bus passes or is being passed by a large truck or other bus. Along major arterials where buses and large trucks are likely to travel, consideration should be given to providing lane widths of 12 ft when possible, or at least 11 ft. This will increase the lateral spacing between buses and other motor vehicles.

3. Turn lanes at intersections along bus routes—The analysis of data from Illinois revealed that 17 percent of bus crashes were turning accidents. Rear-end accidents may occur when adequate separate turning lanes are not available at intersections where buses turn. First, the bus must slow down during right turns and may be rear-ended. When making left turns with no left-turn lane, the bus will often be forced to stop in the left-most through lane and wait for oncoming traffic to clear before turning left into an adequate gap in through traffic. Again, the bus is exposed to the potential for rear-end collisions. For these types of accidents, a potentially effective countermeasure involves adding separate left-turn and right-turn lanes when feasible.

4. Elimination of on-street parking along bus routes—Parked vehicles along bus routes can be associated with several types of bus crashes. These include (a) parked vehicles being struck by the bus, (b) pedestrian accidents as the result of pedestrians stepping or running into the path of the bus from between parked cars, or (c) sideswipe accidents between the bus and other motor vehicles in adjacent lanes (as the result of the bus swerving over the lane line to pass parked vehicles.) To reduce the probability of such accidents, the elimination of on-street parking along selected sections of a bus route is sometimes an effective solution.

5. Adequate paved shoulders or a bus pull-off lane—In suburban and rural areas, some crashes occur when buses stop in the lane to pick up or drop off passengers, thereby resulting in a rear-end collision. Such accidents could be reduced by providing paved shoulders of 8 to 12 ft along such bus routes to allow buses to pull out of the through lane and onto the shoulder to pick up and unload passengers. Where continuous paved shoulders are not feasible, a paved pull-off lane at the bus stop should be considered to allow buses to pull out of the travel lane. Such pull-off lanes are particularly important at bus stop locations where sight distance is severely limited for approaching motorists because of horizontal or vertical alignment. For example, if a bus is stopped in the through lane around a sharp curve, the driver of an oncoming vehicle may not have enough time to see the bus and stop before striking the bus in the rear.

6. Larger traffic signal lenses—The intersection of two roadways is often associated with large numbers of rear-end and turning collisions as a result of conflicting traffic movements. To reduce such intersection collisions involving buses (and other motor vehicles as well), a number of traffic signal–related improvements may be helpful. For example, the use of 12-in. signal lenses instead of the customary 8-in. lenses allows approaching motorists to see the signal more clearly. Vehicles following a transit bus are, therefore, more likely to see a red light and stop behind a bus at the intersection. This is important, because vehicles behind a bus have a limited field of vision of the traffic signal because of the height of the bus and may see the signal of a larger red signal head sooner.

7. Longer clearance intervals—The use of adequate signal clearance intervals can reduce the chance of angle accidents between buses and vehicles at intersections. This is because some intersections are programmed with a minimal amount of yellow time that results in more vehicles running red lights and colliding with vehicles on the cross streets. Angle accidents may be a particular problem for transit buses because of their greater length and greater target area for vehicles coming from cross streets.

8. Separate left-turn phasing—Left-turning buses are involved in accidents more often than right-turning buses. Without left-turn phasing, a left-turning bus must wait in traffic for an adequate gap in oncoming traffic before turning. Under congested conditions, bus drivers may be tempted or forced into making a left turn with an inadequate gap and may be struck by an oncoming through vehicle. This is a particular problem for buses because they are much longer than cars and require a larger gap in traffic to complete a left-turn safely. Separate left-turn phasing stops oncoming traffic, allowing a protected interval for the bus to turn left.

9. Roadway design improvements—Although bus crashes are primarily an urban problem, rural and suburban bus crashes may be reduced by many types of roadway improvements that have shown to be effective for reducing motor vehicle crashes in general. These include providing flatter roadside slopes (to reduce bus rollovers) and clearing roadsides of trees, utility poles, and concrete culverts to the extent possible. Further, guardrail and other roadside hardware should be designed that consider the possibility of bus impacts. On rural roads, adequate widths of lanes (i.e., 11 or 12 ft) and shoulders (paved if possible) and adequate roadway alignment can also be beneficial to bus safety.

10. Improved snow and ice removal—Based on the analysis discussed earlier, bus crashes tend to be more frequent during winter than summer. This may be partly because of the increased snow and ice on the roadways that could contribute to rear-end and other

crashes. Snow removal is a problem in many northern states, but special effort should be made to clear streets of snow and ice promptly along bus routes.

FUTURE RESEARCH NEEDS

The primary future research need is for a study that integrates accident data with high-quality, widespread exposure data. Exposure data could help answer more subtle questions about routes, drivers, and vehicles that could not be answered with the methods of this study. These questions include the following:

- What are the levels of bus exposure (mileage) by bus age, bus type (Interstate versus local transit), type of roadway, and driver factors (age or experience)?
- What types of streets and highways have the highest bus accident rates?
- Are bus accident rates higher at certain times of day and for various types of buses or driver factors?
- What are the effects of specific improvements (routing, bus stop location, geometric and traffic control improvements, etc.) on bus crash rates?

The HSIS data base does not separate local transit and intercity buses. Yet these types of buses are likely to have different levels of exposure and are operated under different conditions. Bus exposure data can be obtained from local transit agency and Interstate bus company records.

The study team also identified several other areas for promising future bus safety research. First, accident data should be obtained from states not included in the HSIS data base. Many states will incorporate a wider range of roadway and weather conditions and increase the sample of bus crashes. This could allow additional conclusions relating accident characteristics to bus crashes and associated injuries.

Another area of needed research involves a more extensive data base, to be obtained from local transit agencies, of noncollision accidents such as falls by passengers. This would allow better comparisons of various bus designs and operating practices on passenger injuries. Research is also needed on accidents in which the bus contributed to an accident but did not collide with persons or other vehicles. For example, pedestrians may step out in front of buses and be struck by passing automobiles. Such accidents would not have appeared in the data base in this study.

ACKNOWLEDGMENTS

This document is disseminated under the sponsorship of the U.S. Department of Transportation, University Transportation Centers Program, through the Southeastern Transportation Center, in the interest of information exchange. The authors appreciate the input and guidance of Roy Field of the Federal Transit Administration throughout this research study.

REFERENCES

1. *General Estimates System 1990.* Report DOT HS 807 781. National Highway Traffic Safety Administration, U.S. Department of Transportation, Nov. 1991.
2. *Fatal Accident Reporting System 1990.* Report DOT HS 807 794. National Highway Traffic Safety Administration, U.S. Department of Transportation, Dec. 1991.
3. Fruin, J., H. F. Huang, C. V. Zegeer, and N. E. Smith. Recommendations for Reducing Noncollision Bus Passenger Injuries. In *Transportation Research Record 1433*, TRB, National Research Council, Washington, D.C., 1994.
4. Dixon, A. H., J. F. Williams, and P. N. Joubert. *Safety Requirements of Bus Seats and Seat Anchorages.* Report CR 25. Department of Transport, Office of Road Safety, Australia, Nov. 1981.
5. Jovanis, P. P., J. E. Schofer, P. Prevedouros, and K. Tsunokowa. *Analysis of Bus Transit Accidents: Empirical, Methodological, and Policy Issues.* Report UMTA-IL-11-0031-89-01. Urban Mass Transportation Administration, U.S. Department of Transportation, June 1989.
6. Johnson, M. A. *Some Characteristics of Buses and Passengers Involved in Stage Carriage Service Accidents, National PSV Accident Survey—Interim Data.* Report K44512. Motor Industry Research Association, Nuneaton, Warwickshire, England, Dec. 1977.
7. Johnson, M. A., T. P. Hutchinson, and J. Zlotnicki. *PSV Accident Survey.* Report K52503, Motor Industry Research Association, Nuneaton, Warwickshire, England, Sept. 1981.
8. Jacobs, G. D., and A. J. Downing. *A Study of Bus Safety in Delhi.* TRRL Supplementary Report 758. Transit and Road Research Laboratory, Crowthorne, Berkshire, England, 1982.
9. Hoffman, E. R. *Low Cost Protection for Bus Passengers.* Department of Transport, Australia, Oct. 1975.
10. Bowman, B. L., and J. E. Hummer. *Examination of Truck Accidents on Urban Freeways. Final Report.* Federal Highway Administration, McLean, Va., June 1989.

This report reflects the views of the authors, who are responsible for the facts and the accuracy of the information presented. The U.S. government assumes no liability for the contents or use thereof.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Estimation of Safety of Four-Legged Unsignalized Intersections

CARL BÉLANGER

In this study, empirical Bayesien methods were applied to the estimation of the safety of four-legged unsignalized intersections. This application can be described as a two-step process. First, multivariate models were developed to estimate the number of accidents from various flow functions at these intersections. The best model was obtained from the product of major and minor flows, raised to a power. Attempts were made to develop models for specific patterns of collisions and to incorporate variables other than traffic flow functions to these models. The modeling results were then combined with the accident count of a four-legged unsignalized intersection to estimate its safety. Results were used to identify blackspot locations and to evaluate the effects of interventions more accurately.

In urban areas, more than half of all accidents occur at intersections, and the corresponding figure for rural environments is about one-quarter (*1*). Given the importance of safety problems at these locations and the relatively small portion of the highway network they represent, interventions aimed at improving the safety of intersections are desirable. To identify sites that have a potential for improvement, knowledge of their long-term mean number of accidents is required. This mean is defined as the "safety" and is denoted by m; its estimate is denoted by $\hat{m}$. Estimates of m are also needed to evaluate the effects of interventions adequately and to determine the success of such actions.

Several methods that have been developed to estimate m are based on the number of accidents observed at the site of interest in a relatively short period. Because of the rarity of accidents and annual variations in the accident count, these estimates are often inaccurate. Also, given that sites are generally chosen for treatment because of a recent poor accident record, $\hat{m}$ is often overestimated. In these situations, the count of accidents in the period after identification will generally revert toward its expected value even if no treatment is applied to the site. This phenomenon, which is called regression to the mean (RTM), introduces significant bias to the conclusions of safety studies (*2–4*). Using longer periods of analysis does not solve the problem because many factors that influence m change over time. Consequently, new methods of analysis were sought. In the past decade, an approach based on empirical Bayesian (EB) methods emerged as a better way of estimating safety. More recently, multivariate statistical analysis was proposed to enhance the benefits of EB methods. We were interested in using these techniques to estimate the safety of unsignalized intersections. To ensure a higher level of homogeneity, a subclass was chosen. Only four-legged intersections that are signed with two stops on minor approaches and have one lane in each direction were selected.

Ministry of Transportation of Québec, 700 René Lévesque Blvd. East, 22nd floor, Québec, Qc., Canada G1R 5H1.

DATA

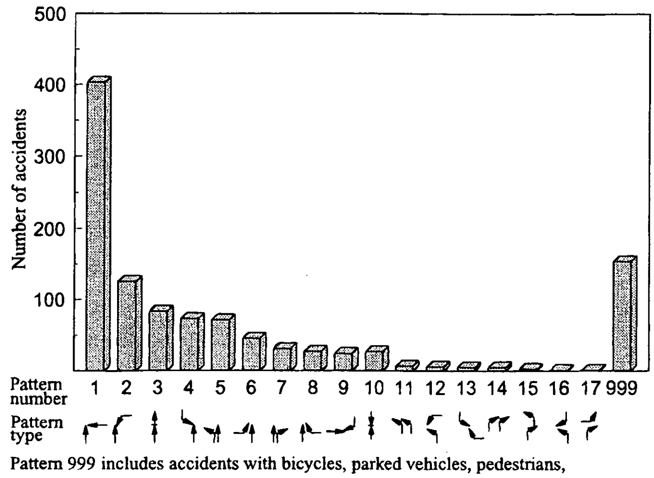
The detailed information about accidents, traffic flow, and geometric characteristics needed for this project was only available for a few sites that had been the object of a safety analysis in recent years. Because these analyses are generally motivated by requests from elected officials to improve sites that are perceived as hazardous, RTM problems are likely to be present. To improve the accuracy of the models, accident data relating to events preceding each site's identification were not used. Only the period following the demand (the number of accidents in the after period is not subject to RTM bias) was considered. With this decision, the establishment of a constant period of analysis became impossible because a sufficient number of accidents could not be gathered during any fixed period. Instead, specific period lengths were determined for each site; consequently, this analysis is based on the number of accidents per day. The sample for this project consists of 149 intersections located in eastern Quebec.

Accidents were considered pertinent if they occurred within 30 m of the intersection or were intersection related. A total of 1084 accidents fulfilled these criteria. Of these, more than 85 percent involved two vehicles. The determination of each pattern of collision and each combination of flows that caused the accident was required, but this information, as coded in the accident file, is unreliable. The list of patterns provided in the accident report form is not exhaustive, and codes are often missing or inconsistent. However, by analyzing the microfilm of each accident report, several of these problems were corrected. Eighteen patterns of accidents were identified (Figure 1). Pattern "999" is a miscellaneous category that includes single-vehicle accidents, accidents with pedestrians, bicyclists, parked vehicles, and reversing vehicles. Right-angle collisions represent 42 percent of the accidents with two or more vehicles.

For each site, a 12-hr count was available, which provided estimates of flows for each of the 12 possible maneuvers at a four-legged intersection: left turn, through, and right turn on each approach. On the basis of data collected from permanent traffic counters, these estimates were converted to daily flow estimates that are representative of an average day, month, and year of the period of analysis. The distribution of flows is shown in Figure 2. It ranges from 388 to 15,942 vpd.

METHODOLOGY

To reduce the regression-to-the-mean bias, EB estimates use not only information from the intersection analyzed but also information from a group of intersections having similar characteristics (called the reference population). The weight attributed to the reference population is a function of its homogeneity. A major diffi-



Pattern 999 includes accidents with bicycles, parked vehicles, pedestrians, vehicles reversing and single vehicle accidents.

FIGURE 1 Number of accidents, by pattern.

culty associated with the use of EB methods consists of defining a reference population that is sufficiently homogeneous to be reliable and yet large enough to improve the estimation. To alleviate this problem, the multivariate approach, as recently proposed by Hauer (*5*) was used. Regression models were first developed to estimate the moments E(m) and VAR(m) that describe the distribution of ms in an imaginary group of intersections having the same characteristics as the site under analysis. Once E(m) and VAR(m) become available, they are combined to the accident history (x) at the intersection of interest to obtain the updated estimate of safety [denoted E($m|x$)] and its variance [denoted VAR($m|x$)].

$$E(m|x) = aE(m) + (1 - a)x$$

$$VAR(m|x) = a(1 - a)E(m) + (1 - a)^2x$$

$$\text{with } a = \frac{E(m)}{E(m) + VAR(m)} \tag{1}$$

Thus, the major task consisted of developing multivariate models to estimate E(m) and VAR(m). In this project, modeling was undertaken in three stages: (a) development of models relating the total number of accidents to various flow functions, (b) development of models relating accidents of a specific type to various flow functions, and (c) evaluation of the effect of variables other than traffic flow.

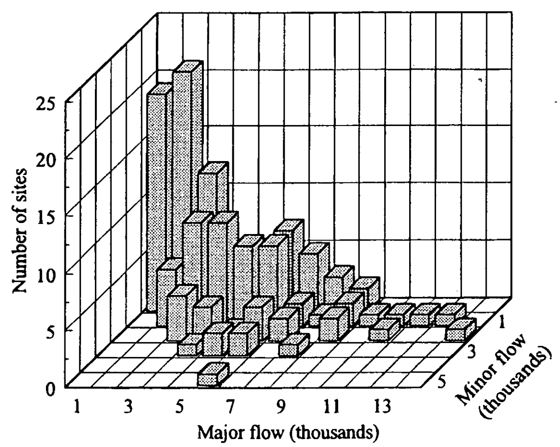


FIGURE 2 Distribution of traffic flow per site.

Most of the regression theory is based on the assumption that the error structure is normal with mean equal to 0 and a constant variance (σ^2); however, this hypothesis is not valid in road safety analysis because residuals tend to increase with larger fitted values. A number of recent studies have concluded that a negative binomial type of error is more appropriate to describe the variations in the number of accidents at several sites. This choice is based on the assumptions that the variations in the number of accidents (x) at any particular location can be described by a Poisson process and that the variations in the levels of safety (m) in a group of similar intersections can be fitted by a gamma distribution (*6*). The GLIM software (*7*) was selected to estimate the coefficients of our models using a negative binomial error structure consistent with the data.

The estimation of E(m) is straightforward because it is obtained directly from the models, but the estimation of VAR(m) from multivariate techniques is less common. It is only recently that a method has been proposed to estimate VAR(m) from the regression results, using the following empirical relationship (*8*):

$$VAR(m) = E(m)^2/k \tag{2}$$

The appropriateness of this relationship was confirmed with the data. As both VAR(m) and k need to be estimated, the process must be iterative, as explained elsewhere (*9*).

GOODNESS OF FIT

In ordinary least-square regression, the coefficient of determination, R^2, is frequently used to express the goodness of fit of a model. It represents the proportion of variation in the observation that is explained by the model and can be calculated in two ways

$$R^2 = 1 - \left(\frac{\text{Unexplained variation}}{\text{Total variation}}\right) \text{ or}$$

$$R^2 = \frac{\text{(Explained variation)}}{\text{(Total variation)}} \tag{3}$$

However, when the variance is not constant (as with the negative binomial distribution), both forms of this equation do not yield identical results, and the R^2 statistic does not constitute a precise estimator of the goodness of fit. Nevertheless, two values of "Pseudo R^2" have been calculated from Equation 3; they constitute a possible range of R^2 values. The difficulty arises in that no equivalent measure of goodness of fit has yet been developed and widely accepted when the error structure is other than normal.

McCullagh and Nelder (*10*) proposed to evaluate the discrepancy of a fit based on the deviance or on the generalized Pearson X^2 statistic. Maycock and Hall (*6*) determined that the expected value of the scaled deviance for a good model having a negative binomial type of error follows a χ^2 distribution with $(n - p)$ degrees of freedom as long as the fitted values are generally larger than 0.5; n is the number of observations, and p is the number of estimated parameters. Larger than expected values of scaled deviance indicate model deficiencies. The appropriateness of adding parameters to a model can be evaluated by comparing decreases in scaled deviance versus decreases in number of degrees of freedom between two models. A decrease in scaled deviance that exceeds the decrease in the number of degrees of freedom justifies the additional complexity of a model. However, when many fitted values are smaller than 0.5, the expected value of the scaled deviance is considerably less than 1 and the χ^2 comparison cannot be used to evaluate the goodness of fit of a model. The Pearson X^2 statistic is calculated from

$$\text{Pearson } X^2 = \sum_i \frac{[x_i - \mathrm{E}(m_i)]^2}{\mathrm{VAR}(x_i)} \tag{4}$$

Miaou et al. (*11*), Bonneson and McCoy (*12*), and Persaud and Dzbik (*13*) evaluated their models on the basis of this statistic, which also follows a χ^2 distribution. McCullagh and Nelder (*10*) mentioned that it may not provide adequate results for limited amounts of data. Hauer (*5*) based model evaluations on the maximum value of the k parameter of Equation 2. Given the relationship between E(m) and VAR(m), models with larger values of k provide a better overall fit because they have a smaller variance. In this research, the evaluation of the adequacy of our models was based on the average behavior of these four indicators: k, scaled deviance, Pearson X^2, and pseudo R^2.

RESULTS

Models for Total Intersection Accidents

In the first stage of regression modeling, relationships between the total number of accidents and various traffic flow functions were explored. To choose functional forms that were coherent with the data, the appropriateness of the selected relationships was verified. The procedure is illustrated with the simple model of the sum of entering vehicles (Q1). A graph of the number of accidents per day versus Q1 was prepared (Figure 3). Sites were ordered in increasing values of Q1 and assembled into groups. Each square on the graph represents an average of 15 sites. The relationship between these two variables is evident: it is almost linear with a hint of downward bend. Regression has been evaluated with the more general functional form of Q1 raised to a power. The best model is

$$\text{Acc/Day} = 3.65 \times 10^{-6} * Q1^{.86} \tag{5}$$

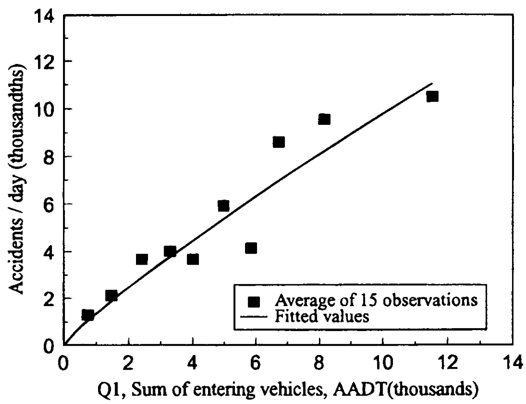


FIGURE 3 Model of total intersection accidents/day versus sum of entering vehicles.

The fitted curve is also shown in Figure 3. A similar approach has been used to develop models that estimate the number of accidents from the product of major and minor flows (F1, F2), the sum of products of conflicting flows (Q2), and the sum of weighted products of conflicting flows (Q3). Although various definitions have been proposed to describe the notion of conflicting flows, no agreement has yet been reached as to its best representation (*14,15*). In the determination of a conflict index, the concern was to ensure that the proposed function provides an adequate representation of accident occurrence in the population of sites. Accordingly, the conflict index consists of the sum of the products of each combination of flows that is involved in ten accident patterns responsible for 95 percent of all collisions involving at least two vehicles. Q2 is calculated from

$$Q2(l) = \sum_{i=1}^{10} \sum_{j=1}^{p} \left[F1(i, j, l) * F2(i, j, l) \right] \tag{6}$$

where

- i = pattern number,
- j = approach number,
- l = intersection number, and
- p = number of occurrences of the same pattern at each intersection.

