

ACCIDENT RATES FOR HEAVY TRUCK-TRACTORS IN MICHIGAN

DANIEL BLOWER, KENNETH L. CAMPBELL, and PAUL E. GREEN

University of Michigan Transportation Research Institute, 2901 Baxter Road, Ann Arbor, MI 48109-2150, U.S.A.

(Received 9 July 1991; in revised form 28 March 1992)

Abstract—Accident rates of heavy truck-tractors are modelled using log-linear methods. The accident data used are a census of truck-tractor involvements in Michigan from May 1987 to April 1988. Travel data used to calculate the rates were produced by a survey of truck-tractors in Michigan covering the same time period. Both the accident and travel data were limited to Michigan-registered tractors operating in Michigan. Log-linear models of casualty and property-damage-only accident rates were developed using number of trailers, road type, area type, and time of day as predictor variables. Overall, differences between tractors with one and two trailers were not significant. Tractors with no trailers (bobtails) have significantly higher accident rates. Characteristics of the operating environment were found to have larger effects on the accident rate than tractor configuration (except for the bobtail). Rates varied by a factor of up to 6.8, depending on the road type. Casualty accident risk at night was 1.4 times the risk during the day. The risk of a casualty accident in rural areas was 1.6 times that of urban areas.

INTRODUCTION

Truck safety issues currently command a large share of the traffic safety interest of policy makers and the general public. Vehicle inspection and commercial drivers licensing programs have recently been put in place or significantly expanded, as have programs to ensure more complete reporting of truck accidents involving interstate carriers. Moreover, the perception that truck combinations such as doubles or triples are inherently unsafe has resulted in proposals to restrict their use.

As attention has focused on truck safety, it has become apparent that little is accurately known about the operating experience of large trucks. A recent Transportation Research Board special report surveyed data sources and noted that existing, publicly available data are so inadequate that there is no broad consensus on even how many trucks there are, much less how many miles they travel or how that travel is distributed across different operating environments (TRB 1990). On the accident side, the picture is mixed. The number of trucks involved in fatal accidents is fairly well established, from the Trucks Involved in Fatal Accidents program at the University of Michigan Transportation Research Institute, building on the Fatal Accident Reporting System (FARS) from the National Highway Traffic Safety Administration (NHTSA). For injury and property

damage accidents, the situation is quite different. Even the total number of truck accidents is uncertain, much less counts by truck configuration, road type, or other factors of interest. States have a difficult time correctly identifying medium and heavy trucks in their accident files. Estimates of total truck accidents from NHTSA's National Accident Sampling System (NASS) vary considerably from year to year. Reports filed with the Office of Motor Carriers (OMC) are another possible source, but, since only about 70% of reportable fatal accidents are filed with OMC, under-reporting is expected to be widespread for injury and property damage accidents, and only interstate carriers are required to report to OMC in the first place. When the fundamental data are uncertain, it is difficult to assess the factors that affect the relative risk of operating different types of trucks.

The present study quantifies the factors that affect the probability of accident involvement for truck-tractors. It focuses on the traffic safety experience of tractors registered and operated in Michigan. Accident and exposure data were collected. Accident rates were calculated for three tractor configurations, taking into account road type, time of day, area type (urban or rural), and accident severity. Multivariate models were developed for the accident rates. This paper will review some recent work in the area, describe the data and the statistical models employed, and present the results of accident rate modelling. It

concludes with a discussion of the results and recommendations for future work.

LITERATURE REVIEW

Much recent research has focused on determining differences between the accident rates of singles and doubles. The Transportation Research Board's 1986 report on doubles evaluated existing research on the safety impacts of the doubles configuration (TRB 1986). Based on the available studies, the report concluded that there was no evidence that one configuration was significantly less safe than the other. The report also stressed that the data available to the studies it relied on covered only some operations of singles and doubles. For example, Chirachavala and O'Day in 1981 used OMC (then BMCS) accident data and travel information from the Bureau of the Census's Truck Inventory and Use Survey (TIUS). In order to match coverage of the two files, the analysis focused on intercity operations by ICC-authorized carriers. The authors found little overall difference in accident rates between singles and doubles, though there were differences for some cargo body styles. Glennon's study used a matched-pair analysis of trips by singles and doubles, using data from a major public carrier. (See literature review in TRB 1986, pp. 310–311). He found that accident rates for the two configurations were virtually identical. The study has been faulted because the data came from private sources, but it was used as part of litigation and withstood the intense scrutiny a trial involves.

In 1988, Stein and Jones employed a case-control technique that has been used more often in clinical studies (Stein and Jones 1988). They used accidents on Interstate highways in Washington, sampling the control group from traffic passing accident sites a week later. Odds ratios calculated from the data indicated that doubles were overinvolved by a factor of two to three times compared with singles. The study also found that bobtails (tractors operating without a trailer) were underinvolved, as were operations at night. The study has generated considerable interest both in its method and its results. Fritzler (1988), in an unpublished paper, questioned the sample sizes of the control group and the extent to which the conclusions could be generalized beyond the specific accident sites. A Canadian study design for assessing the safety impact of front-axle brakes recommended use of the case-control method, but with important modifications on the application of the method and interpretation of the results (Lee-Gosselin, Richardson, and Taylor 1990).

Subsequent studies have used more conventional methods and have arrived at different conclu-

sions from Stein and Jones (1988). A study in western Canada used fleet data to compare singles with doubles and to determine rates for A-, B-, and C-train doubles (Sparks and Bielka 1987). The study concluded that differences in accident rates for singles and doubles were not significant. Calculated rates also suggested that doubles using A-dollies had lower accident rates than B- or C-trains, though the authors noted that this "flies in the face of general belief." In another study, Jovanis, Chang, and Zebaneh (1989) asserted that no previous study "[had] been able to control for differences in how and where doubles operate," that is, controlling for road type, traffic level, and weather. Their study used randomly selected origin-destination terminal pairs used for national less-than-truckload (LTL) carriage. Both singles and doubles operated over precisely the same roads at approximately the same time. They found that accident rates for doubles were somewhat lower than singles on every road type and that the differences, though small in some cases, were statistically significant.

A recent study of fatal truck accident rates used data covering all truck types and roadways (Campbell, Blower, Gattis, and Wolfe 1988). The accident data consisted of a census file of all medium- and heavy-duty trucks involved in fatal accidents in the lower 48 states. The travel data came from a probability-based sample of all trucks registered in the 48 states, and consisted of a survey of the actual usage of the vehicles. The data in both files were collected at a level of detail that permitted calculation of rates by configuration, road type, area type, time of day, and many other factors of interest. One of the conclusions of the study was that the doubles fatal-accident rate is about 10% lower than that of singles, primarily because doubles use interstate-quality roads for a disproportionate share of their travel. If the distribution of doubles travel by road, time, and area type were the same as that of all trucks, the doubles fatal-accident rate would be about 11% higher than singles. The bobtail configuration was significantly overinvolved in all conditions, regardless of whether its travel distribution was adjusted.

The present work employs the same approach, though the focus is restricted to the state of Michigan and the accident threshold is expanded to include all reportable accidents (Blower, Lyles, Campbell, and Stamatiadis 1990). Data used in the study cover the 12-month period from May 1987 through April 1988. Accident counts were developed from a census file of all police-reported accidents in Michigan, classified by configuration, time of day, road type, and area type. The exposure data, collected by telephone interview on a series of randomly selected days, come from a survey of the travel of truck-tractors in Mich-

igan. The mileage data are also classified by the same factors as the accident data, permitting disaggregate rates to be calculated and multivariate models developed, incorporating configuration, road type, time of day, and area type.

DATA

The rates presented here are for Michigan-registered truck-tractors operating on Michigan roads. The mileage figures came from a survey of the actual travel of Michigan-registered tractors in Michigan. Extending the analysis to all truck-tractors on Michigan roads would have involved sampling the entire population of trucks operating in Michigan, regardless of state of registration, which was not feasible. Instead, it was decided to focus on Michigan-registered truck-tractors. Thus, the accident data are limited to accidents in Michigan involving a Michigan-registered tractor, and the travel is limited to travel in Michigan by Michigan-registered tractors.