Because different patterns of accidents have different probabilities of occurrence, a model that takes into consideration the relative risk of a maneuver is likely to provide a better fit than a model allocating the same weight to all products of conflicting flows. To test this hypothesis, weighting indexes (WI) have been calculated. These weighed indexes are obtained by dividing the total number of accidents of a given pattern by the corresponding sum of products of contributing flows

$$WI(i) = \frac{\sum_{j=1}^{p} \sum_{l=1}^{149} \mathrm{acc}(i, j, l)}{\sum_{j=1}^{p} \sum_{l=1}^{149} F1(i, j, l) * F2(i, j, l)} \tag{7}$$

Values of the weighting indexes are shown in Table 1. They range from 0.08 for rear-end collisions to 4.19 for right-angle collisions. The flow function Q3 becomes

$$Q3(l) = \sum_{i=1}^{10} \sum_{j=1}^{p} \frac{WI(i)}{p} * [F1(i, j, l) * F2(i, j, l)] \tag{8}$$

Results of total intersection accident models are summarized in Table 2. Although ranking obtained from each goodness-of-fit indicator is unique, the product of major and minor flows is generally identified as the best functional form.

Modeling by Type of Accidents

A logical approach to modeling consists of relating accidents to the traffic flows that cause the impact. In our population of sites, right-angle collisions (Pattern 1) account for 42 percent of all collisions involving two or more vehicles, and a specific model has been developed for this accident pattern. The pattern second in importance is angle collisions between a through vehicle and a left-

TABLE 1 Weighting Factors, by Patterns

Pattern Number	Pattern Type	Number of accidents	Weighting indexes
1		403	4.19
2		125	1.08
3		83	0.08
4		73	1.00
5		71	0.91
6		45	0.49
7		31	0.25
8		27	0.35
9		24	0.21
10		21	0.46
Total		903	

turning vehicle; 125 collisions of this type have been coded. As causes of these collisions differ depending upon whether the left-turning vehicle is located on the minor or the major approach, these accidents were subdivided into two groups. The resulting subsets were too small to allow the determination of logical relationships based on observed trends of the data, and the goodness of fit was reduced. Instead of proposing several "intuitive models" that would present a poor fit for more than half the data, only right-angle collisions were analyzed in detail, and all remaining collisions were grouped into one aggregate model. The resulting tool to evaluate the safety of an intersection consists of two models: a right-angle model and a remaining patterns model.

Estimation of Contribution of Additional Features

To assess whether variables other than traffic could make a significant contribution to the explanation of accident occurrence, factors were added to the models; these are dummy variables that take distinct integer numbers for each specific value of a variable. For example, to evaluate the influence of flashing beacons on the observed number of accidents, a two-level factor is created: 0 for intersections with flashers and 1 for intersections without flashers. If the coefficients of each value of the factor are different, it means that the factor has an influence on the total number of accidents. The following functional form was used to examine the effect of flashing beacons, sight distance, turning lanes, and speed:

$$\text{Acc/day} = b_0 * F1^{b1} * F2^{b2} * e^{(\text{Factor})} \tag{9}$$

For example, the value of the "flasher factor" for intersections that are equipped with this warning device is 0.17, which indicates that at the same flow these junctions are expected to have 19 percent more accidents than intersections without such a device. However, given the magnitude of the standard error of this coefficient (0.15), the effect is uncertain. As shown in Table 3, similar results were obtained for sight distance, turning lanes, and speed. It should be remembered that regression equations provide relationships that are associative and not causative. That intersections with flashers have on average more accidents does not necessarily mean that beacons reduce the safety. Instead, it could be that they are generally installed at intersections that have a poorer safety performance and that they do not succeed in making these junctions as safe as other similar sites.

Whenever feasible, it is better to assess the effect of a variable by the development of distinct models for each level of a factor. With this data, it was possible to do so for the maximum posted speed at intersection approaches. Models were developed for the 50 and 90 km/hr speed limits. Results are summarized in Table 4.

TABLE 3 Effect of Causal Factors

Characteristic	Factor	Value	Standard error	k
Flashing beacon	1: no 2: yes	0.00 0.17	- 0.15	3.0
Sight distance	1: <100m 2: 100-200m 3: 200-300m 4: >300m	0.00 0.41 0.17 0.45	- 0.24 0.25 0.23	3.3
Turning lanes	1: 2 lanes 2: 2 + RT 3: 2 + LT 4: 2 + LT + RT	0.00 0.10 0.25 0.21	- 0.16 0.20 0.22	3.1
Speed limit	1: 50 Km/hr 2: 90 Km/hr	0.00 0.17	- 0.15	3.6

TABLE 2 Models for "Total Intersection Accidents"

Functional form	k	Deviance (d.f.)	X^2	R^2
$\text{Acc/Day} = 3.65*10^{-6} * Q1^{.86}$	2.50	168.45 (147)	144.01	.42, .50
$\text{Acc/Day} = 5.59*10^{-6} * F1^{.42} * F2^{.51}$	2.95	164.64 (146)	135.88	.47, .56
$\text{Acc/Day} = 4.41*10^{-5} * Q2^{.36}$	2.05	164.75 (147)	146.32	.33, .53
$\text{Acc/Day} = 3.14*10^{-5} * Q3^{.45}$	2.80	166.14 (147)	147.04	.50, .50

TABLE 4 Summary of Results

	Main road speed limit 50 km/hr	Main road speed limit 90 km/hr	All speeds limits
Total intersection models			
a1) Acc/Day=b_0 * Q^{b1}	b_0=4.59E-6 b1=0.83 k =2.70	b_0=4.42E-6 b1=0.83 k =2.70	b_0=3.65E-6 b1=0.86 k =2.50
a2) Acc/Day=b_0 * $F1^{b1}$ * $F2^{b2}$	b_0=1.07E-5 b1=0.34 b2=0.49 k =3.10	b_0=3.37E-6 b1=0.41 b2=0.59 k =5.10	b_0=5.29E-6 b1=0.42 b2=0.51 k =2.95
Pattern models			
b1) Pattern 1 (Right Angle): Acc/Day=b_0 * EXP(b_1 * F1) * $F1^{b2}$ * $F2^{b3}$ (50 km/hr and "all speeds") Acc/Day=b_0 * $F1^{b1}$ * $F2^{b2}$ (90 km/hr)	b_0=2.05E-6 b1=-3.69E-4 b2=0.57 b3=0.46 k =1.95	b_0=6.14E-6 b1=0.32 b2=0.43 k =1.40	b_0=1.09E-5 b1=-6.52E-5 b2=0.26 b3=0.46 k =1.50
b2) Remaining Patterns: Acc/Day=b_0 * $F1^{b1}$ * $F2^{b2}$	b_0=1.97E-6 b1=0.59 b2=0.36 k =3.30	b_0=1.57E-6 b1=0.57 b2=0.45 k =6.20	b_0=1.42E-6 b1=0.65 b2=0.35 k =2.80

ESTIMATION OF E($m|x$) AND VAR($m|x$)

With these models it is now possible to estimate the safety of a four-legged unsignalized intersection when its traffic flow and accident history are known. This estimation consists of two steps

1. Estimation of E(m) and VAR(m) with the multivariate models, and
2. Estimation of E($m|x$) and VAR($m|x$) from Equation 1.

The application of the method is shown with the following example. Suppose that the traffic flows and accident count at a four-legged unsignalized intersection are as indicated in Figure 4. Depending on the availability of the data, different models can be selected to estimate its safety. When the number of collisions by pattern and traffic flow estimates per movement are available, Models b1 and b2 of Table 4 can be used. If instead only the major and minor flows are known, Model a2 should be used. In this case, the calculation is as follows:

Step 1: Estimation of E(m) and VAR(m)

$$\mathrm{E}(m) = (1.07 * 10^{-5} * 4500^{.34} * 2000^{.49}) * 1095$$
$$= 7.74 * 10^{-3} * 1095$$
$$= 8.48 \text{ acc/3 years}$$
$$\mathrm{VAR}(m) = [(7.74 * 10^{-3})^2/3.10] * 1095^2$$
$$= 23.17 \text{ (acc/3 years)}^2$$

Step 2: Estimation of E($m|x$) and VAR($m|x$)

$$\mathrm{E}(m|x) = a\mathrm{E}(m) + (1 - a)x$$
$$\text{with } a = \mathrm{E}(m)/[\mathrm{E}(m) + \mathrm{Var}(m)]$$
$$= 8.48/(8.48 + 23.17)$$
$$= 0.27$$
$$= (0.27 * 8.48) + [(1 - 0.27) * 15]$$
$$= 13.24 \text{ acc/3 years}$$
$$\mathrm{VAR}(m|x) = a(1 - a)\mathrm{E}(m) + (1 - a)^2x$$
$$= (0.27 * (1 - 0.27) * 8.48) + [(1 - 0.27)^2 * 15]$$
$$= 9.67 \text{ (acc/3 years)}^2$$

In this example, the estimate of safety is reduced from 15 to 13.24 acc/3 years, which corresponds to a RTM correction of 12 percent. The larger the difference between the number of accidents at the site and the expected value of the reference population, the larger the correction. Once these estimates are made available, two major tasks can be accomplished: identification of entities that require intervention and evaluation of the effects of road safety interventions.

IDENTIFICATION OF DEVIANT SITES

A site is selected when the difference between its safety and the safety of sites having similar characteristics is unacceptable.

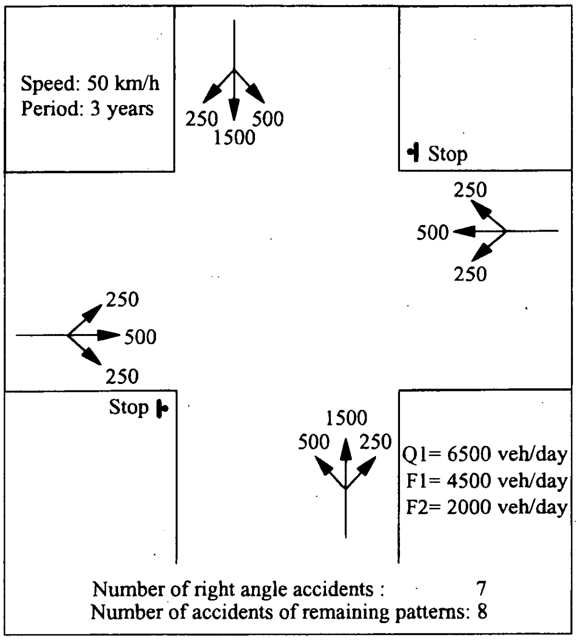


FIGURE 4 Numerical example.

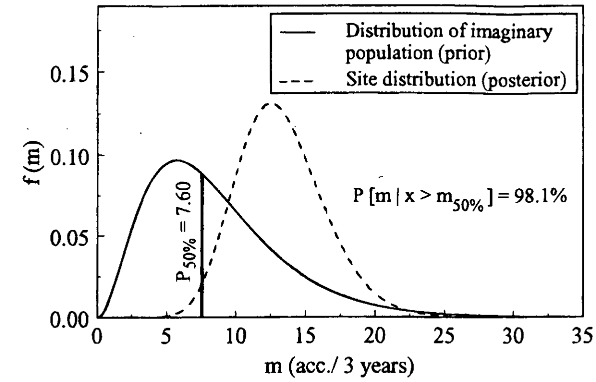


FIGURE 5 Gamma distributions, prior and posterior estimates of safety.

The determination of what is unacceptable should be a function of the resources allocated to the correction of deviant sites. The process is as follows:

- Estimate E(m) and VAR(m) from the multivariate models and plot the probability density function (pdf) of the reference population (gamma distribution).
- On the basis of this pdf, determine the value of m to be used as a point of comparison; the use of the median of the reference population ($P_{50\%}$) is recommended.
- Estimate E($m|x$) and VAR($m|x$) from Equation 1 and plot the pdf of the intersection of interest (gamma distribution).
- On the basis of this pdf, calculate the probability of the $m|x$ of this intersection being larger than the median of the reference population and decide whether the site is deviant.

The previous example is continued to illustrate this method. It has been estimated that E(m) = 8.48 acc/3 years and VAR(m) = 23.17 (acc/3 years)^2. When the density function is plotted one finds that the median of the distribution of ms for an imaginary group of four-legged unsignalized intersections having a major flow of 4500 vpd and a minor flow of 2000 vpd is 7.60 acc/3 years. It was also estimated that intersections with this flow combination that had 15 accidents in the last 3 years have E($m|x$) = 13.24 acc/3 years and VAR($m|x$) = 9.67 (acc/3 years)^2. When the corresponding pdf is plotted one finds that there is only a 1.9 percent probability for this intersection to have m smaller than 7.60 acc/3 years. In other words, there is a 98.1 percent chance that this intersection is less safe than 50 percent of intersections having similar characteristics; consequently, it is selected for treatment. The result is illustrated in Figure 5.

To facilitate the use of the method, a computer program has been developed that estimates the safety of these intersections and identifies blackspots. Information on this program is available from the author.

EVALUATION OF EFFECTS OF INTERVENTIONS

To estimate an intervention's effect on safety, an index of effectiveness (IE) must be calculated. It corresponds to the following ratio:

$$\text{IE} = \frac{\text{Safety in the after period}}{\text{Safety that would have been after, without intervention}} \tag{10}$$

The numerator and denominator of this equation need to be estimated adequately. When the number of accidents is large enough to minimize the effect of random variations, the count of accidents in the after period is a good estimate of the safety after treatment. The estimate of what would have been the safety of the entity in the after period if the intervention had not been implemented is more difficult to obtain because it corresponds to a quantity that cannot be observed directly. A commonly used estimator of the denominator of Equation 10 is the observed number of accidents in the period preceding the intervention, but it often leads to an overestimation of the benefits of our actions. To improve the accuracy of the denominator, two questions must be answered.

1. What was the safety of the entity before treatment?
2. How would the estimate of safety in the before period have changed between the before and after period if the intervention had not been implemented?

The safety before treatment at four-legged unsignalized intersections should be estimated from the multivariate models and the knowledge of the number of accidents at the site, as shown in the previous example. Between the before and the after period, several factors are likely to have changed and to have modified the level of safety at the site: traffic, weather, economy, and so forth. The influence of some factors is unknown and cannot be estimated, but it is important to calculate the effect of factors whose influence is known. For example, the impact of modifications in traffic flows can be estimated from the models. The previous example is continued to illustrate the method.

At the same intersection, 11 accidents have been recorded in a 3-year period following its treatment. In the same period, the average daily traffic increased from 4500 to 5000 vehicles/day on the major street and from 2000 to 2500 vehicles/day on the minor street. The best estimate of safety in the after period is 11 acc/3 years. Earlier it was found that the estimate of the safety of the intersection before treatment is 13.24 acc/3 years. To correct for changes in traffic, use is made of the Model a2 (Table 4). With the after flows, one would expect 9.81 acc/3 years, which represents an increase of 16 percent compared with the original level of traffic. Assuming that only traffic flow changes can be taken into consideration, the estimate of what would have been the safety of the entity in the after period without intervention is equal to $(13.24 * 1.16) = 15.36$ acc/3 years. Accordingly, the index of effectiveness is

$$\begin{aligned} IE &= 11/15.36 \\ &= 0.72 \end{aligned}$$

In other words, the treatment at this intersection is estimated to be associated with a 28 percent reduction in accidents. However, results of similar interventions at several intersections are required to increase the accuracy of this estimated effect.

SUMMARY AND CONCLUSIONS

In this study, multivariate models have been developed that can be used to estimate the safety of four-legged unsignalized intersections in an EB framework. Multivariate models are used to estimate the moments $E(m)$ and $VAR(m)$ of an "average" intersection; this information is then combined with the count of accidents (x) at a specific intersection to calculate its updated estimate of safety, as expressed by $E(m|x)$ and $VAR(m|x)$.

This study confirms the applicability of methodological elements proposed in recent research. The negative binomial error structure was shown to be consistent with the data. Also confirmed by the data is the useful empirical relationship between $E(m)$ and $VAR(m)$; that is $VAR(m) = [E(m)]^2/k$.

Total intersection models and pattern models for three categories of speed were developed: 50 km/hr, 90 km/hr, and all speeds. The 50 and 90 km/hr models are more precise than the all speeds models and should be used whenever possible. When only the total number of accidents and entering vehicles on each approach is known, total intersection models must be used. However, when accidents by pattern and traffic volumes by movements are available, the use of pattern models is preferred. They constitute a more powerful tool of analysis because they can provide a detailed identification of abnormal situations. For example, a site could have a total number of accidents not significantly higher than the average total for similar sites but show an abnormal frequency of right-angle collisions.

In practice, both the total intersection models and pattern models are of interest. Given that it is unlikely that accidents by pattern and detailed traffic flow estimates will be available on a large scale in the near future, total accident intersection models could be used as a first sieve. Data requirement is not as extensive as with pattern models and allows for a wider number of intersections to be considered initially. Detailed information could then be collected on the reduced sample to make possible the use of more precise pattern models.

ACKNOWLEDGMENTS

This research has been made possible by the support of the Ministry of Transportation of Quebec. We are grateful for the collaboration of Ezra Hauer, who supervised the work.

REFERENCES

1. Hodge, G. A., and A. J. Richardson. The Role of Accident Exposure in Transport System Safety Evaluations I: Intersection and Link Site Exposure. *Journal of Advanced Transportation*, Vol. 19, No. 2, 1985, pp. 179–213.
2. Abbess, C., D. Jarrett, and C. C. Wright. Accidents at Blackspots: Estimating the Effectiveness of Remedial Treatment, with Special Reference to the "Regression-to-Mean" Effect. *Traffic Engineering and Control*, Vol. 22, No. 10, 1981, pp. 535–542.
3. Brüde, U., and J. Larsson. *The "Regression-to-Mean" Effect. Some Empirical Examples Concerning Accidents at Road Junctions.* VTI Report 240, 1982.
4. Hauer, E., and B. N. Persaud. Common Bias in Before and After Accident Comparisons and its Elimination. In *Transportation Research Record 905*, TRB, National Research Council, Washington, D.C., 1983, pp. 164–174.
5. Hauer, E. Empirical Bayes Approach to the Estimation of "Unsafety": the Multivariate Regression Method. *Accident Analysis and Prevention*, Vol. 24, No. 5, 1992, pp. 457–477.
6. Maycock, G., and R. D. Hall. *Accidents at 4-Arm Roundabouts.* TRRL Laboratory Report 1120. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1984.
7. Baker, R. J., and J. A. Nelder. *The GLIM System—Release 3.* Rothamsted Experimental Station, Harpenden, U. K., 1978.
8. Hauer, E., and B. N. Persaud. How to Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effect of Warning Devices. In *Transportation Research Record 1114*, TRB, National Research Council, Washington, D.C., 1988, pp. 131–140.
9. Hauer, E., J. C. N. Ng, and J. Lovel. Estimation of Safety at Signalized Intersections. In *Transportation Research Record 1185*, TRB, National Research Council, Washington, D.C., 1988, pp. 48–61.
10. McCullagh, P., and J. A. Nelder. *Generalized Linear Models.* The University Press, New York, 1983.
11. Miaou, S. P., P. S. Hu, T. Wright, A. K. Rathi, and S. C. Davis. Relationship Between Truck Accidents and Highway Geometric Design: a Poisson Regression Approach. In *Transportation Research Board 1376*, TRB, National Research Council, Washington, D.C., 1992, pp. 10–18.
12. Bonneson, J. A., and P. T. McCoy. Estimation of Safety at Two-Way Stop-Controlled Intersections on Rural Highways. In *Transportation Research Record 1401*, TRB, National Research Council, Washington, D.C., 1993, pp. 83–89.
13. Persaud, B., and L. Dzbik. Accident Prediction Models for Freeways. In *Transportation Research Record 1401*, TRB, National Research Council, Washington, D.C., 1993, pp. 55–60.
14. Surti, V. H. Accident Exposure for At-Grade Intersections. *Traffic Engineering*, Vol. 36, No. 3, 1965, pp. 26–27.
15. Peleg, M. Evaluation of the Conflict Hazard of Uncontrolled Junctions. *Traffic Engineering and Control*, Vol. 9, No. 7, 1967, pp. 346–347.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Effect of Resurfacing on Safety of Two-Lane Rural Roads in New York State

EZRA HAUER, DONALD TERRY, AND MICHAEL S. GRIFFITH

In the early 1980s, two kinds of resurfacing projects were undertaken in New York State: Fast track projects involving only resurfacing and reconditioning and preservation (R&P) projects in which roadside and roadway safety improvements have been incorporated with resurfacing. The question was whether following resurfacing the fast track projects (226.7 mi) perform less well, from a safety viewpoint, than the R&P projects (137.2 mi). Findings indicated that in fast track projects safety initially declined, but in R&P projects safety improved. Another conclusion that emerges from this work is that, within the first 6 to 7 years of pavement life, safety improves as the pavement ages. The Empirical Bayes approach to the study of the safety effect has been used. Two methodological innovations may be of interest. First, because the safety effect of resurfacing changes as the pavement ages, it was necessary to find a way to examine changes in safety as a function of time. Second, the accuracy of studies of this kind is often limited by the sparsity of accident data. The method used here allows the use of a long "before" accident history to enhance estimation accuracy.

The effect of resurfacing on the safety of rural two-lane roads was in the eye of a stormy debate in the late 1980s. To clarify the issue, a special TRB study was initiated, culminating in the publication of *Designing Safer Roads* (*1*). A critical review of published evidence conducted for this study by Cleveland (*2*) concluded that, although there is diversity in the findings of the few extant studies, the detrimental effect of resurfacing on safety, if any, is likely to be small.

In the State of New York Department of Transportation (NYDOT) as in all states, road resurfacing is an ongoing activity. In the early 1980s, two kinds of resurfacing projects were undertaken.

- Projects involving only resurfacing are called fast track. These consist of simple resurfacing and restriping. Initially, they did not include shoulder preparation or backing up, replacing guardrail, cutting trees, or other work. These activities were to be done later by the maintenance forces. After a few years, the scope of the fast track projects was enlarged to allow maintenance to catch up.
- Projects in which roadside and roadway safety improvements have been incorporated with resurfacing are called Reconditioning and Preservation (R&P) projects. In addition to resurfacing these may include limited pavement reconstruction and remedies to safety or operational problems. Superelevation, shoulder, drainage, slope flattening, and guide-rail and roadside improvements (removing or relocating fixed objects) are typically included.

E. Hauer, Department of Civil Engineering, University of Toronto, Toronto, Ontario M5S 1A4, Canada. D. Terry, Traffic Engineering and Safety Division, State of New York Department of Transportation, Albany, N.Y. 12232. M.S. Griffith, Office of Safety and Traffic, FHWA, 6300 Georgetown Pike, McLean, Va. 22101.

A before-and-after comparison indicated that there might be a substantial difference between the safety performance of these two kinds of resurfacing projects. To check whether the difference is real, additional data were collected for comparison sites. However, questions still remained about the appropriateness of the comparison groups selected, about a possible regression-to-mean bias, and about the statistical significance of the results. Eventually, the FHWA was asked to assist in resolving the issue whether projects involving only simple resurfacing perform less well, from a safety viewpoint, than similar resurfacing projects where roadside and roadway safety improvements have been incorporated.

This paper is the product of that request. The main aim is to add what has been found for the kinds of treatments used in New York state in the early 1980s to the store of facts about the safety effect of resurfacing. In performing the work, some methodological innovation was required and will be described without burdening the exposition with too much theory. Full details are given in the original report (*3*).

DATA

All data pertain to rural, two-lane, undivided, free-access road sections. The following information has been assembled by officials of NYDOT for each road section:

- The length of the section and the number of intersections in it.
- Traffic counts for the 13 years from 1975 to 1987, factored to represent the AADT in the year of the count.
- The count of fatal, injury, property damage only, fixed object, and intersection accidents for each month of the 13 years from 1975 to 1987.
- If the road section was resurfaced, the month and year in which construction started and ended (mostly in the 1981 and 1982 construction seasons).

The data pertain to 82 fast track projects (226.7 mi, 2.09 intersection/mi), 55 R&P projects (137.2 mi, 4.36 intersections/mi), and 525 comparison and reference road sections (2193.2 mi with 1.92 intersections/mi). During preliminary analysis, a few suspicious traffic volumes, intersection densities, and accident records were identified. Where possible, these were checked and corrected. If verification or correction was not feasible, the data were not used.

PRELIMINARIES

Before analysis could begin, several preparatory activities had to be undertaken.

Estimating AADT

Because traffic and accidents are closely related, changes in traffic have to be accounted for when changes in safety are examined. The accident data to be analyzed are for the 13-year period from 1975 to 1987. Thus, estimates of AADT for the same period were needed.

Traffic counts are conducted on each road section every few years. These counts are then factored to represent the AADT in the year of the count. The task was to fill in all the blanks for the years for which there were no counts. This estimation procedure is described in detail in Appendix A of the work by Hauer (*3*). At the conclusion of this task, estimates of the AADT and its standard deviation for all 13 years for each of the 662 sites were obtained.

Accident Counts

The following questions required answers:

1. What were the effects of changes in 1978 and again in 1985 in the reporting threshold for property damage only accidents?
2. What was the effect of the change that occurred in May 1993 in the definition of an intersection accident?
3. Is the year-to-year variation the same for all accident types to be analyzed?
4. What is the month-to-month variation by accident type?

In response to Questions 1 and 2, no discernible effect could be found for either the changes in the reporting threshold or the definition of what constitutes an intersection accident (*3*). On Question 3, each accident type was found to have its own year-to-year variation. Therefore, when modeling how the expected number of accidents depends on time and traffic, discrete parameters have to be assigned to each year and accident type.

Inasmuch as the safety effect will be estimated as a function of the number of months after the end of construction, information about month-to-month variation is also needed and has been estimated.

Examination of Comparison and Reference Group

In before-and-after studies, the role of a comparison group is to account for changes in safety from the before to the after period that are due to a variety of uncontrolled factors (weather, accident-reporting threshold, driver demography, vehicle fleet, definition of intersection accident, etc.). Therefore, the requirement is that the change in these uncontrolled factors and their effects on safety be the same on the entities in the comparison group and on the treated entities. (Note that here the comparison group is not used to account for changes in traffic flow inasmuch as this can be done better using the available estimates of the AADT.)

A reference group in before-and-after studies serves mainly to account for any bias due to regression to the mean. The requirement is that the expected number of accidents of a treated entity with given traits (geometry, traffic flow) be the same, roughly, as the expected number of accidents of a reference-group entity with identical traits. Although the purpose and use of the comparison group and the reference groups are different, there is no reason why the same group of entities can not serve in both roles, provided that both requirements are met.

It was established earlier that one of the differences among the fast track, R&P, and the comparison and reference road sections was that their intersection densities are very different (2.09, 4.36, and 1.92 intersections/mi, respectively). It was therefore clear that intersection and nonintersection accidents have to be modeled separately.

The 525 road sections which make up the reference and comparison group are composed of 47 sections originally selected for their proximity to fast track projects, 49 sections originally selected for their proximity to R&P projects, and 429 sections selected to represent the entire state. Therefore, the comparison group requirement is likely to be satisfied. To examine the suitability of these road sections as a reference group, nonintersection accidents per vehicle-mi and intersection accidents per intersection-vehicle of treated and not-treated road sections were compared in the 1975 to 1980 period (i.e., before any construction took place). On the basis of such comparisons and in view of the overall similarity in the average rate and its large year-to-year fluctuations, it was concluded that there was no reason to distinguish between the three groups of road sections. Therefore, they were used jointly as one reference group.

In summary, the fast track sites differ from the R&P sites in intersection density. This required modeling intersection and nonintersection accidents separately. Once this was done, all untreated sites were combined to serve as reference group and as comparison group.

Multivariate Modeling

To account for the effect of changes in uncontrolled factors (weather, reporting threshold, demography, etc.), for the effect of changes in AADT, and also for the possible bias due to regression to mean, the parameters of a set of multivariate models need to be estimated. These models link accident counts from 1975 to 1987 to traffic in those years and to variables representing the passage of time. (Details are given in Hauer (*3*), Appendix D). Three models were estimated for three accident types: (a) nonintersection accidents, (b) intersection accidents, and (c) fixed-object accidents.

The models are of the form

$$E(m_{i,y}) = \alpha_y F_{i,y}^{\beta}$$

$$\mathrm{VAR}(m_{i,y}) = [E(m_{i,y})]^2/b \tag{1}$$

where

$m_{i,y}$ = what would be the average number of accidents per mi (or per intersection) of road section i in year y if it were possible to freeze all relevant conditions of year y and repeat them a very large number of times. If there were another road section j, with the same traffic as section i, in the same state, same number of lanes, and so forth, still $m_{i,y} \neq m_{j,y}$ because the two road sections will differ in many traits that are unmeasured and are not featured in the model.

$E(m_{i,y})$ = average of the $m_{i,y}$'s for an imaginary set of road sections that have exactly the same measured and modeled traits (including traffic) as section i.

$\mathrm{VAR}(m_{i,y})$ = variance of these $m_{i,y}$'s.

α_y = parameter for year y that captures the influence of all factors that change from year to year, except for the change in traffic flow. Values are estimated for $\alpha_1, \alpha_2, \ldots, \alpha_{13}$, in which the subscript 1 is for 1975, 2 for 1976, . . . , 13 for 1987.

$F_{i,y}$ = AADT for road section i and year y,

β = manner in which $E(m_{i,y})$ is thought to depend on $F_{i,y}$.

b = parameter required to estimate VAR($m_{i,y}$). The larger the b, the better a model fits a specific data set.

Because $m_{i,y}$ is the expected number of accidents per mi or per intersection, if Road section i is L_i mi long and has N_i intersections, then the mean and variance for nonintersection and fixed object accidents are $L_iE(m_{i,y})$ and L_i^2VAR($m_{i,y}$); the mean and variance for intersection accidents are $N_iE(m_{i,y})$ and N_i^2VAR($m_{i,y}$).

For each accident type there are 15 parameters to be estimated: β, b, and 13 values of α. These were first estimated using data for all 525 road sections. After some outliers were identified and deleted, the parameters were re-estimated. The likelihood function that these parameters maximize is described in Appendix D of Hauer (*3*). To illustrate, the maximum likelihood parameter estimates for nonintersection accidents are given in Table 1.

It is worth noting that the exponent β of AADT is 0.78 for nonintersection accidents (0.71 for intersection accidents and 0.60 for fixed-object accidents). Thus, the relationship between the expected number of accidents and AADT is in each case distinctly nonlinear.

TABLE 1 Parameter Estimates

α_1	0.002844
α_2	0.002885
α_3	0.002745
α_4	0.002550
α_5	0.002662
α_6	0.002634
α_7	0.002479
α_8	0.002699
α_9	0.002601
α_{10}	0.002709
α_{11}	0.002373
α_{12}	0.002541
α_{13}	0.002414
β	0.77606
b	5.571

HOW EFFECT ON SAFETY WAS ESTIMATED

In this section the method used to estimate the effect resurfacing on safety is described. It is somewhat more complex than the more familiar "before-after-with-comparison-group" method. The aim was to (a) use a long accident history to enhance estimation accuracy, (b) account explicitly for changes in traffic flow and for changes in the uncontrolled factors in the "before" and "after" periods, and (c) eliminate the regression-to-mean threat to the validity of the estimates.

These aims can be attained within the Empirical Bayes approach to estimation. In general, the process can be thought to entail four steps.

Step 1

Estimate for each road section what the expected number of accidents per year was during the before-treatment years. Two clues are used for this purpose: (a) the history of accident counts on the road section and (b) the expected count of accidents for road sections with the same traits (AADT, length, number of intersections) in the reference population. This procedure eliminates bias due to possible regression to the mean. The information needed for (b) is the parameters of the multivariate models that link the number of accidents on road sections of a reference population to their AADT, length, and number of intersections. Here Step 1 was based on the following development explained fully in Hauer (*3*):

Let Road section i have accident counts $x_{i,1}, x_{i,2}, \ldots, x_{i,n}$ in Years $y = 1,2, \ldots, n$. Collectively, these form the Vector $\boldsymbol{x}$. The information contained in $\boldsymbol{x}$ can be combined with the information contained in $E(m_{i,y})$ and VAR($m_{i,y}$) obtained earlier into $E(m_{i,1}|\boldsymbol{x})$ using

$$E(m_{i,1}|\boldsymbol{x}) = \frac{b + \sum_{y=1}^{n} x_{i,y}}{a + L_i \sum_{y=1}^{n} C_{i,y}} \quad (2)$$

In this $E(m_{i,1}|\boldsymbol{x})$ is used to estimate the m of Road section i in Year 1, n is the number of time periods for which accident data are used, and a and $C_{i,y}$ are given by

$$a = b/E(m_{i,1}) \quad (3)$$

and

$$C_{i,y} = (\alpha_y F_{i,y}^{\beta})/(\alpha_1 F_{i,1}^{\beta}) \quad (4)$$

Step 2

Using the results of Step 1, predict what the expected number of accidents on that road section would have been in the period after resurfacing if it had not been resurfaced. In this step, one has to account for changes in traffic from the "before" years to the "after" year, as well as for changes in the various uncontrolled factors.

Here, based on the earlier development, $E(m_{i,y}|\boldsymbol{x}) = C_{i,y}E(m_{i,1}|\boldsymbol{x})$, it follows that for a road section that is L_i mi long,

$$E(L_i m_{i,y}|\boldsymbol{x}) = L_i C_{i,y} E(m_{i,1}|\boldsymbol{x}) \quad (5)$$

Step 3

Estimate for that road section what was the expected number of accidents during the after period with resurfacing in place. Compare this to the result of Step 2. Estimate the safety effect.

Step 4

Repeat Steps 1, 2, and 3 for all treated road sections. Combine the results for individual road sections to obtain estimate of mean effect.

To illustrate, estimate $E(m_{1,1})$ and $VAR(m_{1,1})$ for nonintersections accidents of a site that in 1975 (ie., for Year 1) is estimated to have had AADT = 1199 and is 1.6 mi long. From Table 1 α_1 = 0.002844, β = 0.77607, and b = 5.571. The estimate of $E(m_{1,1})$ is $0.002844 \times 1199^{0.77606} = 0.697$ nonintersection accidents/mi per year; $VAR(m_{1,1}) = 0.697^2/5.571 = 0.0872$, which makes a standard deviation of $\sqrt{0.0872} = 0.295$ accidents/mi per year. The estimate of b is 5.571 as given in Table 1. Therefore the estimate of a is $5.571/0.697 = 7.99$.

For this site there are "before resurfacing" accident counts for 7 full years (1975 to 1981) and 3 months in 1982. Thus, $n = 8$. The accident count vector $\boldsymbol{x}$ is 1, 4, 5, 1, 4, 1, 3, 0. Their sum is 19. Because b was found to be 5.571, the numerator in Equation 2 is $5.571 + 19 = 24.571$. For the denominator of Equation 2 one needs values of $C_{i,y}$. These are calculated by Equation 4 and shown in Table 2.

One now can calculate the denominator of Equation 2. The value of a calculated earlier was 7.993. The length of this road section was said to be 1.6 mi. The sum of the C's for 7 full years is 6.325. Only 3 months of accident data for the 8th year are used. Therefore, $C_{1,8} = 0.848 \times 3/12 = 0.212$. This makes the sum of C's to be $6.325 + 0.212 = 6.537$. Thus, the denominator in Equation 2 is $7.993 + 1.6 \times 6.537 = 18.452$. Doing the calculations of Equation 2, one finds that $E(m_{1,1}|1, 4, 5, 1, 4, 1, 3, 0) = 24.571/18.452 = 1.33$ accidents/mi/year in 1975.

Note that if one were to take the raw accident count for the 7 full years, one would obtain $19/(7 \times 1.6) = 1.70$ accidents/mi per year. This amounts to setting $a = 0$ and $b = 0$ and making all C's 1. Doing so means that one does not recognize the variations in traffic from year to year or the variations that go with the passage of time. (This is why the usual advice is not to extend the "before" period beyond three years. The fear is that, if corrections for changes in traffic and other factors are not applied, accident counts from the distant past are of doubtful use when projections are to be made into the "after" period.) The advantage of accounting for changes in traffic and other factors as is done in Equation 2 is to allow the use of a longer "before" history of accidents counts. This enhances the accuracy of estimation. The incorporation of Parameters b and a in Equation 2 reflects the influence of the reference population and ensures that the result is not biased by regression to the mean.

TABLE 2 Calculation of $C_{1,y}$

y	α	$F_{1,y}$	$E\{m_{1,y}\}$	$C_{1,y}$	Cum $C_{1,y}$
1	0.002844	1199	0.697	1.000	1.000
2	0.002885	1201	0.708	1.016	2.016
3	0.002745	1175	0.662	0.950	2.966
4	0.00255	1163	0.610	0.875	3.841
5	0.002662	1098	0.609	0.874	4.715
6	0.002634	1042	0.579	0.830	5.546
7	0.002479	1038	0.543	0.779	6.325
8	0.002699	1038	0.591	0.848	7.173
9	0.002601	1049	0.575	0.824	7.997
10	0.002709	1083	0.614	0.880	8.878
11	0.002373	1104	0.546	0.782	9.660
12	0.002541	1140	0.599	0.859	10.519
13	0.002414	1228	0.603	0.865	11.384

What the expected number of accidents on this site would have been can now be predicted if no resurfacing had taken place. Using Equation 5 one predicts, for example, for 1983 (y = 9, the first full year after resurfacing), that $E(L_1 m_{1,9}|\boldsymbol{x}) = 1.6 \times 0.824 \times 1.33 = 1.76$ accident per year. This prediction has been built up gradually from several pieces of information.

- Accident history of the site during the "before" period taking into consideration the changing AADT in the "before" years and also accounting for the year-to-year change in various uncontrolled factors.
- Distribution of m's at similar sites, based on the multivariate model.
- AADT in 1983 and the effect of the uncontrolled factors for that year.

The next step is obvious. In 1983, after the site had been resurfaced, it recorded three nonintersection accidents. Without resurfacing 1.76 such accidents would have been expected. Thus, for this site and year, there were 1.24 more nonintersection accidents than expected. Since changes in traffic and other factors were accounted for, the noted difference is attributable to resurfacing.

Of course, one can not form an opinion about the safety effect of resurfacing on the basis of one site and one year. The effect will be added up for all sites and examined for all years. The hope is that, by doing so, sufficiently accurate results can be obtained. This is the subject of Step 4. Since the suspicion is that the effect of resurfacing changes with time, an attempt will be made to examine the effect on a monthly not a yearly basis. Indexes of monthly variation are used for this purpose.

EFFECT OF RESURFACING IN FAST TRACK PROJECTS

There is information about 82 fast-track sites, that is, projects involving primarily resurfacing. The effect of resurfacing on safety for three accident types—nonintersection, intersection, and fixed object—will be estimated.

Effect of Resurfacing on Nonintersection Accidents

By using the method in the previous section, the results in Table 3 were obtained. Thus, during the first month after resurfacing, 20.91 nonintersection accidents would be expected if no resurfacing had taken place, and the average pre-resurfacing pavement conditions were to prevail (Column 2). The accumulative sum of the expected numbers is given in Column 3. Actually, 24 such accidents were recorded in the first month after resurfacing (Column 4). Both numbers, 20.91 and 24, are the sum for the 82 fast-track project sites (Column 5). The difference between 24 and 20.91 is the excess number of accidents, Column 6. The last column lists the cumulative excess. Thus, at the end of the third month after resurfacing, the cumulative excess is estimated to be 10.06 nonintersection accidents. The table is interrupted in several places. Because not all 82 sites were resurfaced at the same time, not all had the same length of "after" history. For example, only 40 sites had a history longer than 74 months after resurfacing.

TABLE 3 Summary Calculations

After Months (1)	Expected Accid. (2)	Cumul. Expected (3)	Accident Count (4)	Number of sites (5)	Excess Accid. (6)	Cumul. Excess (7)
1	20.91	20.91	24	82	3.09	3.09
2	23.35	44.26	27	82	3.65	6.74
3	23.68	67.94	27	82	3.32	10.06
4	20.35	88.29	17	82	-3.35	6.71
...						
30	18.07	605.37	17	82	-1.07	124.63
31	17.91	623.27	12	82	-5.91	118.73
...						
41	18.82	829.78	22	82	3.18	138.22
42	17.02	846.80	18	82	0.98	139.20
...						
60	21.08	1203.41	21	82	-0.08	133.59
61	23.57	1226.98	31	82	7.43	141.02
...						
73	16.24	1438.00	12	50	-4.24	106.00
74	16.97	1454.97	20	40	3.03	109.03

The data of Table 3 are shown in Figure 1. The asterisks in Figure 1 belong to the left scale and show the accumulation of the excess of nonintersection accidents with time after resurfacing (the last column in Table 3). The solid line in Figure 1 belongs to the right scale and shows the accumulation of the number of accidents expected without resurfacing. Note that the two scales differ by a factor of 10. To make interpretation easier, the same ratio of left and right scales also will be used in the subsequent figures.

The orderliness of the results is remarkable. For the first 30 months or so, there is an excess averaging 4.15 nonintersection accidents/month. The standard deviation of this average is 0.93. (So, even from the statistical point of view, the excess must be thought real. A null hypothesis that there was no increase in nonintersection accidents is clearly rejected). Over 30 months the excess accumulates to about 125 nonintersection accidents (with a standard deviation of 28). Thus, it is estimated that, if no resurfacing had taken place and if the pre-resurfacing pavement conditions continued to prevail, 125 fewer nonintersection accidents would have been recorded within 30 months. Over the same period of time about 605 nonintersection accidents would be expected without resurfacing. Thus, the increase is of about 21 percent (124/605 = 0.21). After the first 30 months, there is a 10-month transition during which the monthly accident excess gradually diminishes. The detrimental effect of resurfacing appears to vanish after about 40 months. Over these 40 months, more than 135 nonintersection accidents has accumulated, with a standard deviation of 33. Without resurfacing one would have expected to accumulate by that time 810 nonintersection accidents.

After 40 months there is a plateau that lasts until about 63 months after resurfacing. During this period the average monthly excess is 0.28 nonintersection accident. The standard deviation of this average is 0.79 accident. Thus, it appears that on the plateau the number of accidents is approximately what would have been expected without resurfacing but with pavement conditions that prevailed before resurfacing. Following the plateau, there is a gradual decline. That is, beginning with month 64 after resurfacing, there are fewer nonintersection accidents every month than one should expect if the pre-resurfacing pavement conditions continued to prevail. The number of sites having such a long post-resurfacing history is small. Therefore, one can not say whether the noted decline is real. However, inasmuch as similar declines will later be noted for other accident types, the trends appear to have substance. One may speculate that after more than 5 years of service, the pavement condition is on average worse than what it was in the before-resurfacing period. Just as a new pavement was seen to generate an excess of nonintersection accidents, it should not be surprising that old pavements seem to have the opposite effect.

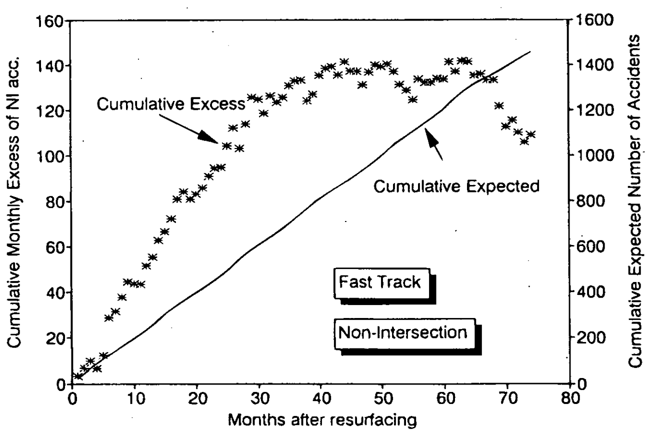


FIGURE 1 Safety effect of resurfacing on nonintersection accidents in 82 fast track projects.