The data modelled consist of contingency tables of accident rates, determined by the factors of interest. The contingency tables are formed by cross-classifying vehicle configuration, road type, area type, and time of day. Three tractor configurations were considered: a tractor with no trailer (bobtail), a tractor with a semitrailer (single), and a tractor with two trailers (double). Road type was classified as either limited access, major artery (other U.S.- and state-numbered routes), and other roads. Time of day was split between day and night, with night defined as the period between 9 P.M. and 6 A.M. FHWA definitions of urban and rural were used to classify area type. Urban areas are those with a population greater than 4,999.

Accident data

All truck-tractor accidents in Michigan were drawn from Michigan State Police accident files for the period May 1987 through April 1988.* State of registration is not coded in the computerized file, so hard copies of all truck-involved accidents were reviewed to identify the registration state. The review also corrected some errors of truck classification. For the study year, 6,002 Michigan-registered truck-tractors were involved in accidents in Michigan. Overall, there were 10,019 tractors, regardless of state of registration, in the accident file, so about 60% of the tractors involved in accidents in Michigan were registered in Michigan.

*Accident data were prepared by Richard Lyles and Polichronis Stamatiadis of Michigan State University as part of a joint MSU-UMTRI project sponsored by the Michigan Office of Highway Safety Planning.

Michigan uses the K-ABC system of classifying injury severity. The threshold for reporting property-damage-only accidents is \$200 in damage. In the analysis, models are presented for casualty and property-damage-only (PDO) accidents. Casualty accidents were defined to consist of accidents that included a K (fatal), A (incapacitating injury), or B (evident injury but not incapacitating) injury. C-injury, complaint of pain, accidents were included with property-damage-only for the analysis.

Travel data

The measure of exposure used is vehicle miles traveled (VMT). The source of the exposure data was a survey conducted by the University of Michigan Transportation Research Institute (UMTRI) between May 1987 and April 1988. A random sample of 1,055 tractors was drawn from the registration files maintained by the Michigan Department of State. The operator of each truck was contacted by telephone four times over the course of the study year for a detailed description of the activities of the truck on the previous day. Information gathered about the truck's use included the total travel for the day, the number and type of trailers pulled, the actual route driven by the truck, and the time of operation.* The route was plotted on a map and the accumulated travel was recorded for each combination of road type (three levels), time of day (day or night), and area type (urban or rural). Since the accident data included only accidents in Michigan, only travel in Michigan was recorded.

Both the vehicle description and route data were subject to careful editing. Each truck Vehicle Identification Number (VIN) was decoded by UMTRI editors to identify the particular model so that the physical description of the vehicle obtained from the owner could be compared with the manufacturer's specifications. The editing staff has extensive experience with and information concerning heavy trucks from the nationwide Trucks Involved in Fatal Accidents survey, currently in its tenth year (Blower, Pettis, and Sullivan 1990). Thus, problems with vehicle descriptions could be spotted and clarifying calls to the truck owner made. Similarly, following the route of the truck on maps allowed gaps in the truck's usage to be identified, so that additional calls could be made to fill in the missing trips.

Overall, interviews for 8,464 trips on 3,603 sample days were completed. The truck-tractors in the study traveled 470,017 miles on those days; the routes for 96.1% of those miles were described in sufficient

*Additional information, not used in this study, included weights and lengths of all units, cargo body type, cargo weight at every point along the route, driver age, and company type.

detail to be broken down by road type, time of day, and area type.

At the vehicle level, the 1,055 trucks in the sample cover the Michigan tractor population well. There were just under 35,000 truck-tractors registered in Michigan during the study year, so vehicle sample weights are relatively low. For the travel data, weights were calculated to inflate the daily mileage to annual estimates. Specific survey dates were randomly assigned to each vehicle for the travel survey. When the vehicle was not used on its sample day, the last previous use was taken and the weights adjusted appropriately.* This technique allowed the travel estimates to be based on the largest amount of travel possible, since every completed trip interview produced mileage data. A stratified sampling procedure was used when the population of registered trucks was sampled. Design effects calculations showed that this procedure had virtually no effect on the variances.

Strengths of the data

Many earlier studies have relied on data that are more narrowly focused and less representative. The Stein and Jones (1988) case-control work sampled accident sites located on interstates. Other work has used BMCS accident data, which suffer both from underreporting and from only partial coverage of the trucking industry. The TIUS file has frequently been used for travel data, but the mileage there consists of self-reported annual totals, which cannot be verified or cross-classified by many factors of interest. Other studies have been restricted either to particular fleets or particular roadways.

The data used in this study represent a significant

*The trip year was divided into four quarters and, optimally, travel data would have been collected on each truck in every quarter. The "survey day weight" to inflate the daily travel to an annual total is calculated by dividing 365 by the number of completed quarters. If a truck was not used on its selected survey date, the interview was done on the last previous use of the vehicle. Then the problem is to determine the weight of such survey days within the quarter. The survey method did not directly determine the interval between trucks' uses, but only the interval between the survey day and the last day of use. However, once a random date was selected, the particular interval between uses of the sample vehicle that spanned the survey day was also specified. Suppose the sample vehicle were used on the 10th and the 20th, so that this particular interval, I , was ten days. Survey dates from the 10th to the 19th would fall in this interval (the 20th begins the next interval). The "half-interval" is estimated by the difference between the survey date and the date the truck was last used. The possible outcomes for the half-interval for this example are the integers 0, 1, 2, 3, . . . , 9. Since the survey date was randomly selected, the probability of each outcome is the same, and the interval, I , is estimated by $I = 2(\text{half-interval}) + 1$. The frequency of use for each quarter is $1/I$. The resulting factor was multiplied by the survey day weight to determine the "annual inflation factor," the weight used (in combination with vehicle weights and an adjustment for nonresponse) to determine annual estimates of travel from the daily trips.

advance over most previous studies. Although the accident and travel data come from separate sources, they are carefully matched in terms of the vehicle population and time period covered by the study. The numerator in the rates comes from a census file of accidents in Michigan involving Michigan-registered truck-tractors. The denominator comes from the travel of a random sample of all Michigan tractors, and is based on the actual travel of those vehicles. Finally, the data are at a level of detail that permits comparison of the relative accident risk of different road types, times of operation, areas of operation, and vehicle configurations, or any combination.

STATISTICAL METHODS

The fundamental hypothesis tested is whether the accident rate is related to vehicle configuration, road type, area type, and time of day. Since this is an observational study, it is not feasible to measure or control for all variables significantly related to accident involvement. Due to the complex nature of the events leading up to an accident, it is expected that there will be some extraneous variation not accounted for by the covariates. Nevertheless, this should not detract from the finding of this study, that an association between accident rates and the covariates of interest, does, in fact, exist.

The data set consists of a $3 \times 3 \times 2 \times 2$ cross-classification. The cross-classifying factors are defined as truck type (single, double, bobtail), road (limited, major, other), time (day, night), and area (rural, urban). The response, or outcome, variable is the accident rate, which is defined as the number of accidents divided by the amount of travel (in millions of miles) for a given classification. Each rate is therefore based on a different amount of travel. The saturated model contains 36 estimable parameters.