Effect of Resurfacing on Fixed-Object Accidents

In general, the detrimental effect of resurfacing on fixed-object accidents was found to be similar to that observed earlier for nonintersection accidents. For the first 30 months, there is an average excess of 3.31 fixed-object accidents per month. The standard deviation of this average is 0.69 accident. This amounts to an excess of nearly 100 fixed-object accidents in 30 months after resurfacing. The standard deviation here is 21 accidents. Over the same period of time about 290 fixed-object accidents would be expected without resurfacing. Thus, the increase is 34 percent. There is a hint of a plateau at about 40 months. The cumulative excess after 63 months is 130 with a standard deviation of 29 fixed-object accidents. As for nonintersection accidents, there is a clear intimation of a decline after 63 months.

Effect of Resurfacing on Intersection Accidents

Two of the 82 road fast-track road sections have no intersections. Therefore, the results here are based on 80 road sections. The accumulation of excess intersection accidents and the number expected without resurfacing is shown in Figure 2.

There are two main differences between the effect of resurfacing on intersection accidents (Figure 2) and its effect on nonintersection accidents (Figure 1). First, the hump that separates the period when more than the expected number of accident occurs from the period when fewer then expected accidents materialize occurs much earlier. Second, the absolute excess is smaller, and therefore the results are not as reliable.

For the first year after resurfacing, there are more intersection accidents than would be expected if the road sections had not been resurfaced. The excess is 2.92 intersection accidents/month with a standard deviation of 1.21 for a total of 35 intersection accidents with a standard deviation of 15. By the end of the first year, 101 intersection accidents would be expected without resurfacing. Thus, there was an increase of 35 percent.

Disregarding the undulations, from the end of the first year until Month 32 there is a plateau. The excess here is 0.01 with a standard deviation of 0.50 intersection accidents per month. From Month 33 to Month 63 (where there are still data for all 80 road sections), each month there are fewer accidents than would be expected if there had been no resurfacing and the pre-resurfacing pavement conditions continued to prevail. The average monthly excess is −2.04 intersection accidents per month with a standard deviation of 0.50.

In Figure 1, a decline was already noted, that is, a negative excess. Here for the first time its magnitude can be estimated. The hypothesis has been advanced earlier that, just as new and smooth pavement is associated with a positive excess, as the pavement ages after a turning point the excess becomes negative. Although the hypothesis is plausible, it is merely a hypothesis. Here one must ask why the turning point for intersection accidents occurs earlier than for nonintersection accidents and by what mechanism can resurfacing and pavement aging affect the frequency of intersection accidents?

One also needs to ask whether there is some factor neglected in the analysis that could have brought about these results. Could perhaps the change in the definition of intersection accidents be responsible for the decline? We think not. Firstly, the decline starts about 32 months after resurfacing while the change in definition occurred (May 1983) perhaps 10 to 20 months after the projects were finished (either in fall 1981 or 1982). Second, no change was detected in the count of intersections accidents coinciding with the change in definition. Third, whatever the effect of the change in definition, it is reflected in the corresponding αs.

EFFECT OF RESURFACING AND OTHER MODIFICATIONS IN R&P PROJECTS

The R&P projects include various additional improvements with resurfacing. Thus, the effect to be estimated is not only of resurfacing but the joint effect of all modifications implemented. There is information about 55 R&P sites. As in the previous section, the joint effect of resurfacing and other improvements on safety will be estimated for three accident types—nonintersection, intersection, and fixed-object.

Effect on Nonintersection accidents

The accumulation of excess nonintersection accidents and of the number of nonintersection accidents expected without resurfacing versus months after resurfacing is shown in Figure 3.

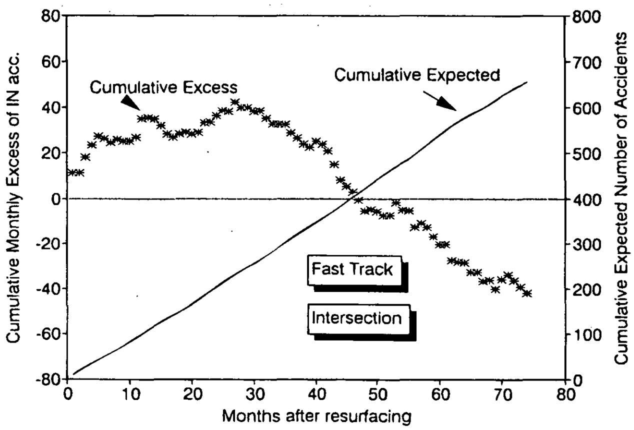


FIGURE 2 Safety effect of resurfacing on intersection accidents in 82 fast track projects.

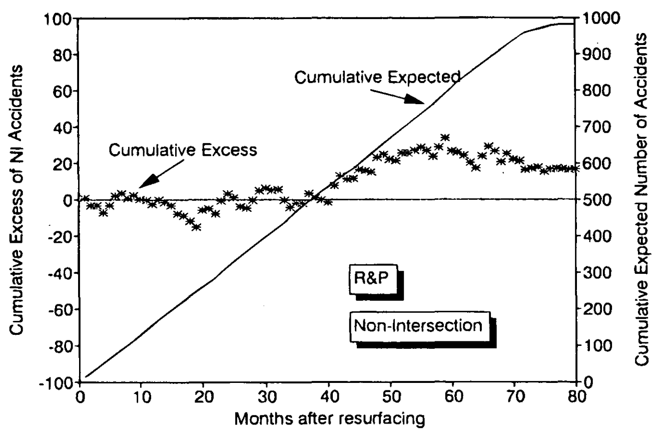


FIGURE 3 Safety effect of improvements on nonintersection accidents in 55 R&P projects.

The comparison is with Figure 1, and the contrast is stark. Without much analysis, one can state that in Figure 3 there is no discernible change in the number of nonintersection accidents from what would be expected if these projects had not been implemented and the pavement condition remained as in the pre-improvement period.

Effect on Fixed-Object Accidents

Although the corresponding figure is not shown here, its examination would show that the joint effect of the resurfacing and other improvements was to make the number of fixed-object accidents just what it would have been if no improvements had been undertaken and the pavement conditions from before the improvement continued to prevail, just as is true for the nonintersection accidents in Figure 3.

Effect on Intersection Accidents

Four of the 55 R&P road sections have no intersections. Therefore, the results here are based on 51 road sections. The accumulation of excess intersection accidents and the number expected without resurfacing are shown in Figure 4.

As is clear, the R&P project improvements are associated with a reduction in the number of intersection accidents for a long period of time. The horizontal tail after Month 70 merely signifies that very few sections have such a long "after" history. The full 51 road sections can be followed only for 60 months.

During that 60-month period, the average number of intersection accidents was reduced by 2.78 per months, with a standard deviation of 0.31 accidents. This accumulates over the 60-month period to a reduction of 167 intersection accidents, with a standard deviation of 19. By that time, 572 intersection accidents would have been expected if the R&P project had not been undertaken and pavement conditions remained constant. This amounts to a 29 percent reduction in intersection accidents.

Again there is the question Is this real? What aspect of the R&P projects can be thought to act on intersection accidents? On one hand, the results here are internally consistent with what has been found for fast-track projects. In fast-track projects, resurfacing was associated with an increase in nonintersection accidents and a slight increase followed by decline of intersection accidents. In R&P projects nonintersection accidents were found to remain stable while intersection accidents declined steadily from the end of resurfacing. Thus, there is a link between how nonintersection and intersection accidents appear to respond. On the other hand, a mechanism by which the kinds of actions undertaken in R&P projects can diminish the number of intersection accidents is unknown.

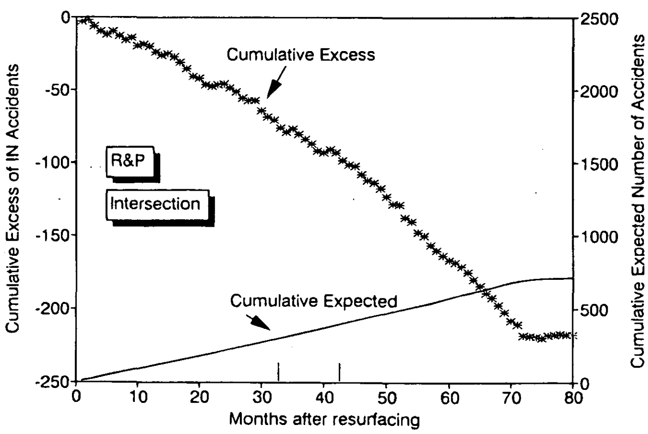


FIGURE 4 Safety effect of improvements on intersection accidents in 51 R&P projects.

SUMMARY AND CONCLUSIONS

In the early 1980s, two kinds of resurfacing projects were undertaken. The task was to ascertain whether projects involving only simple resurfacing (82 fast-track projects, 226.7 mi) perform less well from a safety viewpoint than similar resurfacing projects where roadside and roadway safety improvements have been incorporated (55 R&P projects, 137.2 mi).

The overall answer is yes. In fast-track projects, safety initially declined; in R&P projects safety improved. Since the safety effect of resurfacing changes with the passage of time and differs from one type of accident to another, one can not describe the difference between fast-track and R&P projects by a single number. Table 4 is an attempt at a succinct summary. Another conclusion that emerges from this work is that with the first 6 to 7 years of pavement life, safety improves as the pavement ages.

Retrospective studies of this kind can provide estimates of what the effect on safety of some intervention was. However, to say in such a study how the estimated effect came to be is difficult. There are no data on how speed has changed, how pavement friction was affected, what were the changes in traffic volumes on the crossing legs of intersections, no knowledge of when shoulders were backed up in fast-track projects, or details about what specific improvements were made in which R&P projects. Because the effect of these interventions on safety appears to be large, this lack of explanation is troubling. One would have more confidence in the results if these could be attributed to causes.

Still, the results display a pleasing internal consistency. One can not imagine any element of method or of data analysis that could be incorrect and still leave this internal harmony intact. It is also reassuring that the results presented here are in many ways similar to the results obtained earlier by the NYDOT staff when using part of the data in simple before-and-after comparisons and in comparisons involving the use of control sections.

It appears clear that the kind of resurfacing that went with fast-track projects affected safety differently from the kind of resurfacing associated with R&P projects. This leads to the conclusion that resurfacing as referred to in the professional literature may cover a heterogeneous set of activities. When discussing the effect of resurfacing on safety, one should be specific about the kind of activities performed. Lumping together the safety effect of diverse kinds of resurfacing may give a fuzzy picture.

There are two main novel aspects to the method used in this study. First, it is sensible to expect the effect of resurfacing on safety to change with time as the pavement ages. The method used facilitates the examination of this aspect. Second, the accuracy of studies of this kind is often limited by the sparsity of accident data. The method used here allows the use of a long "before" accident history and enhances accuracy by using information from the reference population.

TABLE 4 Summary of Results

FAST TRACK				R&P		
Non- Intersection		Intersection		Non-Intersection	Intersection	
Mo. 0-30	+21%	Mo. 0-12	+35%	No Change	Mo. 0-70	-29%
Mo. 40-63	0%	Mo. 13-32	0%			
Later	decline(?)	Later	-23%			

The central theme of this study was to assess the effect of the two kinds of treatment on safety. En route, two by-products have been generated. One is a procedure for estimating AADTs for every road section and every year on the basis of traffic counts conducted once in 3 to 4 years. The other by-product consists of the multivariate models, which, with a slight extension, can be the basis of a rational procedure for the identification of hazardous locations.

The work reported here had to be done within time and on budget and is not as complete as it could have been. The procedure for estimating AADT is somewhat ad hoc; the investigation of the correspondence between the reference and comparison groups is limited; with added effort it would have been possible to investigate separately the effect of resurfacing on property damage and on injury accidents; it would also have been possible to examine the effect of the construction period itself.

There is one deficiency that became apparent only after the analysis was completed. The results indicate that as the pavement ages accidents diminish. Because all treated road sections were resurfaced within 1 year of each other, their pavements must have been deteriorating approximately in tandem; they were all in need of repair just before resurfacing and in good shape 5 to 7 years earlier. If so, there is a systematic factor that the analysis in Step 1 neglected. The net effect of this deficiency is that prediction of what would be expected without resurfacing has been produced as if a constant pavement condition prevailed during the entire before-resurfacing period. This logical deficiency applies equally to the fast-track and the R&P projects and is unlikely to affect any of the conclusions materially.

REFERENCES

1. *Special Report 214: Designing Safer Roads.* Transportation Research Board, National Research Council, Washington, D.C., 1987.
2. Cleveland, D. Effect of Resurfacing on Highway Safety: a Synthesis of Prior Research. In *State-of-the-Art Report 6*, TRB, National Research Council, Washington, D.C., 1987.
3. Hauer, E. *The Effect of Resurfacing on the Safety of Two-Lane Rural Roads in New York State.* Final Report, July 1993.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Use of Weigh-in-Motion Scale Data for Safety-Related Traffic Analysis

JERRY J. HAJEK, JOHN BILLING, PHI HOANG, AND ALEXANDER J. UGGE

So far, more than 2000 traffic lanes on North American highways have been equipped with weigh-in-motion (WIM) scales. These WIM scales provide and will continue to provide a large amount of data for individual highway vehicles, such as axle spacing and weights, vehicle length, speed, and headway. Because of their unobtrusiveness and continuous operation, WIM scales provide unbiased, statistically reliable data. The use of WIM data for investigating safety-related highway traffic flow characteristics is examined. WIM technology and its capabilities to generate traffic-monitoring data useful for transportation planning and decision making are described. Examples of data analysis that demonstrate the usefulness of WIM data for investigating safety-related traffic characteristics are provided. They include determination of truck exposure rates and evaluation of vehicle speed and headway distributions as a function of highway facility, vehicle type, daytime and nighttime conditions, and truck load. WIM data are useful in many areas of transportation planning, including safety-related traffic analysis, and should be considered corporate data and managed accordingly.

With the advent of the Strategic Highway Research Program (SHRP) and its national satellite programs, such as Canadian-SHRP, weigh-in-motion (WIM) scales have become commonplace. It is estimated that there are now more than 2000 traffic lanes equipped with WIM scales in North America, with installations in virtually all states and provinces. The Ontario Ministry of Transportation is operating nine in-highway WIM scales.

The WIM scales provide and will continue to provide a large amount of detailed data for individual highway vehicles, such as axle spacing and weights, speed, and headway. This is in addition to the traditional aggregated traffic characteristics such as daily and annual vehicle volumes and equivalent single-axle loads. Considering the effort associated with the installation and operation of WIM scales and with subsequent data retrieval and processing, the wealth of traffic-monitoring data generated by WIM scales should be properly used for as many purposes as possible.

Because of the original association with the SHRP-related pavement research effort, it is often assumed that WIM data are only applicable to pavement performance research. Many potential users of WIM-type data do not know the following:

- Traffic-monitoring capabilities of WIM technology,
- Type of data available, and
- How the data can be used within their area of interest.

Hajek et al. (*1*) demonstrated that WIM data are useful for a wide range of transportation planning and decision-making purposes, including

- Planning of transportation facilities,
- Pavement design and rehabilitation,
- Apportionment of pavement damage,
- Compliance with vehicle weight regulations,
- Development of geometric design standards,
- Compliance and regulatory policy development of truck dimensions,
- Traffic safety analysis,
- Traffic operation and control, and
- Analysis related to highway bridges.

The objective of this work is to make traffic safety researchers and administrators aware of the potential of WIM data in the traffic safety analysis area. Specifically, the objective is to show, by practical examples, that WIM data are also useful and indeed indispensable for fundamental safety-related traffic analysis.

The data used in this study were obtained by three Ontario WIM scales: two scales were located on a freeway (Hwy 402) and one scale on a two-lane highway (Hwy 31). All scales use piezoelectric cable technology (*2*). Compared with static conditions, the scales provide dimensions accurate within 2 to 3 percent, gross vehicle weights within about 5 percent, and axle loads within about 5 to 12 percent. Accuracy depends on vehicle dynamics (e.g., vehicle configuration and speed) and on pavement roughness in the vicinity of the scale.

Highway 402 is a four-lane rural freeway with a speed limit of 100 km/hr. One WIM scale is located in an eastbound right (truck) lane near Sarnia and is referred to as Location 1; the second scale spans both westbound lanes near London and is referred to as Location 2. Low traffic volumes at the two locations on Hwy 402 (about 400 vehicles during a peak daytime hour and about 75 vehicles/hr at night in the right lane) enable a large degree of traffic operational freedom. However, Location 2 is about 2 km downstream from a freeway entrance ramp. Because of highway alignment constraints, traffic may not have always reached its free-flow equilibrium at this location.

Highway 31 is a two-lane rural highway with a speed limit of 80 km/hr. The WIM scale is in both lanes; the WIM in the northbound lane is referred to as Location 3. High traffic volumes on Hwy 31 (about 3500 vehicles per day in each direction) greatly restrict traffic operational freedom, particularly passing opportunities. The highway grade is at a level for 2 or more km before all WIM scale locations.

DESCRIPTION OF TRAFFIC-MONITORING DATA PROVIDED BY WIM SCALES

A typical WIM scale consists of magnetic loops and axle sensors embedded in the pavement and a microcomputer housed in a road-

J. J. Hajek and A. J. Ugge, Research and Development Branch, Ontario Ministry of Transportation, Downsview, Ont., Canada, M3M 1J8. J. Billing, Transportation Technology and Energy Branch, Ontario Ministry of Transportation, Downsview, Ont., Canada, M3M 1J8. P. Hoang, Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1.

side cabinet. Magnetic loops and axle sensors respond to axles passing over the pavement by generating electric signals. The signals are processed by the computer and transformed into engineering parameters for each vehicle. These parameters include a time stamp, instantaneous vehicle speed, vehicle length, distances between consecutive axles, and axle weights.

The knowledge of engineering parameters for individual vehicles, and a knowledge of the impact of vehicle weight and dimension regulations on truck design, can be used to determine vehicle types for classification purposes. A judicious classification scheme using expert system techniques can pinpoint and select specific vehicle types of interest. For example, Figure 1 shows one such scheme designed for identification of fully loaded six-axle tractor-semitrailers with a liftable ("belly") axle, the most common heavy-haul truck configuration in Ontario.

Figure 1 illustrates the following criteria for truck definition:

1. Six-axle trucks;
2. Single (steering), dual (tractor), single (liftable), dual (trailer) axle arrangement;
3. Axles 2 and 3 spacing from 1.07 to 1.83 m;
4. Axles 3 and 4 spacing greater than 4.00 m;
5. Axles 4 and 5 spacing greater than 2.40 m;
6. Axles 4 and 5 spacing greater than Axle 5 and 6 spacing;
7. Axles 5 and 6 spacing from 1.07 to 3.05 m; and
8. Gross weight within 1000 kg of allowable load.

The first item simply ensures the proper number of axles. The second is descriptive and may be redundant. Item 3 covers the known range of drive axle spreads. Item 4 ensures that the filter captures semitrailers at least 10 m (32 ft) long, so it will exclude tractors pulling 7 m (23 ft) tridem container chassis. Items 5 and 6 ensure that the liftable axle is properly separated from the trailer tandem axle, according to Quebec, Ontario, or Michigan regulations. Item 7 covers the known range of trailer axle spreads. The final item ensures that the gross weight is close to the allowable limit.

Considering the variety of vehicles on Ontario's highways and the large samples of vehicles analyzed, there must be a certain level of speculation associated with any vehicle classification scheme based on WIM data. There is no reason to believe that these uncertainties have a significant effect on the observations presented.

It should be stressed that, although WIM-type data can be used to enumerate the population of trucks of any particular basic type (e.g., fully loaded six-axle trucks with one liftable axle, as outlined previously), WIM scales cannot at this time discern certain types of multi-unit trucks, the body style, commodity/load, owner, or other

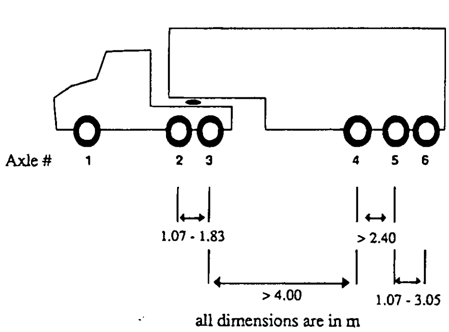


FIGURE 1 Criteria identifying six-axle tractor semitrailers with liftable axles.

information that might identify the specific nature of the truck and its owner. However, the identification of numbers of specific truck configurations is still very important because it allows infrequent or low-probability events to be examined, which is not practical with the small samples that can be garnered by manual survey methods (*3*).

For a more comprehensive assessment of traffic flows, WIM data should be supplemented by data from other sources. Some of these sources and sensor technologies are

- Video imaging for more specific visual vehicle classification (*4*),
- Radar and microwave image processing to identify vehicle shape or bulk,
- Automatic vehicle identification systems that can provide detailed vehicle descriptive data and ownership information (*5*),
- Vehicle height detectors (existing ones usually use a laser beam) to determine vehicle longitudinal profile (and to facilitate detailed vehicle classification),
- Automatic vehicle location systems that can provide information regarding distances traveled by individual vehicles (which can be related, for example, to travel speeds),
- Weather stations to assess weather conditions such as the rate and type of precipitation and visibility, and
- Pavement condition sensors to determine presence of snow or ice.

The use of many of these supplemental sensor technologies require appropriate communication and data integration systems. Spurred by Intelligent Highway Vehicle System needs, work is under way in many of the technology areas to obtain a more comprehensive knowledge of traffic flows.

Finally, data obtained from the individual WIM scales represent only the traffic mix at a specific highway site during a particular monitoring period and not a global picture of traffic flow.

SAFETY-RELATED TRAFFIC ANALYSIS

Safety is a major issue in all debates about changes in highway traffic regulations and in vehicle weights and dimensions. Invariably, it is concluded that adequate information about safety implications of the proposed changes is lacking (*6*). WIM data can contribute to analysis of these issues by providing unique and detailed information on the following:

- Frequency of different vehicle types using highway facilities,
- Driver and vehicle behavior on highway facilities, and
- Truck payload and its distribution.

Frequency of Vehicle Types Using Highway Facilities

The knowledge of accident rates for different truck types is instrumental in identifying the influence of vehicle design parameters on highway safety. The accident rate is defined as the number of accidents divided by the number of kilometers traveled (exposure rate to risk of accident). Because trucks are registered once but may travel in several jurisdictions, vehicle registration systems usually do not provide adequate information for estimating vehicle kilometers traveled by different truck classes, which would help to obtain accident rates for different truck types. Moreover, such estimates are only general and not for specific highway types or locations. WIM data can help establish truck exposure measures, particularly for facilities where WIM scales have been installed. For example, five-axle trucks, consisting of a three-axle tractor with one dual axle semitrailer or 3S2, made up about 60 percent of the total truck volume on Hwy 402 (Location 1) (Figure 2). It would be possible to relate the volume percentage of 3S2s to the percentage of accidents involving the 3S2s on this facility.

Driver and Vehicle Behavior on Highway Facilities

Although highways are designed to serve a mix of vehicle types, the effects of various types of vehicles on traffic operations and safety are not uniform. In this context, factors such as acceleration, speed, headway, passing, merging and other lane changing maneuvers, splash and spray, aerodynamic buffeting, blockage of view, and lateral placement are clearly different among vehicle classes and influence the interaction of the vehicles. Inherent vehicle performance factors related to highway safety are vehicle handling, stability and braking capabilities, and load and load distribution (*7*).

WIM data alone can be used to evaluate driver and vehicle behavior in terms of speed and headway distributions as a function of axle (vehicle) weight and time of day (daytime versus night time). For a more comprehensive assessment, WIM data should be supplemented by data from other sources described previously. In this paper data from a nearby weather station were used to select WIM data for periods when the pavement was likely to be dry.

Vehicle Speed Distribution

Excessive vehicle speed, and particularly speed differentials between different vehicles, is considered a main cause of accidents (*8*). When different vehicle types exhibit different speeds (loaded trucks may travel more slowly than the prevailing traffic, particularly on upgrades), the speed variance of the traffic flow increases. The difference in speed variance has been linked to the increase in overall accident rates (*9*). The primary vehicle characteristic affecting acceleration and speed performance of trucks is the weight/power ratio. Overloaded and speeding trucks may constitute an additional safety hazard.

Examples of vehicle speed distribution for cars and trucks, obtained by the WIM scales during daylight hours and at night, are shown in Figures 3 and 4. Figure 3 shows data for Location 1 (right lane on a four-lane freeway); Figure 4 shows data for Location 3 (northbound lane of a two-lane highway).

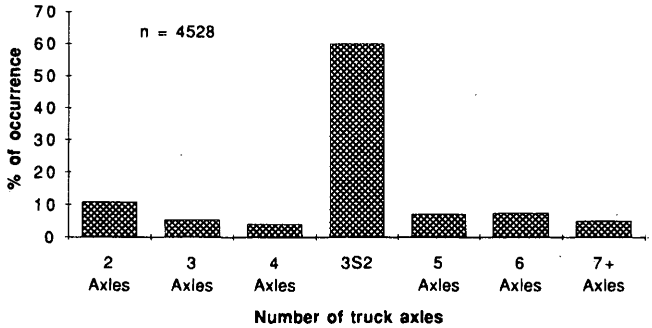


FIGURE 2 Occurrence of different truck types, Hwy. 402, March 19 to 22, 1991.

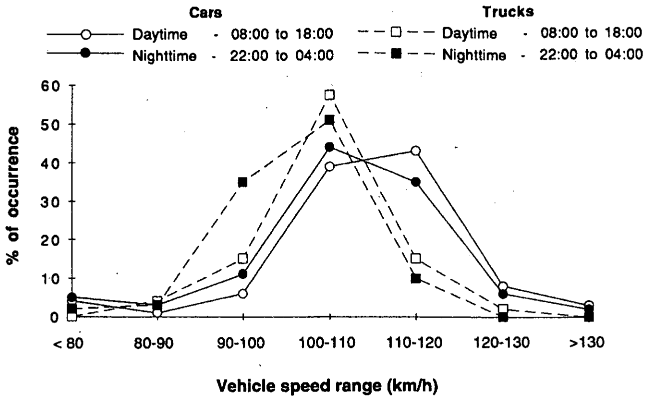


FIGURE 3 Vehicle speed distribution, Hwy. 402, truck lane, March 19 to 22, 1991.

Data for Location 1 were obtained on four consecutive weekdays without any precipitation in March 1991. Overall, data in Figure 3 indicate that truck drivers are more disciplined than car drivers. Some specific observations are

- Highway speed limit is 100 km/hr, and most cars were speeding. During daytime, about 53 percent of all car drivers exceeded 110 km/hr; at night 42 percent of all car drivers exceeded this speed. The corresponding numbers for truck drivers were 16 and 10 percent, respectively.
- Compared with cars, the truck speed distribution is more uniform. Looking at the extremes, during daytime, 1.3 percent of cars had speeds lower than 80 km/hr compared with only 0.3 percent of trucks. At the high end, 1.3 percent of cars (in the right lane) exceeded 130 km/hr compared with 0.1 percent of trucks.
- The more uniform speed distribution observed for trucks is reflected in their lower speed variance. A reduction in speed variance has been linked empirically to a reduction in accident rates (*7*).

Data for Location 3 (Figure 4) were obtained on seven consecutive weekdays in October 1992. It should be noted that the data are only for the northbound lane. Any passing northbound vehicles use

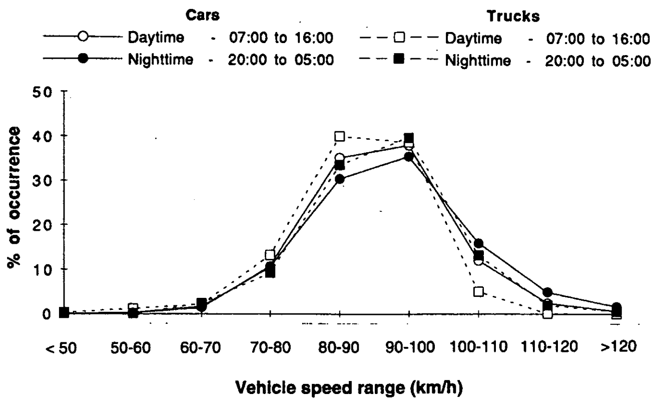


FIGURE 4 Vehicle speed distribution, Hwy 31, northbound lane, October 1 to 5, 1992.

the southbound lane, and their passage is treated as an error by the WIM scale in the southbound lane. There are some similarities between the data in Figures 3 and 4. Most cars are again speeding (about 87 percent of cars exceed the 80 km/hr speed limit during daytime), and truck drivers are again more disciplined than car drivers. However, the differences in the speed distribution attributable to the vehicle type and the time of day are considerably attenuated by the lack of operational freedom caused by the large traffic volumes. In fact, because of the large daytime traffic volumes and contrary to the results given in Figure 3 for Location 1, the daytime (7:00 to 16:00) car and truck speeds on Location 3 are smaller than the night-time (20:00 to 05:00) speeds. For example, the mean night-time truck speed was 91.0 km/hr, and the corresponding daytime speed was only 87.7 km/hr. The higher night-time speeds may be one of the contributing factors to the often-encountered higher night-time accident rates.

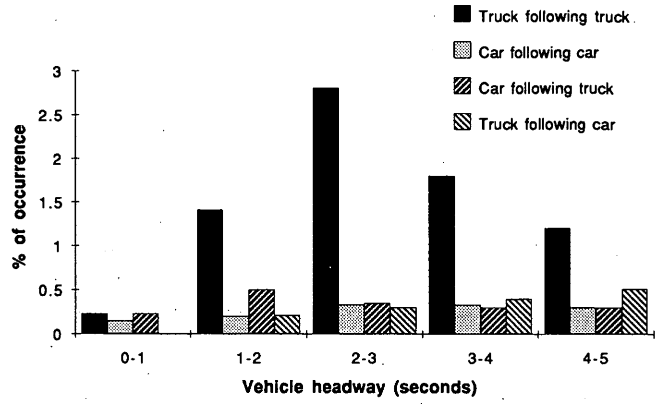


FIGURE 6 Nighttime headway distribution, Hwy. 402, truck lane, March 19 to 22, 1991.

Headway Distribution

According to Ontario's Highway Traffic Act (*10*), maintaining "reasonable and prudent" headway (the time or distance between successive vehicles) is mandatory for all drivers. There is an extra stipulation for drivers of commercial vehicles (trucks) who, while driving at speeds exceeding 60 km/hr, "shall not follow within 60 m of another motor vehicle."

Vehicles traveling at 100 km/hr (27.8 m/sec) would have a front-bumper-to-front-bumper spacing of only 28 m with 1-sec headway. This results in the actual space between vehicles of about 24 m for an average car and, of course, considerably less space for even small trucks. Clearly, headways less than 2 sec do not meet the "reasonable and prudent" stipulation at highway speeds.

Figures 5, 6, and 7 compare the difference in headway distributions of cars and trucks. Figures 5 and 6 are for Location 1 and use the same data set as that used for Figure 3; Figure 7 is for Location 3 and uses the same data set as that used for Figure 4.

Examining first the results for the freeway location (Figures 5 and 6), the greater discipline of truck drivers, as shown by the speed distribution, is also indicated by the headway distribution. Some additional observations are

- During daytime, 7 percent of all cars followed other cars with a headway of 1 sec or less; only 2.5 percent of trucks did so. Nevertheless, considering an average truck speed of 100 km/hr, more than 2.5 percent of all trucks appear to be in violation of the Ontario Highway Traffic Act headway requirement.
- Also during daytime, 3.5 percent of all trucks were following other trucks with a 1-sec headway or less and only 2.5 percent of trucks were following cars with this headway. The difference in the headway distribution for these two cases was found to be statistically significant.
- Trucks tend to travel in convoys. This is particularly evident at night when 7.5 percent of all trucks had headway of less than 6 sec compared to only 1.1 percent for cars.

The headway distribution on the two-lane highway during daytime (Figure 7) shows even more pronounced tendency of cars to follow other vehicles with short, unsafe headways.

- About 36 percent of cars followed other vehicles with the headway of 1.5 sec or less; only 8 percent of trucks did so.
- Compared with the headway distribution obtained for the freeway location, there is a pronounced tendency of trucks to follow other trucks with headway in the range of 1.5 to 5 sec. This can be attributed to the inability of trucks to pass other trucks, thus form-

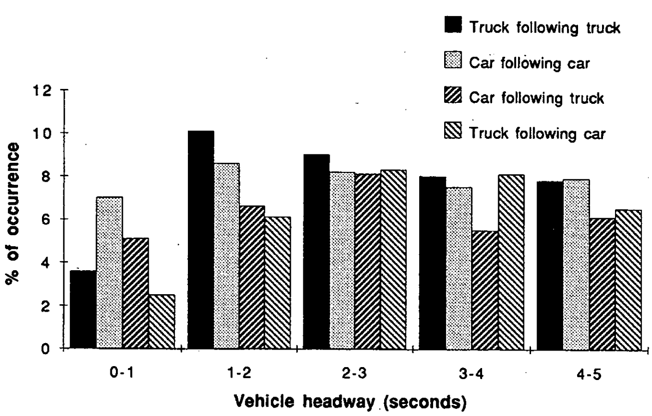


FIGURE 5 Daytime vehicle headway distribution, Hwy. 402, truck lane, March 19 to 22, 1991.

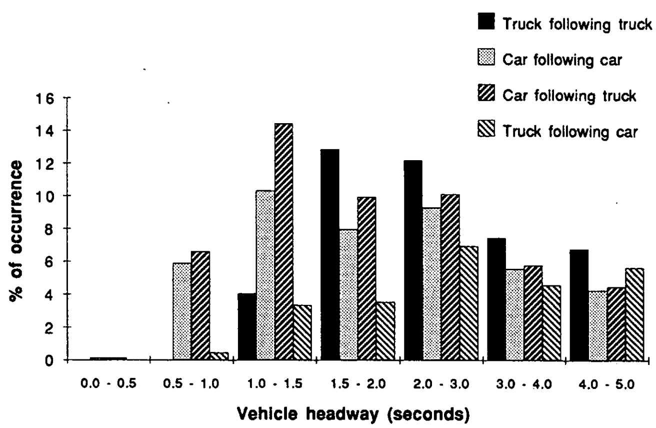


FIGURE 7 Daytime vehicle headway distribution, Hwy. 31, northbound lane, October 1 to 15, 1992.

ing truck convoys. Passing opportunities for cars that follow such convoys may be limited.

Truck Payload and Distribution

Another descriptive parameter useful in accident studies (and provided by WIM data) is the gross vehicle weight and its distribution among axles and, in the case of multiple truck units, between truck units. Braking performance of trucks is not as good as that for cars—braking distances of trucks are significantly longer. A recent investigation (*11*) analyzed braking capabilities of different truck types for unloaded, partially loaded, and fully loaded vehicles. It appears that braking distances can increase substantially when trucks are partially loaded or unloaded and that the magnitude of this degradation in braking capabilities depends on the truck type. A comprehensive analysis of truck accidents in Ontario (*12*) indicated that, except for twin trailers, the highest accident rates were for unloaded straight trucks, semi-trailers, and bobtail tractors. This suggests a link between braking capabilities of unloaded, partially loaded, and fully loaded trucks and accident frequencies, which should be addressed in part by the introduction of antilock brake systems.

Payload distribution can also affect truck operating characteristics. Billing and Hajek (*3*) show how WIM data can be used to evaluate payload distribution on six-axle trucks with one liftable axle. Systematic evaluation of WIM data for loaded and unloaded trucks can provide valuable insights into their operating characteristics. In this work, the loaded trucks were defined as trucks with a payload estimated to be at least one-half of the total allowable load.

Figure 8 uses data obtained on a passing lane at Location 2 (freeway location) to compare the speed distributions for unloaded and loaded six-or-more-axle trucks during daytime. Although the difference in mean speed between unloaded and loaded trucks is only 1.5 km/hr, the corresponding speed variance, with its safety implications, differs by 10 $(km/hr)^2$.

The differences in speed distributions for unloaded and loaded three-or-more-axle trucks during daytime obtained for Location 3 are illustrated in Figure 9. The differences may be attenuated by the relatively high traffic volumes on this highway. Nevertheless, the difference in the mean speed of the unloaded and loaded trucks is 1.4 km/hr, while both loading conditions have a similar speed variance.

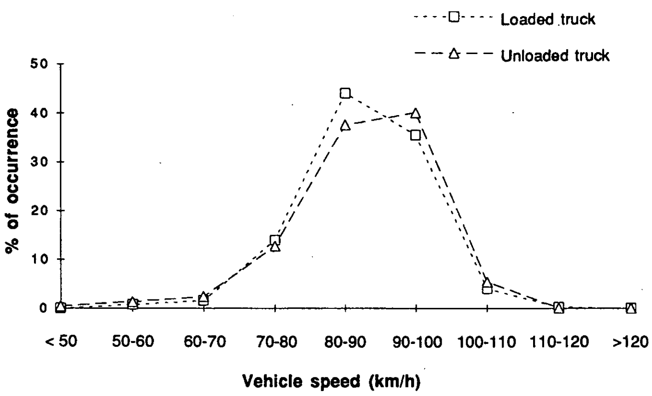


FIGURE 9 Daytime vehicle speed distribution for three-or-more-axle trucks, Hwy. 31, northbound lane, October 1 to 15, 1992.

Examples of headway distribution for loaded and unloaded trucks with six-or-more axles, obtained for the right lane at Location 2 and for the northbound lane at Location 3, are given in Figures 10 and 11, respectively. The headway distribution is different on the two facilities.

On the freeway, the loaded trucks are more likely to follow cars with short headways than are the unloaded trucks. The difference in the headway distribution was found to be statistically significant. The higher occurrence of shorter headways observed for the loaded trucks may be attributed to the lack of engine power available to pass slower moving cars. The unloaded trucks are able to pass these cars using the left (passing) lane.

In contrast, on the two-lane highway, the unloaded trucks are more likely to follow cars with short headways than the loaded trucks. For example, about 8 percent of unloaded six-or-more-axle trucks followed cars with the headway in the 1.5 to 2.0 sec range, and only about 1.5 percent of the loaded trucks did so. It appears that the drivers of the unloaded trucks, because of the availability of spare power, are positioning their trucks to pass slower moving cars. Alternatively, they may simply be better able to maintain the speed with the more nimble cars.

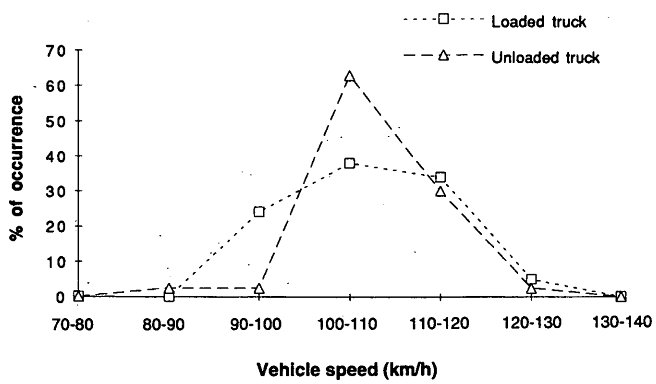


FIGURE 8 Daytime vehicle speed distribution for six-axle trucks, Hwy. 402, passing lane, March 1992.

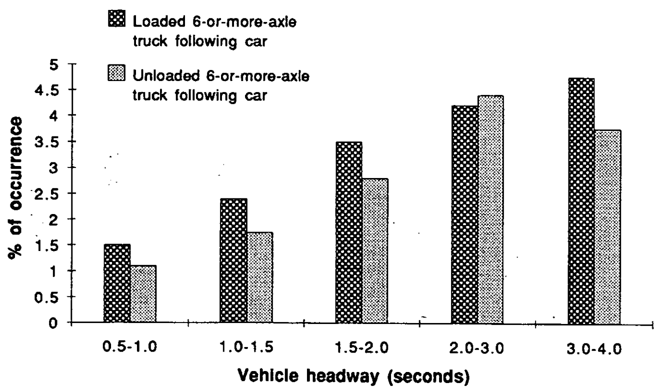


FIGURE 10 Daytime truck headway distribution for six-or-more-axle trucks, Hwy. 402, Lane 1.

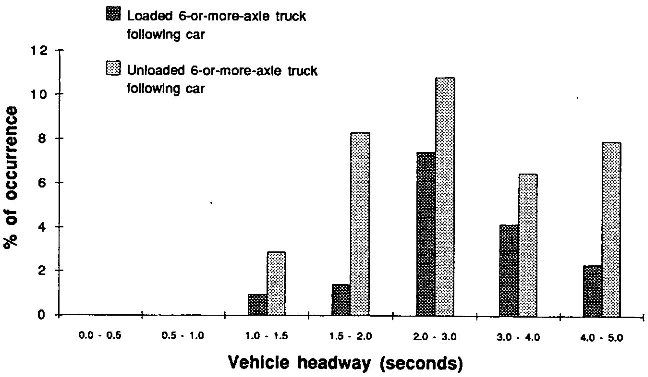


FIGURE 11 Daytime vehicle headway distribution for six-or-more-axle trucks, Hwy. 31, northbound lane, October 1 to 15, 1992.

DISCUSSION OF RESULTS

The discussion on extracting specific information from WIM data to provide insights into safety-related highway traffic characteristics is not exhaustive. It simply illustrates possible use of WIM data in traditional application areas. It is also possible, for example, to study more complex functions of traffic flow such as the relationship among vehicle speed, headway, payload, and weather conditions for different vehicle categories and to provide data to develop and manage police strategies for enforcing traffic regulations.

WIM scales have been also used as a main component of speed advisory systems for truck drivers approaching potentially hazardous conditions, such as

- Long steep downgrades, or steep downgrades with traffic signals or stop signs—The system advises truck drivers, particularly drivers of loaded trucks, of a recommended speed to negotiate the grade safely.
- Freeway ramps with small turning radii—The system advises truck drivers, such as drivers of loaded tanker trucks, of a recommended speed to prevent a rollover (*13*).

Although this work addresses application of WIM data and WIM scales in the traffic-safety area, their use for transportation planning and decision making is much larger and cuts across the organizational structure of any highway agency (*1*).

CONCLUSIONS

1. WIM data are useful for a wide range of transportation planning and decision-making purposes including traffic safety-related applications.
2. Using examples, WIM data have been demonstrated to provide previously unavailable insights about a range of safety-related highway traffic flow characteristics.
3. WIM scales, because of their unobtrusiveness and continuous operation, can provide truly unbiased statistically reliable data, yielding a realistic long-term picture of exposure rates for specific vehicle types and other safety-related traffic flow characteristics. Both concerns are important for identification of relative influence of vehicle design parameters and driver behavior on highway safety.
4. It is imperative that those working in the traffic safety area are made aware of the potential of WIM data for traffic-related safety analysis, and that this potential is further pursued.