Model characteristics

The Poisson distribution is the nominal distribution for counted data. It is commonly used in modelling accident (Joshua and Garber 1990; Buyco and Saccomanno 1988; Saccomanno and Buyco 1988), and worked well in developing the models presented in this paper. The number of accidents for a given observation is assumed to be Poisson distributed with mean μ_i , and observations are assumed to be independent. Since the Poisson mean is required to be positive, an additive model is usually unsatisfactory. The multiplicative model ensures positive fitted values and is plausible for interpreting factor effects as relative risks. This is the motivation for the log-linear model. The model initially takes the form:

$$\begin{aligned} \log(\text{expected number of accidents}) = & \\ & \alpha + \log(\text{travel}) \\ & + (\text{effect due to truck type}) \\ & + (\text{effect due to road type}) \\ & + (\text{effect due to area type}) \\ & + (\text{effect due to time of day}). \end{aligned}$$

It is assumed that the coefficient attached to $\log(\text{travel})$ is unity, although this assumption must be checked. A parameter estimate that is assumed known a priori is sometimes called an offset. After this assumption is checked, and if it is satisfied, real interest focuses on the factor effects. One of the concerns of modelling accident rates is the incorporation of exposure ($\log(\text{travel})$) into the model. By declaring $\log(\text{travel})$ as an offset, it is initially subtracted from the left hand side of the model equation, producing the log of the accident rate, and the model becomes:

$$\begin{aligned} \log(\text{expected accident rate}) = & \\ & \alpha + (\text{effect due to truck type}) \\ & + (\text{effect due to road type}) \\ & + (\text{effect due to area type}) \\ & + (\text{effect due to time of day}). \end{aligned}$$

Of course, important interaction terms are also included in the model. In generalized linear model terminology (McCullagh and Nelder 1989), the natural link function, in this case the log link, provides for a multiplicative model that describes the relationship between the response and predictor variables. The statistical computer software package GLIM (Generalized Linear Interactive Modelling) fits Poisson regression models as a standard option (Aitkin, Anderson, and Francis 1989).

The parameter estimates obtained are maximum likelihood estimates (MLE). Since the maximum likelihood equations are generally nonlinear in the parameters, the MLEs are solved for iteratively. GLIM uses a modification of the Newton-Raphson algorithm to solve for the MLEs and their standard errors. The identifiability constraints placed on the parameters are such that the baseline or referent category is set to 0.

Goodness-of-fit and model checking

Two measures of goodness-of-fit are commonly used. One is the chi-square statistic, and the other is the deviance. Both are approximately distributed as chi-square with degrees of freedom equal to the number of estimable parameters in the saturated model, minus the number of parameters fit in the model. The deviance also may be used to compare two nested models by taking the difference between their devi-

ances and degrees of freedom. This statistic, sometimes called the likelihood ratio statistic, is also asymptotically chi-square and can be used to test whether the addition of further predictor variables significantly improves the fit. Individual residuals also can be examined to determine how well each rate is being predicted. Analogous to the overall measures of fit provided by the chi-square statistic and the deviance are the chi-square residuals and the deviance residuals. The sum of the squared chi-square residuals equals the chi-square statistic. Similarly, the deviance equals the sum of the squared deviance residuals. Patterns in the residuals, or large residuals, are not necessarily reflected in the overall measure.

When the observed numbers of accidents are fairly large, the residuals can be treated as approximately standard normal variables. A normal quantile plot provides a check on their assumed distribution.

RESULTS

Three models were developed from the Michigan accident and travel data. Two models were developed for casualty accidents and one for property damage accidents. One of the casualty accident models used all three levels of vehicle configuration—bobtail, single, and double. The second casualty accident model excluded bobtails. Dropping bobtails allowed the development of a model with fewer interaction terms and a better fit. The property damage accident model also used only two levels of vehicle configuration, single and double.

Casualty accidents. The 24-cell model

The simpler casualty model will be presented first. The data for this model are displayed in Table 1. There were 24 rates or "cells" in the data matrix. Each cell is determined by the cross-classification of the independent variables in the model. Road type was divided into three types—limited access, major artery, and "other" roads. Area was classified as either urban or rural, and time was divided into day and night. Tractor configuration was either single or double. Excluding bobtails from the model does not neglect a major tractor configuration. In Michigan, singles and doubles account for 98.8% of tractor travel. Bobtails account for the remaining 1.2% of tractor miles.

Table 1 also shows half 95% confidence intervals for each observed rate in the 24-cell rate matrix. Confidence intervals were calculated based on the estimated variances for the travel and accident numbers. Cells with relatively little travel or few accidents have large confidence intervals. Since the observations fit by the models are accident rates, the large confidence

Table 1. Accidents, travel, and accident rates for the 24-cell casualty accident model

| Cell | Truck type | Road type | Time | Area | Casualty accidents | Travel | Rate/ 10 ⁶ miles | Half 95% C.I. | |
|------|------------|-----------|---------|-------|--------------------|-------------|--------------------------------|------------------|---------|
| 1 | Single | Limited | Day | Rural | 188 | 204,433,874 | 0.9196 | 0.1436 | |
| 2 | | | Urban | 107 | 177,250,749 | 0.6037 | 0.1207 | | |
| 3 | | Major | Limited | Night | Rural | 63 | 41,949,294 | 1.5018 | 0.4130 |
| 4 | | | | Urban | 23 | 29,883,757 | 0.7696 | 0.3311 | |
| 5 | | Major | Major | Day | Rural | 241 | 128,647,023 | 1.8733 | 0.2851 |
| 6 | | | | Urban | 92 | 59,822,202 | 1.5379 | 0.3364 | |
| 7 | | Major | Major | Night | Rural | 61 | 17,642,351 | 3.4576 | 1.0507 |
| 8 | | | | Urban | 19 | 6,838,521 | 2.7784 | 1.3648 | |
| 9 | | Other | Other | Day | Rural | 200 | 31,765,363 | 6.2962 | 1.1633 |
| 10 | | | | Urban | 118 | 59,730,974 | 1.9755 | 0.3917 | |
| 11 | | Other | Other | Night | Rural | 22 | 1,289,058 | 17.0667 | 8.7425 |
| 12 | | | | Urban | 19 | 3,775,431 | 5.0325 | 2.5808 | |
| 13 | Double | Limited | Day | Rural | 21 | 23,163,210 | 0.9066 | 0.4345 | |
| 14 | | | Urban | 13 | 21,162,524 | 0.6143 | 0.3564 | | |
| 15 | | Major | Limited | Night | Rural | 11 | 9,473,358 | 1.1612 | 0.7912 |
| 16 | | | | Urban | 1 | 3,474,259 | 0.2878 | 0.5844 | |
| 17 | | Major | Major | Day | Rural | 31 | 15,040,022 | 2.0612 | 0.9069 |
| 18 | | | | Urban | 11 | 5,529,527 | 1.9893 | 1.2980 | |
| 19 | | Major | Major | Night | Rural | 5 | 2,400,560 | 2.0828 | 2.1343 |
| 20 | | | | Urban | 0 | 459,525 | 0.0000 | 0.0000 | |
| 21 | | Other | Other | Day | Rural | 26 | 3,207,263 | 8.1066 | 4.2706 |
| 22 | | | | Urban | 17 | 4,951,688 | 3.4332 | 1.9036 | |
| 23 | | Other | Other | Night | Rural | 2 | 224,036 | 8.9271 | 13.1945 |
| 24 | | | | Urban | 6 | 340,844 | 17.6034 | 16.2369 | |

intervals in the observed rates are more likely to have large residuals from the model fitting.

Table 2 shows the coefficients and standard errors for the parameters in the model that was developed. The model has a scaled deviance of 18.98 with 14 degrees of freedom, which is consistent with a good overall fit. That is, treating the scaled deviance as a chi-square with 14 degrees of freedom, the overall distribution of accidents predicted by the model does not differ significantly from the observed distribution of accidents. Note also that the model is able to account for most of the variation in the data with only two interaction terms. Twice the standard error gives a quick approximation of a 95% confidence interval for the coefficients. All are significant at the .05 level

Table 2. Parameter estimates for the 24-cell casualty accident rate model

| Coefficient estimate | Standard error | Parameter |
|----------------------|----------------|---------------|
| -13.89 | 0.06672 | Baseline* |
| 0.09039 | 0.08857 | Double |
| 0.7060 | 0.08856 | Major artery* |
| 1.925 | 0.09271 | Other roads* |
| 0.3655 | 0.1154 | Night* |
| -0.4720 | 0.1025 | Urban* |
| 0.1754 | 0.1666 | Major-night |
| 0.6106 | 0.1911 | Other-night* |
| 0.2826 | 0.1473 | Major-urban |
| -0.6347 | 0.1442 | Other-urban* |

Scaled deviance = 18.982, 14 degrees of freedom.

*Statistically significant at the .05 level.

except the main effect for doubles and the interactions of major artery roads with night and urban areas, respectively.

The "baseline" case is a tractor pulling a single trailer on rural limited access roads during the day. The other coefficients reflect the effect of the various factors relative to the baseline case.

Overall, road type has the largest effect on the relative risk of different operating environments in this model. This can be seen most directly by examining the coefficients for the main effects of the parameters presented in Table 2. The largest coefficients are associated with major and, especially, "other" roads. For "other" roads, the accident rate is 6.8 times ($e^{1.925}$) greater than the accident rate on limited access roads. Major arteries have 2.03 times ($e^{0.706}$) the risk of limited access roads. Of course, this only compares the baseline case of single, day, rural, limited access with the same case on "other" and major roads, respectively. Nevertheless, the coefficients are included in calculating the rates for all the "other" and major artery cells, so the main effect of road type is clear. Of the factors included in the model, road type has by far the largest effect on accident rates.

Time of operations and area are also important factors in the model. For area type, the coefficient for urban is -0.472 , indicating that relative to the baseline case the risk of involvement in a casualty accident is $e^{-0.472} = .62$ of the baseline (rural) case. The effect of urban operations is to lower the casualty accident rate, in general. This could be due to lower

travel speeds in urban areas. Night, in contrast, is associated with higher accident risks. The parameter for the main effect of night is 0.3655, which means the relative risk is 1.44 times higher at night.

Two interaction terms improved the fit of the model appreciably: road type with time of day and road type with area type. In the case of the "other" road type, the effect is quite large. The coefficient of the "other" road-night interaction indicates that the combination of "other" roads and night increases the risk beyond the sum of the separate effects of night and "other" roads. Thus, the rate for singles on "other" rural roads at night is higher than limited access rural roads during the day (the baseline case) by $e^{(1.925+0.3655+0.6106)}$ or over 18 times. On the other hand, the interaction effect of "other" roads and urban areas is negative, with approximately the same magnitude as the "other"-night interaction. This indicates that the harmful effect of "other" roads at night is primarily in rural areas, where operating speeds are higher than on comparable urban roads.

In contrast to the finding that road type, area, and time of day have a significant impact on the casualty accident risk, configuration has very little impact. The coefficient for the doubles configuration is 0.09039 or about a 9% higher risk. The standard error of the estimate is 0.08857, so the coefficient is not statistically different from zero. Even if the effect is real,

it is relatively small and far outweighed by the operating environment, particularly road type.

Table 3 shows rates predicted by the model, with upper and lower bounds of a 95% confidence interval, along with the rates actually observed in the data for each of the 24 cells in the casualty accident rate matrix. Overall, the table shows the wide variation in accident rates for different combinations of factors in the model. The range is from predicted rates of 0.6 casualty accidents per million miles of travel for doubles on limited access urban roads in the daytime to over 18 for doubles on rural, night, "other" roads. The factors in the model clearly have a very large effect on the probability of an accident. Most of the confidence intervals for the singles cells are reasonably tight. Singles are the dominant tractor configuration so there is adequate data for most of the cells. Doubles account for about 10% of all tractor travel and sample sizes for some cells are small. For example, there were no casualty accidents recorded for the doubles combination on major arteries at night in urban areas and only 459,000 miles of estimated travel. Nevertheless, for most cells the confidence intervals are satisfactory and demonstrate clear differences in the rates from cell to cell.

The deviance residuals were also examined to see if additional independent variables were required for the model. The residuals are displayed in Table 4.

Table 3. Observed and predicted rates and 95% confidence intervals for the 24-cell casualty accident rate model

| Cell | Truck type | Road type | Time | Area | Observed rate | Predicted rate | Lower bound | Upper bound* |
|------|------------|-----------|-------|-------|---------------|----------------|-------------|--------------|
| 1 | Single | Limited | Day | Rural | 0.9196 | 0.9273 | 0.8136 | 1.0568 |
| 2 | | | | Urban | 0.6037 | 0.5784 | 0.4879 | 0.6857 |
| 3 | | Major | Night | Rural | 1.5018 | 1.3364 | 1.0849 | 1.6462 |
| 4 | | | | Urban | 0.7696 | 0.8336 | 0.6517 | 1.0662 |
| 5 | | Other | Day | Rural | 1.8733 | 1.8785 | 1.6701 | 2.1131 |
| 6 | | | | Urban | 1.5379 | 1.5545 | 1.2957 | 1.8649 |
| 7 | | Night | Rural | Day | 3.4576 | 3.2264 | 2.5921 | 4.0160 |
| 8 | | | | Urban | 2.7784 | 2.6698 | 2.0458 | 3.4840 |
| 9 | | Day | Rural | Other | 6.2962 | 6.3536 | 5.5848 | 7.2283 |
| 10 | | | | Urban | 1.9755 | 2.1008 | 1.7877 | 2.4688 |
| 11 | | Night | Rural | Day | 17.0667 | 16.8617 | 12.5417 | 22.6697 |
| 12 | | | | Urban | 5.0325 | 5.5753 | 4.1328 | 7.5214 |
| 13 | Double | Limited | Day | Rural | 0.9066 | 1.0150 | 0.8307 | 1.2403 |
| 14 | | | | Urban | 0.6143 | 0.6331 | 0.5031 | 0.7967 |
| 15 | | Major | Night | Rural | 1.1612 | 1.4628 | 1.1384 | 1.8797 |
| 16 | | | | Urban | 0.2878 | 0.9124 | 0.6867 | 1.2124 |
| 17 | | Other | Day | Rural | 2.0612 | 2.0563 | 1.6964 | 2.4926 |
| 18 | | | | Urban | 1.9893 | 1.7015 | 1.3373 | 2.1650 |
| 19 | | Night | Rural | Day | 2.0828 | 3.5316 | 2.7071 | 4.6073 |
| 20 | | | | Urban | 0.0000 | 2.9223 | 2.1460 | 3.9794 |
| 21 | | Day | Rural | Other | 8.1066 | 6.9546 | 5.6853 | 8.5074 |
| 22 | | | | Urban | 3.4332 | 2.2996 | 1.8330 | 2.8849 |
| 23 | | Night | Rural | Day | 8.9271 | 18.4568 | 13.2583 | 25.6935 |
| 24 | | | | Urban | 17.6034 | 6.1028 | 4.3614 | 8.5393 |

*The confidence intervals for this table are calculated from the rates predicted by the model. The confidence intervals for the raw rates, e.g. in Table 1, are calculated from the raw accident frequencies and truck travel.