REFERENCES

1. Hajek, J. J., J. R. Billing, and G. Kennepohl. Applications of Weigh-In-Motion Data in Transportation Planning. In *Transportation Research Record 1364*, TRB, National Research Council, Washington, D.C., 1992.
2. *IRD Weigh-in-Motion Piezoelectric 1060P User's Manual.* International Road Dynamics, Inc., Saskatoon, Saskatchewan, Canada, 1990.
3. Billing, J. R., and J. J. Hajek. *Weigh-in-Motion Scale Data as a Basis for Geometric Design Standards Development, Weight Compliance Assessment, and Regulatory Policy Development.* Ontario Ministry of Transportation, Downsview, Canada, Sept. 1990.
4. Ronning, R. Video Imaging for Vehicle Classification. Report FHWA/CA/TO/TDS-92-1 *Proc., National Traffic Data Acquisition Conference*, FHWA, U.S. Department of Transportation, Oct. 1992, pp. 185–189.
5. Walton, M. C. A Concept of IVHS in Commercial Vehicle Operation: The HELP/Crescent Program. Report FHWA/CA/TO/TDS-92-1 *Proc., National Traffic Data Acquisition Conference.* FHWA, U.S. Department of Transportation, Oct. 1992, pp. 625–629.
6. McGee, H. W. Accident Data Needs for Truck Safety Issues. In *Transportation Research Record 1052*, TRB, National Research Council, Washington, D.C., 1986, pp. 146–150.
7. *TRB Special Report 211: Twin Trailer Trucks.* TRB, National Research Council, Washington, D.C., 1986.
8. Garber, N. J., and R. Gadiraju. Impact of Differential Speed Limits on the Speed of Traffic and the Rate of Accidents. In *Transportation Research Record 1375*, TRB, National Research Council, Washington, D.C., 1992.
9. Garber, N. J., and R. Gadiraju. *Speed Variance and its Influence on Accidents.* AAA Foundation for Highway Safety, Washington, D.C., July 1988.
10. The Highway Traffic Act (as amended to May 1989), Government of Ontario, Toronto, Canada, May 1989.
11. Parker, D. J., and B. G. Hutchinson. *Large Truck Braking at Signalized Intersections.* Report TDS-88-01. Ontario Ministry of Transportation, Downsview, Canada, Dec. 1988.
12. Buyco, C., F. F. Saccomanno, and A. Stewart. *Factors Affecting Truck Accident Rates in Ontario.* Department of Civil Engineering, University of Waterloo, Canada, Jan. 1987.
13. *Software User's Manual IRD WIM Data Collection System, Version 7.4.3.* International Road Dynamics, Inc., Saskatoon, Saskatchewan, Canada, 1993.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.

Highway Accident Data Analysis: Alternative Econometric Methods

PATRICK S. MCCARTHY AND SAMER MADANAT

The past decade has seen significant advances in econometric modeling including the analysis of disaggregate data, the structure of discrete response models, the treatment of simultaneity in linear models, the specification of models based on pooled time series–cross sectional data, and the estimation of models in truncated and censored samples. Furthermore, the data sets available in the field of highway safety include significant amounts of detailed information. However, to date, highway safety analyses using these data sets have not fully exploited state-of-the-art econometric methodologies. The applicability of recently developed econometric methods to the field of highway safety analysis is illustrated. It is anticipated that such applications will improve the accuracy of traffic accident models and lead to more effective policies and investment decisions in the area of highway safety.

Highway safety is an area of research characterized by a disparity between data collection and data analysis. At the state and federal level, significant amounts of detailed information are routinely collected on highway traffic accidents. The amounts and types of data collected are of particular interest to the research community because they enable the researcher to investigate aspects of highway safety using state-of-the-art statistical and econometric methodologies. However, despite significant econometric advances during the past decade that potentially have important implications for improving understanding of factors that affect highway safety, there has been relatively little research identifying and evaluating the potential gain from these new methodologies. This paper constitutes a small step in this direction.

In the following sections, several areas, including policy endogeneity, cross sectional heterogeneity, and small numbers problems, are identified that illustrate problems with existing methodologies and offer alternative econometric techniques to correct the problem. In addition, other econometric issues including sample truncation and ordinality of accident severity data are discussed to illuminate often implicit assumptions associated with existing methodologies.

POLICY ENDOGENEITY AND HIGHWAY SAFETY

Consider the following equation:

$$y_t = \alpha + \beta' x_t + e_t \qquad t = 1, \ldots, T \tag{1}$$

where

y_t = highway safety outcome (e.g., fatality rate),
x_t = vector of k explanatory (exogenous) variables,
e_t = error with mean 0 and constant variance
α = parameter, and
β = parameter vector with k elements.

Assuming away problems of heteroscedasticity and autocorrelation, ordinary least squares (OLS) estimates of the unknown parameters will be best linear unbiased estimates (BLUE). Assume that the kth explanatory variable reflects a policy that was enacted to enhance highway safety (e.g., speed limits laws, mandatory seat belt use laws, minimum drinking age laws, etc.). If the policy was truly exogenous (the original reduction of speed limits in 1975 was a response to the oil crisis instead of to highway safety concerns), then the resulting parameter estimates will be BLUE. Alternatively, however, suppose that the policy was a reaction to concerns about highway safety. Then x_{kt} is itself a function of a set of explanatory variables including v_t. For example, reluctance in the United States to increase speeds after the oil crisis ended was a response to the life-saving effects of the lower speed limit. In this case, Equation 1 is actually a system of two equations that can be expressed as

$$y_t = \alpha + \beta' x_t + e_t \qquad t = 1, \ldots, T \tag{2a}$$

$$x_{kt} = \gamma + \delta' z_t + \Phi y_t + u_t \qquad t = 1, \ldots, T \tag{2b}$$

where

x_{kt} = kth explanatory variable in x_t,
z_t = vector of k' explanatory variables,
u_t = error term with mean 0 and constant variance,
γ = constant term,
δ = vector of k' parameters, and
ϕ = parameter of the endogenous variable y_t.

If the estimation sample is a time series data set, then one could apply Granger causality tests to check for endogeneity between y_t and x_{kt}. Granger (*2*) and Sims (*3*) developed tests to evaluate the direction of causality. To test the hypothesis that "x_{kt} does not cause y_t," regress y_t on lagged values of y_t and lagged values of x_{kt}; run a second regression of y_t on lagged values of y_t only. An F-test based upon the error sum of squares in the unrestricted and restricted regressions, respectively, can be used to test the null hypothesis. A similar set of regressions is run with x_{kt} as the dependent variable and lagged values of x_{kt} and y_t as the explanatory variables. In this case, the null hypothesis is "y_t does not cause x_{kt}." To conclude that "x_{kt} causes y_t," it is necessary that the null hypothesis is rejected in the first set of regressions and accepted in the second set of regressions.

If the true state of the world is Equation 2 but the structural relationship between x_{kt} and y_t is ignored, then the OLS estimates of Equation 1 will produce biased and inconsistent parameter estimates (*1*). To avoid the endogeneity bias, the analyst typically

P. McCarthy, Department of Economics; S. Madanat, Department of Civil Engineering, Purdue University, West Lafayette, In 47907.

implements either of two strategies, depending upon the objective. First, the analyst wants to capture the influence of y_t on x_{kt} but the structural relationship between x_{kt} and y_t is not important. In this case, x_{kt} in Equation 2b is simply substituted in Equation 2a and the resulting reduced form equation is estimated. It is important to recognize, however, that in this case the parameter estimates are not capturing the structural impacts of the explanatory variables on highway safety but instead the combined effects of these variables from the system of equations. Second, if one is concerned with the structural effect that a particular (endogenous) policy has on highway safety policy, then the simultaneous structure must be estimated and the structural parameters recovered.

To illustrate, consider the effect that recent relaxation of rural Interstate highway speed limits has on highway safety. Although enacting higher speed limits reflects a state's underlying demand for higher limits, it is also likely that a state's demand for higher speeds depends on the extent to which the affected roads are currently safe. Define speed law to be a variable that reflects the relaxation of rural interstate speed limits. If this variable is entered in an estimating equation for highway safety as a dummy variable that equals 0 in the 55 mph environment and 1 in a 65 mph post-law environment, the resulting estimates could be biased if enactment of relaxed speed limits were, at least in part, a response to changes in highway safety on the 55 mph roads. If so, the relaxed speed limit on rural Interstates is endogenous and the coefficient estimate on a speed limit dummy variable is biased.

To account for endogenous effects of highway safety on speed limit legislation, consider the following system of three equations (*4*):

$$y = \beta_1 x_1 + \alpha(\text{speed law}) + \gamma A + e_1$$
$$A = \beta_2' x_2 + e_2$$
$$A^* = A + \kappa(y^*) \tag{3}$$

where

y = measure of highway safety,
y^* = y in the absence of the relaxed speed limit,
A = latent variable reflecting attitudes towards relaxed speed limits
A^* = latent variable that reflects the demand for speed limit relaxation,
speed limit = dummy variable that equals 1 when the speed limit was relaxed and 0 otherwise,
x_1 = vector of explanatory variables for highway safety,
x_2 = vector of explanatory variables for A, and
e_i $(i = 1,2)$ = error term. The relationship between A^* and speed law is given by the following:

$$A^* = A + \kappa(y^*) > 0 => \text{speed law} = 1$$
$$A^* = A + \kappa(y^*) < 0 => \text{speed law} = 0$$

Thus, the hypothesis is that highway safety depends upon a set of explanatory variables, x_1, the speed limit, and the state's preferences for higher speeds, which depend upon a set of explanatory variables, x_2. The state's demand for higher speeds, in turn, reflects its attitudes toward higher speeds and the incidence of accidents, y^*, in the lower speed environment. This produces two estimating equations

$$y = \beta_1' x_1 + \gamma\,(\beta_2' x_2) + \alpha(\text{speed law}) + u_1 \tag{4}$$

and

$$A^* = \beta_2' x_2 + \kappa[\beta_1' x_1 + \gamma(\beta_2' x_2)] + u_2$$
$$= \kappa(\beta_1' x_1) + (1 + \kappa\gamma)\,\beta_2' x_2 + u_2 \tag{5}$$

where $u_1 = e_1 + \alpha e_2$ and $u_2 = e_2 + \kappa e_1$. Estimation is a two-step process. First, Equation 5 is estimated where A^* is replaced with speed law. The predicted value of the dependent variable in Equation 5, speed law, gives the demand for relaxing the speed limit and replaces speed law in the estimating Equation 4. Note that the full set of structural parameters can be recovered. From Equation 4, estimates of β_1' and α are obtained; dividing the coefficient of x_1 in Equation 5 by the coefficient of x_1 in Equation 4 gives κ; dividing the coefficient of x_2 in Equation 4 by that in Equation 5, and knowing κ, enables one to solve for γ; from Equation 4, knowledge of γ produces β_2'.

In the estimations, the components of x_1 would include standard determinants of highway safety (e.g., young drivers per capita, alcohol consumption, per capita income, and time trend). x_2 reflects a state's attitudes towards raising the speed limit and could be measured by two variables: excess speed, defined as the extent to which observed average speeds on 55-mph roads exceed the 55 mph limit, and speed variance, the variance of speed on 55-mph roads. Since the primary benefit of an increased speed limit is travel time savings, observed average speeds that are above the mandated 55-mph limit are tantamount to the driving population revealing its demand for higher speeds. This suggests that excess speed is positively correlated with drivers' sentiments toward raising the speed limit. On the other hand, authorities have a responsibility for providing safe driving environments and will not be inclined to raise the speed limit if it is believed to compromise highway safety. Thus, the net effect of excess speed on the demand for raising the speed limit is ambiguous and depends upon the magnitudes of these two effects.

Lave (*5*) has shown that increases in speed variance, all else constant, reduce highway safety. Consistent with this, the demand for raising the speed limit would be expected to be negatively related to speed variance.

Using this methodology, Saffer and Grossman (*6*) estimate a model in which highway safety and a state's drinking age policy are endogenous. McCarthy and Ziliak (*7*) use a similar framework to analyze the simultaneity between highway safety and the formation of Mothers Against Drunk Driving chapters.

TRUNCATION

Most analyses of the policy effects on highway safety base these results upon an OLS model (simple or reduced form) in which the dependent variable is some measure of highway safety. Because highway safety policy strives to reduce the incidence of the most serious accidents, namely, those involving a fatality, a frequently used measure of highway safety is some function of highway fatalities (fatal accidents, fatality rate, fatalities per capita.) By limiting the analysis to accidents involving a fatality, the sample is truncated from below because it excludes observations on all individuals who have experienced nonfatal accidents in the sample period. Thus, the estimates of the effect of a policy on highway safety are likely to be biased. Figure 1 illustrates this graphically. By excluding those accidents below severity level SC, the effect of increasing speed on highway safety is seen in the figure to bias the slope parameter downward and the intercept upward.

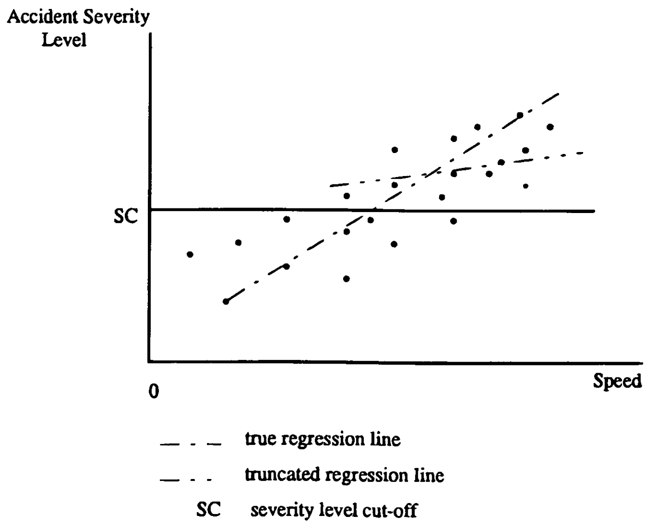


FIGURE 1 Potential bias from truncation.

To quantify the potential bias, consider a problem in which the dependent variable y is highway safety, x is a vector of K independent variables and β is a vector of parameters to be estimated. The underlying sample is a cross section of states or counties. Oftentimes, an OLS model of the following form is estimated

$$y_i = \beta' x_i + u_i \qquad i = 1, \ldots, \mathrm{N}$$

where u_i is a normally distributed error term with 0 mean and constant variance. Let y_i be measured as the number of fatalities or the fatality rate for Cross Section i. Define SC to be the level of truncation (e.g., AIS severity level) such that all accidents for which $y_i \leq \mathrm{SC}$ (e.g., AIS ≤ 4) are eliminated. The density function for the truncated variable y_i is

$$g(y_i) = \frac{(1/\sigma)\phi[(y_i - \beta' x_i)/\sigma]}{1 - \Phi\,[(\mathrm{SC} - \beta' x_i/\sigma)]} \qquad y_i > \mathrm{SC}$$
$$= 0 \qquad \text{otherwise} \tag{6}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density function and the distribution function of the standard normal, respectively (*4*). The log-likelihood for this function is given as

$$\log \mathrm{L} = -\mathrm{N}\ln[(2\pi)^{\frac{1}{2}}\sigma] - (\tfrac{1}{2})\sum_n \left(\frac{y_i - \beta' x_i}{\sigma}\right)^2 - \sum_n \ln\left[1 - \Phi\left(\frac{\mathrm{SC} - \beta' x_i}{\sigma}\right)\right]$$

which can be shown to be globally concave. Standard Newton-Raphson techniques can be used to obtain the maximum likelihood parameter estimates for β and σ. Once estimated, the parameter estimates are used to obtain the conditional mean and variance of y_i. In particular,

$$E(y_i | y_i > \mathrm{SC}) = \beta' x_i + \sigma\lambda(t_i)$$
$$V(y_i | y_i > \mathrm{SC}) = \sigma^2[1 - \delta(t_i)] \tag{7}$$

where $\lambda(t_i) = \phi(t_i)/1 - \Phi(t_i)$ and $\delta(t_i) = \lambda(t_i)(\lambda(t_i) - t_i)$ [which lies in the open unit interval (*1*)] and $t_i = \mathrm{SC} - \beta' x_i/\sigma$. Thus, the bias is $\sigma\lambda\,(t_i)$, which can be shown to be increasing in SC. By excluding accidents with severity levels below SC, the effect of explanatory variables, including policy variables, on highway safety will be biased. To determine the marginal effect of an increase in x_{ki} ($i = 1, \ldots, I$; $k = 1, \ldots, K$), differentiate the conditional mean with respect to x_{ki}. This gives

$$\text{marginal effect of } x_{ki} = \beta_k[1 - \delta(t_i)] \tag{8}$$

which is less than β_k since $\delta(t_i)$ lies between 0 and 1. In the subpopulation, the marginal effect of each of the k explanatory variables on y_i is less than the coefficient β_k.

To illustrate the potential importance of this bias, return to the graph showing the effect of increasing speeds on highway safety. If the analyst is concerned only with the subpopulation of accidents above severity level SC, then the marginal effect identified in Equation 8 is the relevant effect; alternatively, if the objective is to identify the effect for the entire population, then β_k is the relevant effect. Thus, an inference on the highway safety effects of higher speeds drawn from an analysis based upon a subpopulation of fatal accidents will understate the effect if applied to the entire population.

In the literature, criteria for truncation include accident severity (*8–10*), age of driver (*11–14*), alcohol involvement (*15–17*), number of vehicles involved (*18*), and vehicle size (*19–22*).

ENDOGENOUS STRATIFICATION

A related problem is endogenous stratification. As indicated, most models identify fatalities (or fatality rate, fatal accident rate) as a measure of highway safety to the general exclusion of other accidents (serious injury, minor injury, and property damage accidents). In this case, the sample is stratified according to accident severity. Lower severity accidents are often undersampled or completely absent. A generalization of the truncated sample is to analyze a model that identifies various severity strata and their sampled proportions. This may be an important stratagem for reducing costs because of the size of accident data files (statewide as well as nationally). In California, for example, there are more than 500,000 accidents annually.

To illustrate, suppose that a researcher uses the regression model $y_i = \beta' x_i + u_i$ ($i = 1, \ldots, N$) to identify the determinants of statewide highway safety. u_i is normally distributed with mean 0 and variance σ^2. Instead of obtaining the complete set of accident records, the analyst takes a p_1 and p_2 percent sample of fatal ($y_i > \mathrm{SC}$) and nonfatal accidents ($y_i \leq \mathrm{SC}$) respectively. Note that for the truncated model in the previous section, p_1 percent = 1 and p_2 percent = 0.

The density function for y_i is now given as (*4*)

$$g(y_i) = \frac{p_1 f(y_i)}{p_1 Pr(y_i > S) + p_2 Pr(y_i \leq S)} \qquad y_i > \mathrm{SC}$$
$$= \frac{p_2 f(y_i)}{p_1 Pr(y_i > S) + p_2 Pr(y_i \leq S)} \qquad y_i \leq \mathrm{SC} \tag{9}$$

where $f(y_i)$ is the density of y_i in the population. Substituting for $f(y_i)$ and t_i (defined as in the previous section) gives

$$g(y_i) = \frac{p}{p + (1-p)\Phi(t_i)} \frac{1}{\sigma} \phi\left(\frac{y_i - \beta' x_i}{\sigma}\right) \qquad y_i > \text{SC}$$

$$= \frac{1}{p + (1-p)\Phi(t_i)} \frac{1}{\sigma} \phi\left(\frac{y_i - \beta' x_i}{\sigma}\right) \qquad y_i \leq \text{SC} \qquad (10)$$

where $p = (p_1/p_2)$. These expressions can be used to form the log-likelihood function, which is then maximized with respect to β and σ (if p is not known, the log-likelihood function can be maximized with respect to p as well). The conditional means for this model are

$$E(y_i|x_i, y_i > \text{SC}) = \beta' x_i + E(u_i|y_i > \text{SC})$$

$$= \beta' x_i + \sigma \frac{\phi(t_i)}{1 - \Phi(t_i)} \qquad (11a)$$

$$E(y_i|x_i, y_i \leq \text{S}) = \beta' x_i + E(u_i|y_i \leq \text{S})$$

$$= \beta' x_i - \sigma \frac{\phi(t_i)}{\Phi(t_i)} \qquad (11b)$$

and the unconditional mean is a weighted average of the conditional means

$$E(y_i|x_i) = \beta' x_i + \sigma \frac{p_1\phi(t_i) - p_2\phi(t_i)}{p_1[1 - \Phi(t_i)] + p_2\Phi(t_i)}$$

$$= \beta' x_i + \sigma\, \gamma(t_i) \qquad (12)$$

Similar to the comments made in the previous section, depending on whether the analyst is concerned about the marginal effect of x_{ki} on the estimation subpopulation or its effect on the entire population, the appropriate marginal effect is obtained by differentiating Equations 11 and 12, respectively, with respect to x_{ki}. It can be shown that the marginal effect of the unconditional mean with respect to x_{ki} is $\beta_k\,[1 - \delta'(t_i)]$ where $\delta'(t_i)$ equals $\gamma(t_i)\,[\gamma(t_i) - t_i]$, which is similar to the expression for $\delta(t_i)$ given below Equation 7. Also note that if $p_1 = 1$ and $p_2 = 0$, then the marginal effect obtained from Equation 12, $\beta_k\,[1 - \delta'(t_i)]$ is identical to the conditional marginal effect (Equation 8) for the truncated model in the previous section. That is, $\delta'(t_i) = \delta(t_i)$.

Although this would appear to be a useful procedure for obtaining meaningful highway safety results while reducing the effort and computational burden associated with analyzing statewide or national accident records, the authors are not aware of any studies in the highway safety literature that use this methodology.

CROSS SECTION–TIME SERIES

Continuous Dependent Variable

State and national highway agencies routinely collect highway safety data that are organized into monthly and annual reports. In that these reports often discriminate by state, by county within state, by type of road, by various socioeconomic characteristics, and along numerous other dimensions, the information represents a panel of data—a time series of data across a set of cross section units.

When a time series of cross sections is available, ordinary least is generally not appropriate because it ignores the heterogeneity in the cross-sectional units. There are generally two methodologies for estimating panel data. First, a fixed-effects approach includes dummy variables for each of the cross-sectional units. This model assumes that differences between cross-sectional units can be captured by a parametric shift in the regression line. If all cross-sectional units are represented in the sample (e.g., all states in the nation, all vehicle types, all times of day), then a fixed-effects approach may be appropriate because it embodies all the differences among the cross-sectional units. However, if the cross sections represent a sample from a larger population (e.g., a subset of states), then it may be more appropriate to assume that the cross-sectional heterogeneity is randomly distributed across cross-sectional units. This latter approach represents a random effects specification.

In general, a cross section–time series model can be expressed as

$$y_{it} = \alpha + \sum_{j=1}^{k} \beta_j x_{it,j} + \epsilon_{it} + \eta_i \qquad (13)$$

where

y_{it} $(i = 1, \ldots, N; t = 1, \ldots, T)$ = highway safety outcome for cross section i and time period t,
$x_{it,j}$ $(i = 1, \ldots, N; t = 1, \ldots, T; j = 1, \ldots, k)$ = jth explanatory variable for cross section i and time period t,
α = constant term and $\beta_j (j = 1, \ldots, k)$ is a parameter that reflects the marginal effect of the jth explanatory variable on the highway safety outcome,
ϵ_{it} = error term associated with cross section i and time period t with mean 0 and constant variance, and
η_i = term specific to cross section i.