Table 4. Observed and predicted frequencies, and residuals for the 24-cell casualty accident model

| Cell | Truck type | Road type | Time | Area | Observed frequency | Predicted frequency | Deviance residual* | |
|------|------------|-----------|-------|-------|--------------------|---------------------|--------------------|--------|
| 1 | Single | Limited | Day | Rural | 188 | 189.570 | -0.114 | |
| 2 | | | | Urban | 107 | 102.521 | 0.439 | |
| 3 | | | Night | Rural | 63 | 56.061 | 0.909 | |
| 4 | | | | | Urban | 23 | 24.911 | -0.388 |
| 5 | | | Major | Day | Rural | 241 | 241.674 | -0.043 |
| 6 | | Urban | | | 92 | 92.991 | -0.103 | |
| 7 | | Night | | Rural | 61 | 56.922 | 0.534 | |
| 8 | | | | | Urban | 19 | 18.257 | 0.173 |
| 9 | | | Other | Day | Rural | 200 | 201.824 | -0.129 |
| 10 | Urban | 118 | | | 125.484 | -0.675 | | |
| 11 | Night | Rural | | 22 | 21.736 | 0.057 | | |
| 12 | | | | | Urban | 19 | 21.049 | -0.454 |
| 13 | Double | Limited | | Day | Rural | 21 | 23.511 | -0.527 |
| 14 | | | | | Urban | 13 | 13.398 | -0.109 |
| 15 | | | Night | Rural | 11 | 13.858 | -0.797 | |
| 16 | | | | | Urban | 1 | 3.170 | -1.426 |
| 17 | | | Major | Day | Rural | 31 | 30.927 | 0.013 |
| 18 | | Urban | | | 11 | 9.409 | 0.505 | |
| 19 | | Night | | Rural | 5 | 8.478 | -1.294 | |
| 20 | | | | | Urban | 0 | 1.343 | -1.639 |
| 21 | | | Other | Day | Rural | 26 | 22.305 | 0.762 |
| 22 | Urban | 17 | | | 11.387 | 1.549 | | |
| 23 | Night | Rural | | 2 | 4.135 | -1.168 | | |
| 24 | | | | Urban | 6 | 2.080 | 2.207 | |

*The deviance residuals are scaled.

Overall, there does not seem to be any pattern that would indicate systematic lack of fit. Generally, the cells with large residuals also have large confidence intervals in the observed rates. The large residuals are primarily in the doubles cells, where sample sizes are small. Another test for unspecified variables that are correlated with the residuals is to examine the normal plot of the residuals. The plot of the residuals should show a straight line through the origin. The plot of the residuals for this model (not shown) has some upward curve in the tails, but for the most part, the data satisfy the requirements of the GLIM method. Similarly, the method assumes that the offset variable, in this case travel, is linearly related to the dependent variable, such that when the offset is incorporated into the model as an independent variable, the coefficient is one. When that is tested in this model, the travel coefficient is 0.6096 with a standard error of 0.2977. Although the coefficient is somewhat less than one, one is well within the 95% confidence interval.

Casualty accidents. The 36-cell model

Adding bobtails (tractors operating without a trailer) as a third category in the vehicle-type variable increased the number of cells by 12 and resulted in a more complicated model. The final model developed used all the same parameters as the 24-cell model discussed above, but several additional parameters were required to achieve a good fit. The additional parameters were for two sets of interactions involving vehi-

cle type. The first is between vehicle type and road type, the second between vehicle type and area type. Even with the additional parameters, the fit of the model is not quite as good as the previous model. The 36-cell model has a scaled deviance of 30.928 with 19 degrees of freedom. Treating the scaled deviance as a chi-square, this is within the critical value at the .10 level, and just over the critical value at the .05 confidence level. Thus, overall, the model does not quite achieve significance at the 95% level.

Table 5 shows the additional bobtail data for the 36-cell model. The data for the singles and doubles remain the same as in Table 1. Since many of the results and most of the data are the same for both the 24-cell and the 36-cell casualty model, only the additional data and results associated with bobtails will be presented.

The model derived using the GLIM package had 17 parameters (see Table 6). All the parameters included in the 24-cell model were used, and in addition there were terms for the interactions between vehicle type and road type, and between vehicle type and area type. The effect of adding observations for bobtails is to add variance that requires additional terms in the model. The additional terms are all interactions between vehicle type and other independent variables. It seems that adding another level to vehicle type complicates the relationship among vehicle, road, and area because bobtails interact differently from singles or doubles with those two factors.

Table 5. Accidents, travel, and accident rates for bobtail cells of the 36-cell casualty accident model

| Cell | Truck type | Road type | Time | Area | Casualty accidents | Travel | Rate/ 10 ⁶ miles | Half 95% C.I. |
|------|------------|-----------|-------|-------|--------------------|-----------|--------------------------------|------------------|
| 25 | Bobtail | Limited | Day | Rural | 7 | 2,096,174 | 3.3394 | 2.8188 |
| 26 | | | | Urban | 9 | 2,627,142 | 3.4258 | 2.4602 |
| 27 | | | Night | Rural | 2 | 236,833 | 8.4448 | 19.3254 |
| 28 | | | | Urban | 0 | 371,614 | 0.0000 | 0.0000 |
| 29 | | Major | Day | Rural | 12 | 2,099,071 | 5.7168 | 4.0829 |
| 30 | | | | Urban | 4 | 929,995 | 4.3011 | 4.4500 |
| 31 | | | Night | Rural | 0 | 67,400 | 0.0000 | 0.0000 |
| 32 | | | | Urban | 0 | 68,364 | 0.0000 | 0.0000 |
| 33 | | Other | Day | Rural | 22 | 258,223 | 85.1977 | 52.5523 |
| 34 | | | | Urban | 7 | 1,439,441 | 4.8630 | 3.7881 |
| 35 | | | Night | Rural | 7 | 58,385 | 119.8938 | 183.7319 |
| 36 | | | | Urban | 4 | 92,941 | 43.0381 | 48.0053 |

As in the 24-cell casualty model, the baseline case is a tractor pulling a single trailer on rural, limited access roads during the day. The other coefficients measure the effect of their respective factors relative to the baseline case. For most parameters that appear in both the 24-cell and 36-cell models, the coefficients and standard errors are very similar.

The largest standard errors for the terms in the model are associated with bobtails. This is a result of the relatively small amount of data for the bobtails. Three of the bobtail cells have zero casualty accidents. The bobtail configuration also has less travel than doubles or singles. Since the rates are based on less data, the variance is greater.

More of the terms in the 36-cell model were not statistically significant than in the 24-cell model. In some cases, the lack of statistical significance is due to the small size of the effect. For example, the double-major artery roads coefficient is 0.1324 with a stan-

dard error of 0.2204. Even if the standard error were much smaller, the effect of the interaction would still be slight. In other cases, the failure of some of the coefficients to attain significance is associated with the relative lack of data concerning bobtails, rather than the size of the effect. For example, in these data the interaction between bobtails and major artery roads is -0.5508 , which is substantial. The standard error of the coefficient is 0.3531, which means, using twice the standard error to approximate a 95% confidence interval, the coefficient is not statistically different from zero. The interactions are in the model because they improve the fit substantially, though not every term is significant.

Including bobtails in the vehicle type variable allows the rates for the three major tractor configurations to be compared directly. The coefficient estimates in Table 6 show that the risk of the bobtail configuration is quite large, with a coefficient second only to the coefficient for "other" roads in magnitude, and, despite the small amount of bobtail data, highly significant. Compared with the baseline case, the rate for bobtails on limited access, rural roads during the day is $e^{1.701}$ or 5.5 times higher. The rates for bobtails are higher than the singles or doubles configurations in every cell in the data matrix. The importance of this model is to show that the bobtail is the most unsafe of the tractor configurations. The coefficient for the doubles configuration is small and not statistically significant.

Table 7 shows the accident rates per million miles predicted by the model for the bobtail cells in the data matrix, along with the lower and upper bound of the 95% confidence interval and the rate observed in the data. As expected, confidence intervals (C.I.) for bobtail cells are relatively wide. For example, the rate for the baseline case is 0.94, with a C.I. from 0.83 to 1.08. The bobtail rate for the same cell is 5.2, 3.2–8.4 C.I. The bobtail rates are based on less travel and fewer accidents because bobtails are used

Table 6. Parameter estimates for the 36-cell casualty accident rate model

| Coefficient estimate | Standard error | Parameter |
|----------------------|----------------|----------------|
| -13.87 | 0.06758 | baseline* |
| -0.1726 | 0.1708 | double |
| 1.701 | 0.2542 | bobtail* |
| 0.7024 | 0.0908 | major artery* |
| 1.907 | 0.09434 | other roads* |
| 0.3732 | 0.1141 | night* |
| -0.4631 | 0.1026 | urban* |
| 0.1324 | 0.2204 | double-major |
| 0.4494 | 0.2166 | double-other* |
| -0.5508 | 0.3531 | bobtail-major |
| 0.5211 | 0.2955 | bobtail-other |
| 0.1561 | 0.1655 | major-night |
| 0.5926 | 0.1811 | other-night* |
| 0.2451 | 0.1890 | double-urban |
| -0.6448 | 0.2642 | bobtail-urban* |
| 0.2628 | 0.1443 | major-urban |
| -0.7167 | 0.1401 | other-urban* |

Scaled deviance = 30.928, 19 degrees of freedom.

*Statistically significant at the .05 level.

Table 7. Observed and predicted rates and 95% confidence intervals for the bobtail cells of the 36-cell accident rate model

| Cell | Truck type | Road type | Time | Area | Observed rate | Predicted rate | Lower bound | Upper bound |
|------|------------|-----------|-------|-------|---------------|----------------|-------------|-------------|
| 25 | Bobtail | Limited | Day | Rural | 3.3394 | 5.1635 | 3.1627 | 8.4301 |
| 26 | | | | Urban | 3.4258 | 1.7053 | 0.9408 | 3.0910 |
| 27 | | | Night | Rural | 8.4448 | 7.4994 | 4.4411 | 12.6639 |
| 28 | | | | Urban | 0.0000 | 2.4767 | 1.3256 | 4.6275 |
| 29 | | Major | Day | Rural | 5.7168 | 6.0087 | 3.6491 | 9.8940 |
| 30 | | | | Urban | 4.3011 | 2.5808 | 1.3255 | 5.0252 |
| 31 | | | Night | Rural | 0.0000 | 10.2017 | 5.9163 | 17.5914 |
| 32 | | | | Urban | 0.0000 | 4.3818 | 2.1697 | 8.8494 |
| 33 | | Other | Day | Rural | 85.1977 | 58.5525 | 40.0221 | 85.6626 |
| 34 | | | | Urban | 4.8630 | 9.4440 | 6.0956 | 14.6319 |
| 35 | | | Night | Rural | 119.8938 | 153.8151 | 101.8391 | 232.3185 |
| 36 | | | | Urban | 43.0381 | 24.8091 | 15.2517 | 40.3555 |

much less than the other tractor configurations. Bobtails had three cells with no accidents. Nevertheless, the bobtail rates are still clearly much higher than singles or doubles.

The data were checked to see if they satisfy the assumptions of the model. A check of the residuals (not shown) show that the larger residuals are concentrated in the bobtail and double cells. In general they appear to be related to lack of data rather than to any systematic pattern. The normal plot of the residuals also showed them to be normally distributed. When travel is introduced into the model as an independent variable, the coefficient is 0.771 with a standard error of 0.2197, which is well within bounds.

Property damage accidents. The 24-cell model

Initial models included a variable for accident severity, broken down as casualty and property-damage-only (PDO). The models developed with accident severity as an independent variable required many 3- and 4-way interaction terms, which use up most of the degrees of freedom and are very difficult to interpret. The large number of interaction terms involving accident severity made clear that there were essential differences in the relationship of the other factors to accident risk for casualty as compared with PDO accidents. It was decided to separate casualty accidents from property damage accidents and to model each individually.

Table 8 shows the property damage accident data that were modelled. The travel numbers are, of course, the same as in the previous two models. Note that only singles and doubles are considered. Confidence intervals for the PDO rates are somewhat tighter due to the larger sample sizes in the numerator. For example, the half-width of the confidence interval for the baseline cell is about 15% of the casualty rate, as compared with 10% of the PDO rate.

Table 9 shows the coefficients and standard errors for the parameters. The model has a scaled deviance of 11.183 with 8 degrees of freedom, which is a very good fit. Unfortunately, achieving that fit required five two-way interactions and one three-way interaction. This model is more complicated than the previous ones. Its interpretation is more complicated as well.

As in the previous models, the main effect of doubles is negligible and not statistically significant. However, note that the vehicle type variable appears in interaction terms with road type and with time of day. For both the double-“other” and the double-night terms, the coefficient is negative, which means that the interaction serves to lower the relative risk compared to singles for the same conditions. For example, the rate per million miles for doubles on “other” roads at night in rural areas is $e^{(-12.78 + .04898 + 2.134 + .1113 - .4497 - .5114 + .6593)} \times 10^6 = e^{(-10.78752)} \times 10^6 = 20.6786$. For singles in the same environment, the rate is 61.4851. (See Table 10.) What is the explanation for this result? One possibility is that doubles pulling van trailers, which account for about half the travel of doubles, more often operate at night along regular routes. Such operations have the advantage of permitting a schedule that allows the driver to adjust to night work, as well as the familiarity of a regular route.

Just as in the casualty models, the largest coefficients are associated with road type. In fact, the effect of “other” roads is larger for property damage accidents than for casualty accidents, although among PDOs the effect is greater for singles than doubles. If we just vary road type to “other” roads for the baseline case, the relative risk increases by almost 8.5 times ($e^{2.134}$). For doubles, the increased risk is 5.4 times ($e^{2.134 - .4497}$). Leaving aside these differences by configuration, the overriding point is that road class

Table 8. Accidents, travel, and accident rates for the 24-cell property-damage-only accident model

| Cell | Truck type | Road type | Time | Area | PDO accidents* | Travel | Rate/10 ⁶ miles | Half 95% C.I. | | |
|------|------------|-----------|-------|-------|----------------|-------------|----------------------------|---------------|---------|---------|
| 1 | Single | Limited | Day | Rural | 580 | 204,433,874 | 2.8371 | 0.2839 | | |
| 2 | | | | Urban | 348 | 177,250,749 | 1.9633 | 0.2329 | | |
| 3 | | | Night | Rural | 137 | 41,949,294 | 3.2658 | 0.6639 | | |
| 4 | | Major | Day | Rural | Urban | 40 | 29,883,757 | 1.3385 | 0.4462 | |
| 5 | | | | | Rural | 730 | 128,647,023 | 5.6744 | 0.6229 | |
| 6 | | | | Urban | 353 | 59,822,202 | 5.9008 | 0.7396 | | |
| 7 | | | Night | Rural | 121 | 17,642,351 | 6.8585 | 1.6775 | | |
| 8 | | | Other | Day | Rural | Urban | 45 | 6,838,321 | 6.5804 | 2.2763 |
| 9 | | | | | | Rural | 748 | 31,765,363 | 23.5477 | 3.2871 |
| 10 | | Urban | | | 808 | 59,730,974 | 13.5273 | 1.3774 | | |
| 11 | | Night | | Rural | 67 | 1,289,058 | 51.9759 | 19.4676 | | |
| 12 | | Urban | | 49 | 3,775,431 | 12.9787 | 4.7527 | | | |
| 13 | Double | Limited | | Day | Rural | 65 | 23,163,210 | 2.8062 | 0.8907 | |
| 14 | | | Urban | | 53 | 21,162,524 | 2.5044 | 0.8089 | | |
| 15 | | | Night | Rural | 14 | 9,473,358 | 1.4778 | 0.9186 | | |
| 16 | | Major | Day | Rural | Urban | 3 | 3,474,259 | 0.8635 | 1.0417 | |
| 17 | | | | | Rural | 81 | 15,040,022 | 5.3856 | 1.8179 | |
| 18 | | | | Urban | 30 | 5,529,527 | 5.4254 | 2.3984 | | |
| 19 | | | Night | Rural | 12 | 2,400,560 | 4.9988 | 3.8179 | | |
| 20 | | | Other | Day | Rural | Urban | 5 | 459,525 | 10.8808 | 10.7835 |
| 21 | | | | | | Rural | 60 | 3,207,263 | 18.7075 | 8.1617 |
| 22 | | Urban | | | 36 | 4,951,688 | 7.2702 | 3.1123 | | |
| 23 | | Night | | Rural | 4 | 224,036 | 17.8543 | 19.4321 | | |
| 24 | | Urban | | 2 | 340,844 | 5.8678 | 8.6718 | | | |

*Accidents that involved only C injuries, complaint of pain, are included with property-damage-only accidents.

has a very large influence on the probability of a property damage accident, just as it does with more serious accidents.