In the absence of any cross-sectional heterogeneity, η_i is equal to 0, and OLS is used to estimate the model. For a fixed-effects specification, η_i is a parameter that is estimated along with β_j, where η_i represents a parallel shift in the regression line for cross-section unit i. In a random effects model, η_i is assumed to be a random term with mean 0 and constant variance that is specific to cross section unit i. Notice that cross-sectional heterogeneity is confined to the error term in the random effects model, whereas in the fixed effects model it is explicitly represented as a parametric shift in the regression line.

In general, there are advantages and disadvantages to either approach. A fixed-effects specification entails a potentially large decrease in degrees of freedom if there are a high number of cross sections in the sample. In addition, fixed effects models cannot be estimated if any of the explanatory variables is constant throughout the sample period. On the other hand, if the fixed effects parameters are correlated with the included variables but omitted from the model, then a random effects specification leads to biased parameter estimates (*23*). Hausman (*24*) developed a specification test, based on a chi-squared statistic, to test the null hypothesis that the cross section–specific parameters in a fixed effects model are independent of the included explanatory variables. Accepting the null hypothesis would be consistent with a random effects specification, whereas rejecting the null hypothesis would argue for a fixed effects specification.

There have been a number of recent examples in the literature (*14,16,17*) of panel data analyses using accident data.

Discrete Dependent Variable

An interesting variation of that problem occurs when the dependent variable takes on very small integer values. For example, suppose a researcher is interested in modeling the incidence of countywide alcohol-related fatal accidents or countywide fatal accidents among teenagers. If this study were undertaken by state transportation departments, it is likely that in many states there would be a large number of counties in which very few or no fatal accidents occurred. As an example, consider Indiana, which has 92 counties. In 1989, there were 99 alcohol-related fatal accidents and 2923 alcohol-related injury accidents statewide, which represents an average of just over 1 and 31 alcohol-related fatal accidents and injury accidents, respectively.

One methodology for modeling these accidents is to estimate a logit model that defines the dependent variable y_{it} to be one if an accident in cross section i and time period t involved a fatality and 0 otherwise. In particular, the probability of a fatal accident is given by

$$P(\text{fatal accident}) = P(y_{it} = 1) = \frac{e^{\beta' x_{it}}}{1 + e^{\beta' x_{it}}}$$

$$i = 1, \ldots, N;\ t = 1, \ldots, T \qquad (14)$$

Similar to the continuous case discussed, if the cross-section units are heterogeneous and the heterogeneity is ignored, then estimating Equation 14 will lead to inconsistent parameter estimates.

Consider an alternative model that incorporates cross-sectional parameters, α_i ($i = 1, \ldots, N$) to reflect the underlying heterogeneity. Then Equation 14 becomes

$$P(\text{fatal accident}) = P(y_{it} = 1) = \frac{e^{\alpha_i + \beta' x_{it}}}{1 + e^{\alpha_i + \beta' x_{it}}}$$

$$i = 1, \ldots, N;\ t = 1, \ldots, T \qquad (15)$$

For large N and small T (≤ 5) Chamberlain (25) devised a method for estimating this model that is based on conditional maximum likelihood functions that do not depend on the heterogeneity parameters. Moreover, on the basis of a Hausman test of the null hypothesis that the cross-section units are homogeneous, it is possible to test a standard logit specification in Equation 14 against the alternative specification given by Equation 15.

To date, the authors are aware of no studies in highway safety using a panel logit methodology.

MODELS WITH ORDINAL DEPENDENT VARIABLES

Accident data generally obtained from police records are disaggregate data. However, when such data are used to analyze the effect of various factors on accident severity, they are usually aggregated and analyzed by using classical statistical methods such as multivariate regression. These methods are limited to the analysis of continuous variables, such as the total number of accidents, or the total number of fatalities and hence require that the data on individual accidents be aggregated before analysis. This is especially the case when accident data are recorded on an ordinal instead of a continuous scale. For example, the National Safety Council (26) devised a scheme for injury classification—no injury, possible injury, non incapacitating injury, incapacitating injury, and fatal injury.

With such a scale, an order is established between different categories of injury, but the distance between any two numbers on the scale is of unknown size. As such, these data cannot be analyzed by using traditional statistical methods, except by aggregation and subsequent loss of information.

In a recent paper, Nassar et al. (27) proposed using a sequential logit approach for modeling accident severity using disaggregate accident data. Such a model structure implies that an accident moves up the scale of severity, starting from the least severe. After each move, the accident either moves up one more notch or stays at its current level of severity. By assuming independence across error terms of the different logit models, the authors end with a model that is a product of binary logits. Each logit model is of the form

$$P(S_m | S_{@m-1}) = \frac{\exp(\sum_j \beta_{jm} X_{jm})}{1 + \exp(\sum_j \beta_{jm} X_{jm})} \qquad (16)$$

$P(S_m | S_{@m-1})$ = probability of experiencing injury severity level m given that the impact is sufficient to produce at least an injury of severity level $m - 1$.
X_{jm} = impact of factor j on severity level m.
β_{jm} = coefficient associated with factor j on severity level m.

With such a modeling approach, the richness of information available at the disaggregate level is exploited. Sequential choice models, however, are restrictive in the sense that they assume independence of the error terms across moves, for each accident, which may be an unrealistic assumption.

To relax the assumption of independence that the sequential logit approach imposes on the error terms, models with ordinal dependent variables such as the ordered logit should be used. Such an approach was specifically developed for models in which the dependent variable is ordinal, such as the accident severity ratings described (28). These models do not assume that the observed rating is the result of a sequence of move-ups; instead, the assumption is that the ratings represent a discretization of an underlying latent severity scale that is continuous. By using such an approach, it is possible to estimate jointly the parameters of the different severity factors and the thresholds that separate the successive severity ratings on the underlying latent scale. Mathematically, let the continuous underlying accident severity be denoted by y^*. Then, we have that

$$y^* = \sum_j \beta_j X_j + \epsilon$$

where

X_j = impact factor j,
β_j = coefficient associated with factor j, and
ϵ = random disturbance.

The process giving rise to the observed severity levels S_m ($m = 1, \ldots, M$) may be viewed in terms of y^* crossing some of the $M - 1$ threshold values. Specifically, we have that

$$m = 1 \text{ if } -\infty < \sum_j \beta_j X_j + \epsilon < t_1^*$$

$$m = 2 \text{ if } t_1^* < \sum_j \beta_j X_j + \epsilon < t_2^*$$

•

•

•

$$m = M \text{ if } t_{M-1}^* < \sum_j B_j X_j + \epsilon < +\infty$$

Thus, the probability of observing an injury severity S_m is

$$P(S_m) = P(t_{m-1}^* < \sum_j \beta_j X_j + \epsilon < t_m^*)$$
$$= P(\epsilon < t_m^* - \sum_j \beta_j X_j) - P(\epsilon < t_{m-1}^* - \sum_j \beta_j X_j)$$

If the error terms ϵ's are independently and identically distributed logistically, then the probabilities of various severity are given by an ordered logit model (*28*).

CONCLUSION

An overview has been presented of the potential application of some recent developments in econometric methodology to the field of highway safety analysis. Although no empirical work was presented, the data required to perform the analyses discussed are readily available to highway safety researchers.

In addition to the presented methodologies, there have been other modeling techniques, including empirical Bayesian analysis and Poisson methodologies (and variants thereof), which have been successfully although not frequently used to study the effect of traffic improvements at highway intersections (*29–32*). However, because these techniques are more familiar to traffic safety analysts than those identified here, they have been omitted from the overview.

Because major policy and investment decisions are often made by state and federal agencies on the basis of the results of highway safety analyses, the importance of accuracy in such analyses can hardly be overemphasized. By using state-of-the-art econometric methods such as those described herein, researchers can improve the level of accuracy in highway safety analysis.

REFERENCES

1. Greene, W. H. *Econometric Analysis.* Macmillan Publishing Company, New York, 1990.
2. Granger, C. W. J. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, Vol. 37, 1969, pp. 424–438.
3. Sims, C. A. Money, Income, and Causality. *American Economic Review*, Vol. 62, 1972, pp. 540–552.
4. Maddala, G. *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, Cambridge, England, 1983.
5. Lave, C. Speeding, Coordination, and the 55 MPH Speed Limit. *American Economic Review* Vol. 75, 1985, pp. 1159–1164.
6. Saffer, H., and M. Grossman. Drinking Age Laws and Highway Mortality Rates: Cause and Effect. *Economic Inquiry*, Vol. 25, 1987, pp. 403–417.
7. McCarthy, P., and J. Ziliak. The Effect of MADD on Drinking-Driving Activities—An Empirical Study. *Applied Economics*, Vol. 9, 1990, pp. 1215–1227.
8. Crandall, R., et al. *Regulating the Automobile.* The Brookings Institution, Washington D.C., 1986.
9. Graham, J., and S. Garber. Evaluating the Effects of Automobile Safety Regulation. *Journal of Policy Analysis and Management*, Vol. 3, 1984, pp. 206–221.
10. Zlatoper, T. Regression Analysis of Time Series Data on Motor Vehicle Deaths in the United States. *Journal of Transport Economics and Policy*, Sept. 1984, pp. 263–274.
11. Williams, A., et al. The Effect of Raising the Legal Minimum Drinking Age on Youthful Auto Fatalities. *Journal of Legal Studies*, Vol. 12, 1983, pp. 169–179.
12. Cook, P., and G. Tauchen. The Effect of Minimum Age Drinking Age Legislation on Youthful Auto Fatalities, 1970–77. *Journal of Legal Studies*, Vol. 13, 1984, pp. 169–190.
13. Males, M. The Minimum Purchase Age for Alcohol and Young-Driver Fatal Crashes: A Long-Term View. *Journal of Legal Studies*, Vol. 15, 1986, pp. 181–217.
14. Saffer, H., and M. Grossman. Beer Taxes, the Legal Drinking Age, and Youth Motor Vehicle Fatalities. *Journal of Legal Studies*, Vol. 16, 1987, pp. 351–374.
15. Wilkinson, J. Reducing Drunken Driving: Which Policies Are Most Effective? *Southern Economic Journal*, Vol. 54, 1987, pp. 322–334.
16. Chaloupka, F., et al. *Alcohol Control Policies and Motor Vehicle Fatalities.* Working Paper 3831, National Bureau of Economic Research, Inc., 1991.
17. McCarthy, P. The Effect of Higher Rural Interstate Speed Limits on Alcohol-Related Accidents. *Journal of Health Economics*, Vol. 11, 1993, pp. 281–299.
18. Colon, I. The Alcohol Beverage Purchase Age and Single-Vehicle Highway Fatalities. *Journal of Safety Research*, Vol. 15, 1984, pp. 159–162.
19. Campbell, K. L., et al. *Analysis of Accident Rates of Heavy-Duty Vehicles.* Final Report. University of Michigan Transportation Research Institute, Ann Arbor, 1988.
20. Stein, H. S., and I. S. Jones. Crash Involvement of Large Trucks by Configuration: Case Control Study. *American Journal of Public Health*, Vol. 78, 1988, pp. 491–498.
21. Jovanis, P., H. Chang, and I. Zabaneh. A Comparison of Accident Rates for Two Truck Configurations. In *Transportation Research Record 1249*, TRB, National Research Council, Washington D.C., 1989, pp. 18–29.
22. Mingo, R., et al. *Safety of Multi-Unit Combination Vehicles: Report on Findings.* Final Report, Association of American Railroads, Sept. 1990.
23. Judge, G. G., et al. *Introduction to the Theory and Practice of Econometrics.* 2nd Edition. John Wiley and Sons, Inc., New York, 1988.
24. Hausman, J. Specification Tests in Econometrics. *Econometrica*, Vol. 46, 1978, pp. 69–85.
25. Chamberlain, G. Analysis of Covariance with Qualitative Data. *Review of Economic Studies*, Vol. 47, 1980, pp. 227–238.
26. National Safety Council. *Manual of Classification of Motor Vehicle Traffic Accidents.* 5th edition, 1989.
27. Nassar, S., F. Saccomanno, and J. Shortreed. Nested Logit Model of Road Accident Severity. Presented at 72nd Annual Meeting of Transportation Research Board, Washington, D.C., 1993.
28. Barnard, P. How Some Advances in Econometrics Might Be Used in Road Crash Analysis. Australian Road Research Board Report, ARR 171, 1989.
29. Hauer, E. An Application of the Likelihood/Bayes Approach to the Estimation of Safety Countermeasures Effectiveness. *Accident Analysis and Prevention*, Vol. 15, 1983, pp. 287–298.
30. Hauer, E. On the Estimation of the Expected Number of Accidents. *Accident Analysis and Prevention*, Vol. 18, 1986, pp. 1–12.
31. Ben-Akiva, M., A. Ceder, L. Cheng, and M. Liss. Estimation of Intersection Traffic Safety Improvements. Final Report, Massachusetts Department of Public Works, 1988.
32. Hauer, E., and B. Persaud. How to Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effect of Warning Devices, In *Transportation Research Record 1114*, TRB, National Research Council, Washington, D.C., 1990, pp. 131–140.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.

Application of Automated Records Linkage Software in Traffic Records Analysis

KARL KIM AND LAWRENCE NITZ

Following a brief discussion of the underlying theory of records linkage, an automated record linkage software package called Automatch is examined along with its various applications. Features, hardware, and user requirements are discussed, and user support and interfaces are detailed and commented on. An example linking crash reports to ambulance records is described. After other possible applications and uses for this software are described, additional issues about records linkage are raised. The report is part of ongoing research carried out by the Hawaii Crash Outcome Data Evaluation System (CODES) project, funded by the National Highway Safety Traffic Administration, U.S. Department of Transportation. The purpose of the CODES Project is to link crash, EMS, hospital, claims, and long-term care data to conduct analyses on the effectiveness of seat belts, motorcycle helmets, and other traffic safety interventions.

At one level, linking records from different data bases raises interesting and complex questions about privacy and the uses of computerized data. When one considers the many data bases that have been created and adds the possibility of linkages among these data bases, frightening, Orwellian images could be invoked. At a second level, there are other questions involving how to accomplish such a task. Recent developments in the theory and methods of records linkage, including the release of a software package called Automatch (*1*), serve to enhance the feasibility of records linkage. These technological developments may in turn spark more debate and discussion on issues of the uses of data and appropriateness of linking diverse data bases together. This report deals principally with the second-order concerns—that is, how to use available methods and technology to carry out records linkage. While outside the scope of this paper, the basic concerns about the appropriateness and ethics of data linkage must also be addressed. Recent advances in technology point to some areas of concern that are summarized in the conclusions.

Perhaps every social science researcher has at one point linked or tried to link two different data bases. Typically, the data bases were collected by different agencies for different purposes, but they provide valuable information. Studies have linked land-use data with tax data, typically at the parcel level. Transportation data (car ownership, drivers licenses, etc.) could also be linked to housing data bases and to zoning data bases. School enrollment figures are often pooled with other data bases to derive estimates of population change. Other social data such as health statistics, crime surveys, and many different surveys and opinion polls are often used for purposes for which they might not have been initially designed. This is the nature of data collection—the cardinal rule is often to use existing data bases before expending the time and resources to gather what would amount to essentially the same information.

Department of Urban Planning, University of Hawaii, Porteus 107 2424 Maile Way, Honolulu, Hawaii 96822.

More data are becoming available in computerized form so that it is not at all unusual to pass machine-readable data (tape, diskette, or CD-ROM) between different users. At the same time, more users are becoming computer literate with the proliferation of PCs, workstations, and statistical packages. Yet the merging of different data bases still poses some basic difficulties. First, surveys and other data bases generally protect anonymity so that unique identifiers such as name or social security number and so forth are not used. Second, even if name, street address, or other identifiers are available, there are still problems with matching records because of inconsistencies across sources in data entry and editing procedures. For example, the use of initials instead of full names, different abbreviations for street names, and the usual assortment of misspellings and other errors in the data base make exact matches impossible.

Several years ago, a survey on attitudes toward helmet laws in Hawaii was conducted. To construct a sample, two different data bases were linked: the vehicle registration file and the operator's license file. For the city and county of Honolulu in 1989, there were 8,514 registered motorcycles. (Military Personnel were excluded from the population.) For the same year, 13,595 persons held motorcycle licenses. It is not expected that everyone who has a motorcycle license owns a motorcycle and vice versa. Yet in terms of producing the best sample of motorcyclists, it appeared reasonable to construct a single file consisting of those who both were licensed and owned motorcycles. Because of misspellings, differences in punctuation, and other differences in the information contained in the two files, few records could be exactly matched. A matching strategy was devised to organize records in both files around the name field, then to match on the basis of last name, first name, street address, and zip code using the Statistical Analysis System (SAS) statistical package. Once the exact matches were located by computer, all remaining pairs were reviewed manually. On the basis of name and address, only 2,970 cases were matched, less than 35 percent of the registered motorcycles.

Manual review took many hours and, serious problems are associated with this procedure. Certain people, particularly those who tended to move or change addresses frequently, were more likely to be excluded, which could introduce certain biases. Some individuals in the ownership file owned more than one motorcycle and therefore showed up as duplicates in the ownership file but as unique records in the license file. Finally, uncontrolled error was introduced by the manual review process—in addition to being tiresome work, the process of comparing records to identify a match is tedious, particularly because it is difficult to devise a comprehensive set of decision rules without reviewing all the data. Each of these problems, from the duplicate records in the registration file to the problems associated with manual comparison of records, could have been handled more effectively with the Automatch software.

THEORY OF RECORD LINKAGE

Although a detailed mathematical discussion on the theoretical developments of record linkage is outside the purview of this paper, it is important to note that there have been many important theoretical developments. The Automatch software builds on the work of Felligi and Sunter (*2*), Newcombe and Kennedy (*3*), Newcombe (*4*), Jaro (*5*), and others. Although there are a few commercially available programs for records linkage, the Automatch program is experiencing growing popularity—not just among those interested in records linkage but also among those involved in more specialized activities of geocoding and data base management. The geocoding and record unduplicating features of the program will be discussed later.

To conceptualize the theory behind the Automatch software, one must begin with two different files. Each file contains fields of fixed length and a finite number of records. For records to match on these two different files they must share one or more equivalent fields. If every common field contributes to linking, the larger the number of common fields in the two files, the greater the opportunities for linking the two files. The record-matching process involves pairing records from the two files and determining whether a given record pair can be considered a match or a nonmatch. For any two files, there are always many more unmatched pairs than true matches. In two files, each of which contains exactly 500 records, the possible number of record pairs would be 500 × 500, or 250,000 possible record pairs. Because there are only 500 records in each file, the maximum number of matches one could hope to produce is 500 (assuming no duplicates in either file and perfect matches between the two files). The basic idea is to use common fields in both fields to match records. Each of the matching fields has certain properties that affect its performance as a matching variable. Some fields (e.g., date of birth, name, social security number) contain many different possible values. A match on one or more of these fields greatly increases the probability of a match between two records. On the other hand, many of these fields have a higher likelihood of errors and inconsistencies at the time data are collected and entered. Other fields, such as gender, zip code, political party affiliation, or other attributes with a limited number of possible values, may be more accurately entered but do little by themselves to increase the likelihood of matching record pairs. Of course, taken together matches on many individual fields help increase confidence of an overall match between records that have been paired. The matching algorithm involves determining the extent to which any individual field, as well as the summation of all fields used in matching, contributes to the probability of a true match.

FEATURES OF AUTOMATCH SOFTWARE

Automatch was recently developed and marketed by Matthew A. Jaro (MatchWare Technologies, Inc.) of Silver Spring, Maryland. Jaro is a computer scientist who left the U.S. Bureau of the Census to form a software development firm. Although MatchWare Technologies has many of the problems associated with small start-up ventures, one advantage of its small scale is that customers can deal directly with the developer. Slick packaging and carefully edited training manuals received from most vendors are less valuable than the personalized and informed user support received from MatchWare. There are not many users—in part because records linkage tends to be a more specialized field within social science research (few basic courses on data base management and statistical analysis include record linkage as a topic). Moreover, although there is tremendous potential for new uses and abuses of this technology, it is, for the most part, an emerging technology that has not yet been widely implemented.

Automatch is currently available in a PC version running on the MS-DOS or OS/2 operating system. It is also available in Unix versions for running on workstation environments. There are some differences between the PC and Unix versions of Automatch, but many are a function of operating system and hardware characteristics instead of program differences. Obvious differences are processing speeds and memory management. Although the PC version can run with just 640K, the performance is acceptable with only relatively small data bases. On the other hand, running Automatch on a workstation allows for the handling of much larger files. For example on a Sparc 10, a 70,000 record file was matched against a 9,000 record file in under 5 min.

In some respects, the PC version is more user-friendly than the Unix version tested. With a color menu–driven system, the PC version of Automatch can be used by most who are familiar with data base management systems. The PC version was found to be especially good for training purposes. Users must be able to define file structures clearly, name variables, specify types and lengths, and understand the basic principles of records linkage. If one could not carry out the records linkage manually, it would not be possible to instruct the machine to do so.

Automatch is a collection of specialized programs that operate by indexing instead of sorting the original files to be matched. To use Automatch, a certain amount of file preparation must be done. The amount of preparation will depend on the nature of the data collected as well as inputting, editing, error checking, and other data management practices. The files must be standard ASCII files, with each line delimited with a carriage return. Records must be of fixed size. Automatch does not support records or fields of variable length. Automatch will support most character, numeric, date, street address, and other types of variables. There is a procedure for defining missing value codes, although Automatch does not recognize the SAS use of "." as a missing value. Automatch is not a substitute for a data base management system or a statistical analysis system. Although one byproduct of a matching exercise is the identification of errors in the files being matched, Automatch is not equipped to correct or modify the original files directly. Automatch calculates certain statistics and distributions, which are specific to the matching procedures and not particularly useful for description, analysis, or modeling of data. The Unix version, unlike the PC version, is not at present a menu-driven system, so users must be familiar with a good text editor to write the control files required to run the program. Users of the Unix version of Automatch must have some elementary programming skills as small files are written and compiled to control the linkage process. Anyone who has written batch command files in DOS or has written control files in statistical packages such as SAS or statistical package for social sciences (SPSS) should have no difficulty mastering the Automatch system.