The main effect for urban is $-.3383$, indicating that the general effect of urban in all relevant cells is to lower the accident risk compared with rural areas. This is similar to the finding in the casualty accident models, though the urban effect for PDOs is somewhat less, which could be due to an association be-

tween property damage accidents and traffic density. Note also that the coefficient for the night-urban term is strongly negative. Urban traffic is typically slower than rural, and urban areas are also more likely to be lighted at night. Moreover, traffic densities are lower at night.

Night, in contrast to the effect for casualty accidents, is not associated with a large increase in the PDO rate. The main effect of night is 0.1113, which means that when day is replaced with night for the baseline case, the risk is only about 12% higher. Moreover, the coefficient is no longer significant. The effect of night is primarily in interaction with other factors in the model. The coefficient for the night-“other” road type interaction is 0.6593, indicating that this combination increases the relative risk beyond the separate effects of night and the “other” road type. On the other hand, the coefficient for night-urban is almost as strongly negative (-0.5163). This is consistent with the interpretation that the night-“other” road problem is primarily in rural areas, where the lack of lighting compounds the problems associated with “other” roads. Urban roads are more commonly lighted at night.

Table 10 shows the property damage accident rates predicted by the model, the lower and upper bounds of the 95% confidence intervals for those rates, and the rate actually observed in the data. Table 11 presents the deviance residuals from the model.

Table 9. Parameter estimates for the 24-cell property-damage-only accident rate model

| Coefficient estimate | Standard error | Parameter |
|----------------------|----------------|-------------------|
| -12.78 | 0.04066 | baseline* |
| 0.04898 | 0.09551 | double |
| 0.6952 | 0.05461 | major artery* |
| 2.134 | 0.05412 | other roads* |
| 0.1113 | 0.09295 | night |
| -0.3383 | 0.06360 | urban* |
| -0.05986 | 0.1307 | double-major |
| -0.4497 | 0.1379 | double-other* |
| --0.5114 | 0.1772 | double-night* |
| 0.09939 | 0.1304 | major-night |
| 0.6593 | 0.1536 | other-night* |
| 0.3757 | 0.08884 | major-urban* |
| -0.2382 | 0.08042 | other-urban* |
| -0.5163 | 0.1844 | night-urban* |
| 0.4868 | 0.2555 | major-night-urban |
| -0.2808 | 0.2647 | other-night-urban |

Scaled deviance = 11.183, 8 degrees of freedom.

*Statistically significant at the .05 level.

Table 10. Observed and predicted rates and 95% confidence intervals for the 24-cell property-damage-only accident rate model

| Cell | Truck type | Road type | Time | Area | Observed rate | Predicted rate | Lower bound | Upper bound | |
|------|------------|-----------|---------|--------|---------------|----------------|-------------|-------------|---------|
| 1 | Single | Limited | Day | Rural | 2.8371 | 2.8196 | 2.6036 | 3.053 | |
| 2 | | | | Urban | 1.9633 | 2.0103 | 1.8188 | 2.222 | |
| 3 | | | Rural | 3.2658 | 3.1514 | 2.6718 | 3.717 | | |
| 4 | | Major | Day | Urban | 1.3385 | 1.3407 | 0.9935 | 1.809 | |
| 5 | | | | | Rural | 5.6744 | 5.6506 | 5.261 | 6.069 |
| 6 | | | | Urban | 5.9008 | 5.866 | 5.3006 | 6.492 | |
| 7 | | | Night | Rural | 6.8585 | 6.9757 | 5.8728 | 8.286 | |
| 8 | | | | | Urban | 6.5804 | 7.0313 | 5.3279 | 9.279 |
| 9 | | | | Other | Day | Rural | 23.5477 | 23.8252 | 22.2118 |
| 10 | | Urban | 13.5273 | | | | 13.3867 | 12.503 | 14.333 |
| 11 | | Night | Rural | | | 51.9759 | 51.4851 | 40.7464 | 65.054 |
| 12 | | | | | Urban | 12.9787 | 13.0357 | 9.9039 | 17.158 |
| 13 | Double | | Limited | | Day | Rural | 2.8062 | 2.9611 | 2.4629 |
| 14 | | Urban | | | | 2.5044 | 2.1112 | 1.7406 | 2.561 |
| 15 | | Major | | Night | Rural | 1.4778 | 1.9846 | 1.4136 | 2.786 |
| 16 | | | Urban | | | 0.8635 | 0.8443 | 0.5454 | 1.307 |
| 17 | | | Day | | Rural | 5.3856 | 5.5895 | 4.6704 | 6.689 |
| 18 | | | | Urban | | 5.4254 | 5.8025 | 4.7651 | 7.066 |
| 19 | | | | Other | Night | Rural | 4.9988 | 4.1376 | 2.8395 |
| 20 | | | Urban | | | | 10.8808 | 4.1706 | 2.6706 |
| 21 | | Day | Rural | | | 18.7075 | 15.9585 | 13.0574 | 19.504 |
| 22 | | | | | Urban | 7.2702 | 8.9666 | 7.3253 | 10.976 |
| 23 | | | Night | | Rural | 17.8543 | 20.6786 | 13.5052 | 31.662 |
| 24 | | Urban | | | | 5.8678 | 5.2357 | 3.3069 | 8.289 |

Overall, they look very good, except for some of the doubles cells. The largest residual is for the double-major-night-urban cell, where the model predicted 1.9 accidents, but five actually occurred. In such a small cell, a very few accidents either way can make a large difference. Overall, the model is quite successful in fitting the data, particularly for the singles cells.

DISCUSSION

Recently there has been a growing controversy over the relative safety of singles and doubles. A series of studies has found either no differences or that doubles present a significantly greater risk of accident involvement than singles. Most previous work has been limited to a particular road type or to a restricted range of operations. The data used in this analysis broadly cover the uses of truck-tractors in Michigan. Though there are regional variations in the typical usage of some tractor configurations, the Michigan experience should not differ greatly from the national experience. Almost 70% of the Michigan trucks are operated by interstate carriers. Almost 50% of the travel of Michigan-registered tractors is accumulated out of state. The Michigan truck population differs from the rest of the nation primarily because of Michigan's generous weight laws—combination vehicles can be licensed to operate at gross weights up to 164,000 pounds. Nevertheless, for the level of analysis presented here, the Michigan truck population ad-

equately represents trucking nationally. To the extent that the Michigan truck population differs from that elsewhere, it is that Michigan allows combinations up to 164,000 pounds. This fact would hardly be expected to result in a safer tractor population or to reduce differences between configurations. Moreover, none of the factors included in the models are directly related to gross weight.

The data presented here show that the differences in relative risk between singles and doubles are small compared to other factors. In the 24-cell casualty accident model, the coefficient for the main effect of doubles is 0.09, which implies about a 10% higher risk for doubles. That difference is not statistically significant, since the standard error of the estimate is about equal to the estimate. In property damage accidents, the size of the effect is even smaller, with an increased risk of about 5%. In each model, the risk for doubles is not significantly different from singles. If the doubles configuration does pose a higher overall risk of accident involvement, the increased risk is about 10%. An interpretation of no overall difference in accident risk is also consistent with the models.

There are some differences between singles and doubles when road type is taken into account. It has been argued that the similarity in the overall rates of singles and doubles can be explained by the fact that doubles travel more than singles on limited access roads, which are significantly safer than other roads in the highway system (Carsten 1987; Campbell et al.