The program is designed so that users begin by assigning a project name that is used in all steps of the linkage procedure. The program generates various extensions that identify all the files for a given project. A first step in Automatch involves the creation of data dictionaries for the files to be matched. The data dictionary defines the location of the file, the record size, as well as variable names, positions, lengths, and missing value codes. The prepared dictionaries for the files are then compiled into binary format. It is important to note that if changes are made to the dictionary or the underlying data set the dictionary must be recompiled.

In addition, users must prepare a match specification file to control the matching algorithm. It identifies the type of program, sets out a blocking scheme to subset the file, and lists matching variables and the matching procedure to be used with each one.

Automatch contains three different types of matching programs: (a) MATCH—for matching between records on two files, (b) GEOMATCH—for matching a file to a geographic reference file (e.g., the 1980 Census DIME map files or the 1992 Census TIGER map and street address files), and (c) UNDUP—for identifying duplicate records within a single file.

The Automatch procedure is efficient because it breaks the matching process into two distinct steps: (a) blocking the data in each file into small groups using a few variables that partition the total file into subsets of similar cases and (b) indexing the blocks and running the match comparisons within each block using probabilities of matches developed in the indexing stage.

In this way a 70,000 case file can be tested for possible matches against a 100,000 case file without examining 7 billion possible match pairs, that is, without actually testing every unlikely case. Auto crashes involving male 45-year-old drivers need not be tested against ambulance calls to pick up female 16-year-old injury victims. Thus, by restricting the range of comparisons to a block of plausible cases, the number of actual tests is dramatically reduced.

Blocking involves the creation of homogeneous subsets formed around variables such as age or place of residence. The more blocks that are created, the smaller they will be, and, therefore, the more efficient will be the matching procedure. Automatch recommends block sizes of 100 records per file. The PC version has a block limitation of 32,400 pairs in a block (180 records per block). The best variables for blocking are those with a large number of possible values and a high degree of reliability.

Automatch also requires the user to specify the variables to be used for matching and the cutoff values for declaring matches. The matching variables must be different from those selected for blocking. The program accepts a variety of different types of variables (character, numeric, time, odd or even interval, etc.). Depending on the type of variable selected for matching, different approaches to comparison are used. For example, with character fields, a character-by-character comparison is carried out with shorter fields padded with trailing blanks to match the length of the longer field. Automatch also provides for an uncertainty character field in which tolerance for phonetic errors, transpositions, random insertions, deletions, and other differences between two fields can be set. Numeric fields involve a straight algebraic numeric comparison in which leading spaces are converted to zeros and numbers are compared. This is particularly useful for record data that have ill-defined columns or out-of-place number values. Automatch also has a delta percent comparison in which differences between fields should be measured in percentages. There are also allowances for interval data and odd or even intervals (useful for geocoding applications).

After specifying the fields to be used for matching in terms of their names and types, the user must specify two different subjective probabilities, *m* and *u*. The *m* probability is the probability that the field agrees, given that the record pair is a match. The *u* probability is the probability that the field agrees at random. Although the user must provide an initial estimate of these probabilities, some guidelines are given that are helpful. It is easier to begin by estimating the *u* probabilities. For a field such as gender, where there are only two possible outcomes, male and female, the probability is .5. Estimating the probabilities for fields with more possible values, say, zip code or date of birth, may be more difficult, but it forces the user to think about the characteristics of the particular fields selected for matching. In a similar manner, the user should estimate the *m* probabilities. This probability can be estimated by subtracting the error of the field from one and typically ranges between .9 and .99. The prospects of having to estimate these probabilities may seem somewhat daunting, but Automatch contains a program that can be run to update this probability (after a match run is executed) that is based on the actual characteristics of data included in the matching files.

After specifying the probabilistic matching parameters, the user must also specify the cutoff weights that signify the threshold levels for an acceptable automated match and those cases that require clerical review. The weights for a given field are calculated by taking the log to the base 2 of the ratio of *m* and *u* probabilities (if the fields agree) and the log (base 2) of the ratio 1-*m* and 1-*u* (if the fields disagree). In this way, fields that agree receive positive weights and those that disagree receive negative scores. A composite weight for the record pair is calculated by summing all of the individual field weights. The program produces a histogram of these composite weights. Records that have a high positive weight are assumed to match and those that have a low or negative weight are assumed to be nonmatches. Based on the distribution for all comparisons, users are able to discriminate between matches and nonmatches (see Figure 1). On the basis of this distribution, cutoff weights can be established, and those cases that require clerical review can be identified.

Only occasionally will users be able in a single pass to determine the matching specifications and produce a satisfactory match. It is clear that, with each pass of the matching algorithm, more information about the data is gleaned and can be incorporated into the selection of blocking and matching variables as well as in the selection of appropriate cutoff weights. By design, Automatch is meant to be iterative; it may take several passes before initiating clerical review.

The clerical review process involves classifying record pairs as matches or residuals (nonmatching records). A report-generator program is built into Automatch to facilitate clerical review. This program enables the user to view records and additional fields defined in the data dictionary to make an assessment about whether a record pair is indeed a match. The clerical review program allows the user to examine not only potential matches but also duplicate records that may have ended up in either of the two comparison

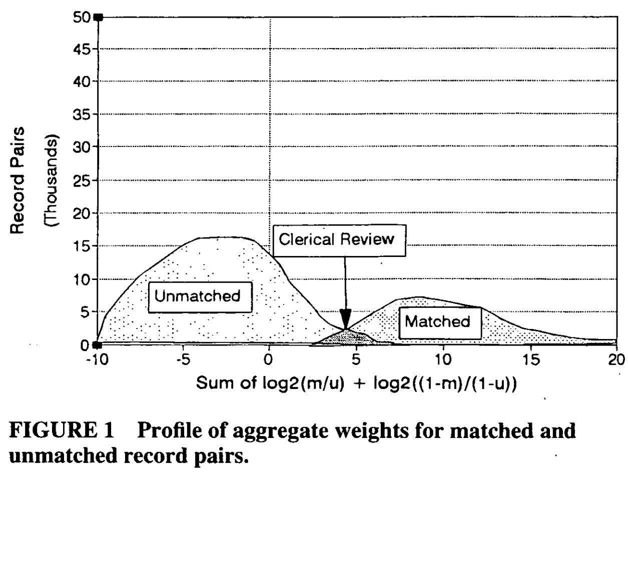


FIGURE 1 Profile of aggregate weights for matched and unmatched record pairs.

files. One feature of the clerical review algorithm is a history file, which keeps a record of all the decisions made by the clerical reviewer so that if the matcher program is rerun the user will not have to view the same records that have been previously reviewed.

The final step in Automatch generally involves producing an output file of the matched records that can then be imported into another package such as SAS or SPSSx for statistical analyses. Other special features are built into the Automatch program for handling geocoded data and producing other specialized reports related to the matching procedure. MatchWare Technologies has also developed address standardizers and other specialized programs and routines that are useful in file preparation and records linkage.

EVALUATION OF AUTOMATCH

In this section Automatch's performance on matching two data bases from Hawaii—the state Motor Vehicle Accident (MVA) file and the state Emergency Medical Services (EMS) Data Base—is described. Based on experience, an evaluation of Automatch's performance is provided. The data were matched as part of requirements for a federal grant to build a Crash Outcome Data Evaluation System (CODES), administered by the Department of Transportation. The Hawaii CODES project involves linking crash data to ambulance transport (EMS) data, hospital data, insurance data, and other information on traffic crashes in Hawaii during 1990. The purpose of the project is to build a linked data base on which models explaining crashes, driver behavior, and the effectiveness of safety devices on reducing injury and fatality can be tested.

The MVA data contains computerized records on all major traffic accidents in Hawaii. Data are collected by police officers called to the scene of a traffic collision. Data on driver, vehicle, occupant, and roadway characteristics are filled out on a paper form, which data in turn are entered into a computer system maintained by the Department of Transportation, Hawaii. In 1990, there were approximately 27,000 major traffic crashes, involving some 45,000 drivers, and an additional 30,000 occupants. The data suffer because it is collected by police officers under not ideal conditions and then entered by keypunchers who have few resources with which to check or verify the work. On the other hand, there are only four counties in Hawaii, and the data are more centralized than in many other states with more local law enforcement agencies.

The EMS data are also collected on a statewide basis and maintained by the Department of Health, Hawaii. This data base draws information from a dispatch card, which is filled out when a call for ambulance service comes in, and an EMS report, which is completed by ambulance attendants called to the scene of an accident. When the nonemergency, nontraffic-related ambulance runs are excluded from the EMS data base, approximately 9,000 ambulance runs must be accounted for.

The steps in records linkage involve cleaning the two data bases, preparing the fields for matching, devising a matching strategy, running several passes with the Automatch program, conducting clerical review, and preparing various summary reports.

In matching EMS records to MVA records, more time and effort went into the preparation, editing, and cleaning of the files than in conducting the matching procedure. Part of the reason for this has to do with the nature of public data bases that are maintained more for individual records reporting than for data analysis and modeling. A basic hardware problem involved extricating and decoding the MVA data from an antiquated Wang minicomputer system before it could be read on to the Sparc 10 workstation. For the MVA data, the data files were prepared, using SAS, by recording variables into a usable format and writing them to a text file using the "PUT" command. For the EMS data, dBASE was initially used because the data had been entered into a relational data base management system. Eventually, the data were transferred into SAS so that comparable matching variables could be constructed.

The blocking strategy was dictated by the nature of error in the original data. Blocks were to be as small as possible to provide near certainty that matches that were not physically possible would be prohibited. The most important variables for blocking were county and date. These were good variables because in Hawaii each of the four counties consists of separate islands, isolated by ocean. The date field was systematically edited and verified by EMS personnel. We also used gender as a blocking variable. These blocks enabled a match with greater certainty on the variables such as age, time of the incident and service, and location codes. The matching strategy produced a distribution of matched and unmatched pairs, including exact matches, duplicates, and clerical (manual) review cases (see Figure 2).

At the outset, it is important to note that Automatch's performance was impressive. First, few products comparable in cost or flexibility are available. Second, there is an underlying mathematical basis for the matching algorithm that is based on probability theory and enables the user to specify error ranges and accompanying levels of tolerance. Also the user is led through a logical sequence of data definition, developing a blocking and matching strategy, adjusting or correcting the strategy based on information generated through the match procedure, and can set parameters for clerical review. Automatch encourages the user to think systematically about the data that are being matched. Third, the level of technical support and the quality of customer service offered by MatchWare Technologies, Inc., has been superior.

```
******************************************************************
*
*  OUTPUT STATISTICS FOR MATCH: sec
*  PASS: 1
*
*     69072     Records on file A
*      9395     Records on file B
*         0     A residuals from previous pass
*         0     B residuals from previous pass
*     65795     A records read
*      9334     B records read
*      2012     Blocks processed
*         0     OVERFLOW blocks
*       173     Maximum A block size (including overflow)
*        30.8   Average A block size (not including overflow)
*        29     Maximum B block size (including overflow)
*         4.6   Average B block size (not including overflow)
*      6786     Matched pairs
*       227     EXACT matched pairs
*       537     Clerical pairs
*      4144     A duplicates
*         4     EXACT A duplicates
*       171     B duplicates
*         2     EXACT B duplicates
*     57605     A residuals (including skips & missing)
*      1901     B residuals (including skips & missing)
*      3745     A records skipped
*        30     B records skipped
*
******************************************************************
```

FIGURE 2 Sample output statistics from Automatch.

The documentation is clearly written and provides enough for most users to start using the Automatch software, but the documentation is thin (approximately 60 pages) and may not be enough for those who are doing records linkage for the first time. The program has been used several times, and the documentation now appears all the more clear and straightforward. The documentation falls short because no example is worked out all the way from start to finish with all inputs statements and screen outputs. The documentation tends to be one-sided, as it provides fairly good instructions in terms of statements and commands but leaves out the system responses, which, unless one has been through the entire matching procedure, are not the most informative. In using the system, the most common response was, "Now what?" In the spirit of DOS and Unix, in Automatch no response is a good response.

Another area with some degree of mystery involves the generation of program files and object files. Automatch generates several different files, so it would be useful to provide a more clear discussion of what files are created and how each is used in the system. The directory contents and watches were checked periodically to determine the effects of various programs and match runs to determine what new files were being created and modified. One would also like to have more information about the actual matching algorithm. Although Jaro's work (*5*) provides a basic understanding of what is going within the program, the documentation does not rehearse the algorithm in the context of a worked-out problem. More discussion is needed about the various user decisions that influence the matching procedure. Here too, an example or two worked all the way through from beginning to end, replete with the determination of composite weights and cutoff scores, would help bridge the gap between documentation and implementation. The concerns about the documentation are minor because these appear easily correctable deficiencies. Jaro (*5*) provided excellent technical support when needed.

Areas in which there is more room for improvement are user interfaces, screen calls, and the transitions between one program and the next. Although the PC version, with its menu-driven format is more user-friendly than the Unix version used, user interface support can be improved. It would be nice, for example, to have pull-down menus with program templates to serve as guides not only for writing individual programs but also for showing the sequence from one step to the next. Error messages could be improved so that debugging would be easier. Screen prompts emerge when submitting and executing commands, but many program files are prepared in batch format. When a program bombs, it is sometimes difficult for new users to figure out from error messages what went wrong. It would be useful to build a program editor into Automatch specific to the Automatch language so that illegal entries and inconsistencies would be flagged before compilation of the program.

Once one has a basic understanding of the principles of records matching and how to interpret the information provided by Automatch, then the actual records linkage becomes more challenging. One learns how to use Automatch by formulating blocking strategies, identifying matching variables, estimating the m and u probabilities, adjusting cutoff weights, and working the data bases to minimize the number of clerical reviews and retaining high confidence that the algorithm has done a good job of matching.

Special care is required in any form of raw-data procedure in which the underlying data are thought of as variables or measurements and would normally be indexed by named variables in a statistical analysis system or data base manager. This applies here as well. If the data sets to be merged are small and contain relatively few distinct variables, it is possible to create files by merging that encompass entire records. If the records are large, however—100 or 200 field records—it will be more efficient to extract a subset of variables that constitute the blocking, matching, and ID indicators for the set required to marry the subset back to the original records exactly. The user is tempted simply to use the sequence number of the record as the key for this purpose. This is not advisable for at least two reasons: (a) The order of the key is dependent on the order of the original file; if the file is transformed or, worse, sorted, the order is lost. (b) Any attempt to match the "unmatched" cases on a second pass, once the original matched records have been merged into a new, combined data set, is likely to define the "unmatched" cases as a subset of the original cases. The positional ID number of cases in the subset will be different from the ID number in the original file. (It is possible to execute successful matches on, say, 8,000 of 64,000 records from File A to File B, then return to try to match the 56,000 unmatched records from File A. To bring the results of this second matching process back into the master data set and to marry the cases correctly, unique identifiers must be carried into the subset used in the matching procedure.

It is easy to avoid the traps of this procedure by executing multiple blocking and matching steps within a single run of Automatch. In this way the criteria for matching can be upgraded on the basis of experience or new information, yet only one process will be used to merge the matched files back onto the master data base file.

Though Automatch has some rough edges, the product is fairly easy to use. Most computer-literate individuals can master the program in a few hours, provided that they start with a simple matching problem (with many good fields for comparison) and then graduate to a more difficult and realistic matching exercise. The flexibility of the program allows matching of different types of data sets and includes many special features that emerge more as one interacts with it. It is likely that most data base managers would find Automatch to be something that, once used, would be difficult to live without.

Other Applications and Uses

Automatch was developed for postenumeration surveys conducted by the U.S. Bureau of the Census. Undercounting of certain groups (minorities, non-English speakers, etc.) is suspected in some areas. Follow-up surveys are typically conducted in these areas. Automatch enables the comparison of individual records (between the original and follow-up survey) to find out which people were not counted the first time but were enumerated the second time around, to produce an estimate of undercounting.

Follow-up surveys, longitudinal questionnaires, and other applications that involve matching pairs for study over time could benefit from the use of a program such as Automatch. This is particularly useful when errors in data entry or substantial changes in population characteristics over time are concerns (*6*). Automatch enables matching to go beyond merely the use of one or two identifiers and permits many different kinds of variables to be used in records matching.

Another procedure in Automatch that identifies duplicate records would have many potential applications, from purging mailing lists of duplicates to removing duplicate records before updating a data base or performing statistical analyses on it.

The geocoding applications involve matching a particular data file with a reference file. For example, one could match data on crash locations (typically coded in terms of street name and mile marker or cross street) with latitude and longitude data from computerized street index files or geocoded reference files. As Geographic Information System (GIS) technologies continue to expand, more reference files (e.g., Census 1992 TIGER map files, Census Summary Tape File data organized by ZIP codes, etc.) have become available in a variety of formats. Automatch can be used with GIS and mapping technologies to bring this sort of summary data into a format suitable for mapping.

For data base management, Automatch may be particularly useful in those circumstances in which there is much transaction processing. Organizations with large data bases in which information is continually being updated and altered may find use for probabilistic matching for error detection and postaudit review. At present, Automatch is set up only for batch processing, yet one could imagine ways of applying the algorithms in a more interactive fashion.

Other applications for Automatch may be in the field of criminal justice research, where one could examine the relationship between, say, traffic citations, traffic collisions, criminal activity, and other forms of deviant behavior (e.g., DUI, drug use, etc.). In the future, firms specializing in records linkage might emerge—similar to the emergence of those that provide geocoding services in response to the growth of geographic information systems. Many different applications can be envisioned in urban and regional planning, for example, linking transportation data to land use and marketing studies, social services data to census data, and environmental quality studies to data on land use and ownership.

CONCLUSIONS

Automatch does much to elevate the level of sophistication of data base managers and others who enter, clean, and match data. Too often, the business of data base management and records linkage has been kept in the dark ages—that is, although there is much statistical, graphics, and presentation software, really new tools in data base management have been rare. Automatch is an exception. It provides data base managers a new arsenal of programs for matching data, identifying duplicate records, and handling assorted problems typically associated with geocoded data.

Automatch also opens doors for researchers and statistical modelers looking for ways of combining data bases. Through records linkage, new and interesting analyses can be carried out. Gaps among agencies, disciplines, time periods, and data sources can be bridged through records linkage. The potential uses and abuses of this technology are great. The prospect of linking specific public record data bases (property ownership or voter registration files) to attitude survey, market research, health, or financial reporting data bases presents enormous ethical and political challenges. With this software and some understanding of probabilistic records linkage, even files in which many of the common identifiers (e.g., name, social security number, etc.) have been stripped can be linked to other public data files (which could contain names, addresses, and other person-level identifiers). Although this paper is meant to provide an overview of the technology, there are also important questions of what constitutes appropriate and legal data linkages and major questions about maintenance of confidentiality and the uses of data for purposes other than for what they were collected. Certainly, one response to the technology may be to make it more difficult than ever to gain access to computerized data.

A more complete discussion of the ethics of records linkage must come before widespread application of this technology—although as often happens with innovation, progress often precedes policymaking. Records linkage is definitely on its way to becoming a more widespread practice. For planners to appreciate its potential and limitations more fully, more systematic discussion about appropriate plans, policies, and standards of practice for automatic records linkage must occur. Educators have an especially important role to play not only in teaching the technology of records linkage but also in conveying a critical understanding of ethical concerns as well.

Because of the existence of probabilistic matching software, real data hounds will undoubtedly discover ways of improving the quality and coverage of information that will only serve to improve and expand upon the nature and levels of analysis and model building. MatchWare Technologies, Inc., has already developed a small impressive set of clients, ranging from the U.S. Department of Transportation to various public- and private-sector organizations around the world. We predict that Automatch—in its present form and versions beyond—will become more widely used and that the practice of probabilistic records linkage, with all its opportunities and challenges, is here to stay.

ACKNOWLEDGMENTS

Automatch is available from MatchWare Technologies, Inc., 14637 Locustwood Lane, Silver Spring, Maryland 20905. The program was tested and run on a Sun Sparc 10 workstation at the Department of Urban and Regional Planning, University of Hawaii. Support for this research was provided by the Hawaii CODES Project, a cooperative research agreement between the U.S. Department of Transportation, National Highway Safety Administration, and the University of Hawaii. The authors acknowledge the support of the Department of Transportation, Hawaii, and the EMS Branch of the Department of Health, Hawaii. Graduate students in the Department of Urban and Regional Planning at the University of Hawaii (Richard Kirschenbaum, George Nabeshima, and John Valera) helped in the preparation of data files used in this procedure.

REFERENCES

1. *Automatch: Generalized Record Linkage System, User's Manual, Version 1.4.* MatchWare Technologies, Inc., Silver Spring, Md., 1993.
2. Felligi, I. P., and A. B. Sunter. A Theory for Record Linkage, *Journal of the American Statistical Association*, Vol. 64, 1969, pp. 1183–1210.
3. Newcombe, H., and J. M. Kennedy. Record Linkage, *Communications of the Association for Computing Machinery*, Vol. 5, 1962, pp. 563–566.
4. Newcombe, H. *Handbook of Record Linkage: Methods for Health and Statistical Studies, Administration, and Business*. Oxford University Press, Oxford, England, 1988.
5. Jaro, M. A. Advances in Records-Linkage Methodology as Applied to Matching the 1985 Census of Tampa, Florida. *Journal of the American Statistical Association*, Vol. 84, No. 406, 1989, pp. 414–420.
6. Newcombe, H. B., M. E. Fair, and P. Lalonde. The Use of Names for Linking Personal Records, *Journal of the American Statistical Association*. Vol. 87. No. 420, Dec. 1982, pp. 1193–1024.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.