Table 11. Observed and predicted frequencies and residuals for the 24-cell property-damage-only accident model

| Cell | Truck type | Road type | Time | Area | Observed frequency | Predicted frequency | Deviance residual | |
|------|------------|-----------|-------|-------|--------------------|---------------------|-------------------|--------|
| 1 | Single | Limited | Day | Rural | 580 | 576.412 | 0.149 | |
| 2 | | | Urban | 348 | 356.322 | -0.443 | | |
| 3 | | Major | Day | Night | Rural | 137 | 132.2 | 0.415 |
| 4 | | | | Urban | 40 | 40.067 | -0.011 | |
| 5 | | Major | Day | Rural | 730 | 726.934 | 0.114 | |
| 6 | | | | Urban | 353 | 350.915 | 0.111 | |
| 7 | | Major | Day | Night | Rural | 121 | 123.067 | -0.187 |
| 8 | | | | Urban | 45 | 48.084 | -0.450 | |
| 9 | | Double | Other | Day | Rural | 748 | 756.817 | -0.321 |
| 10 | | | | Urban | 808 | 799.6 | 0.297 | |
| 11 | | | Other | Day | Night | Rural | 67 | 66.367 |
| 12 | Urban | | | | 49 | 49.215 | -0.031 | |
| 13 | Limited | | Day | Rural | 65 | 68.588 | -0.437 | |
| 14 | | | | Urban | 53 | 44.678 | 1.209 | |
| 15 | Major | | Day | Night | Rural | 14 | 18.8 | -1.160 |
| 16 | | | | Urban | 3 | 2.933 | 0.039 | |
| 17 | Major | | Day | Rural | 81 | 84.066 | -0.336 | |
| 18 | | | | Urban | 30 | 32.085 | -0.372 | |
| 19 | Major | | Day | Night | Rural | 12 | 9.933 | 0.635 |
| 20 | | Urban | | 5 | 1.916 | 1.850 | | |
| 21 | Other | Day | Rural | 60 | 51.183 | 1.199 | | |
| 22 | | | Urban | 36 | 44.4 | -1.304 | | |
| 23 | | | Rural | 4 | 4.633 | -0.301 | | |
| 24 | Other | Day | Night | Urban | 2 | 1.785 | 0.158 | |

1988). The greater number of low-risk miles tends to balance out the higher risk of doubles on nonlimited access roads. The models presented here lend some support for that interpretation. In the casualty accident model that includes bobtails (36-cell), there is a significant term for the doubles-“other” road interaction. The coefficient is 0.4494, which means that the effect of the doubles-“other” road interaction is to increase risk by 56% compared with singles on that road type. Note, however, that in this model the main effect of doubles is -0.1726 , so the net result is that doubles rates on “other” roads are 32% higher than the baseline case on “other” roads. The interaction is also weak. It does not appear at all in the simpler casualty accident model, and it just barely attains significance in the 36-cell model.

On the other hand, for property damage accidents, doubles actually show significantly lower rates than singles in some circumstances. The interaction term for double-“other” roads in the PDO model is -0.4497 . The interaction term for doubles on major arteries is also negative, though small and not statistically significant. Doubles also do better at night, with an interaction term of -0.5114 . The predicted and observed PDO rates in Table 10 show that doubles have lower rates than singles in every “other” road cell. This does not imply that doubles are safer than singles on all road types. The good showing of doubles off limited access roads for property damage accidents is probably because, given an accident, dou-

bles have a somewhat higher probability of producing a casualty. But the data presented here are not consistent with large differences in the relative risk of singles and doubles.

The tractor configuration for which there are substantial differences in relative risk is the bobtail. The main effect for bobtails is among the largest coefficients in the models, almost as large as the coefficient for “other” roads. Bobtails appeared in some interaction terms that mitigated their high relative risk in some situations, but tractors operating without a trailer had much higher accident rates than singles or doubles in every cell of the data matrix. In terms of the absolute number of bobtail accidents, bobtails are less of a problem than singles or doubles. There were 74 bobtail casualty accidents in the study data file, 144 doubles, and 1,153 singles. But, clearly bobtails are a much less safe configuration than singles or doubles. There is much more room for safety improvement in terms of reducing bobtail rates to bring them into line with the rates for singles and doubles.

While the single/doubles question has generated the most interest, operating environment is much more important than vehicle configuration in determining the relative risk of an accident (except for bobtails where it is comparable in magnitude). Among the environmental variables that could be incorporated into the models, road type had the largest impact. The safest roads in the highway system are limited access roads. Compared with limited access

roads, other classes of roads present much higher operating risks for truck-tractors. For example, in the casualty model, the main effect of major artery roads is to increase the risk by about a factor of two, while the risk on "other" roads is almost seven times as high. These are substantial effects, far greater than the singles-doubles difference.

The effect of the different types of road is not the same in all operating environments. Urban areas tend to lower the risk on "other" roads. This may be because operating speeds are lower on urban "other" roads than on rural ones. On the other hand, night tends to raise the risk on "other" roads, beyond the risk that night has overall. For casualty accidents, this interaction has about the same magnitude as the "other"-urban interaction, but it is in the opposite direction. Thus, in urban areas at night, the higher risk of night on "other" roads is canceled out by the lower risk of urban areas on "other" roads, leaving just the main effect of night. The explanation could be that urban roadways are typically lighted at night. The higher risk of "other" roads at night is manifested primarily in rural areas, where operating speeds are higher and where the roadway is typically not lighted.

Area type is also significant in the accident risk model. Urban areas have lower rates than rural areas. The effect is large and significant, for both casualty and PDO accidents. Urban areas have higher traffic densities, but they also have generally lower operating speeds. These two factors pull in opposite directions and express themselves in complex ways. Lower speeds doubtless contribute to a lower probability of injury or death, given an accident. The lower speeds also may contribute to fewer accidents overall. In the casualty accident model, the main effect of the urban factor is to lower the accident risk by about 60%. The effect is not uniform, however—the "other" road-urban interaction term is also strongly negative, indicating that the risk of "other" roads is much higher in rural areas, where speeds are higher, than in urban.

With regard to PDO accidents, traffic density seems to play a greater role in the accident rate. Operating speeds on urban roads are lower than on rural ones, which should lower the accident rate, but traffic densities are higher, which has the opposite effect. The net result is that while urban areas tend to reduce the risk relative to rural areas, the effect is not as strong as in the casualty accident models. The "other" road-urban interaction is still negative but also less strong, and the major-urban interaction is strongly and significantly positive. The association between increased risk of PDO accidents and traffic density probably explains these results. The interpretation is strengthened by the fact that the coefficient

for night, when densities are lower, is small and not statistically significant.

The influence of time differs between casualty accidents and property damage accidents. In casualty accidents, night is associated with a large increase in the relative risk. Clearly, it is not night itself that "causes" accidents, but factors associated with night. For example, the likelihood of driver fatigue is higher at night, the proportion of drinking drivers is higher, and the ability to see and react to hazards in unlighted areas is reduced. The result of these and other factors associated with night is to increase substantially the relative risk of accident involvement. The main effect of night increases the risk of casualty accidents by about 45% compared to day. On lower-quality roads, the combination of the factors associated with night and the more restrictive geometry of "other" roads results in much higher rates, beyond the separate effects of night and the "other" road type. That is, the interaction term for the "other" roads-night combination is positive and substantial, amounting to an additional 85% increase in risk beyond the effects of night and "other" roads alone.

For property damage accidents, the main effect of night is small and not statistically significant. Traffic densities are lower at night, which may account for the weak effect of night. There are some interesting interactions, however. The "other" roads-night interaction is large and positive, while the night-urban interaction is large and negative. This indicates that PDO rates at night on "other" roads in rural areas are substantially higher than in urban areas, which are generally lighted. "Other" roads in rural areas are typically not lighted which, along with the lower quality of "other" roads, results in substantially increased accident risk despite the lower traffic density. In urban areas, PDO rates for singles on other roads are about the same at night as they are during the day.

CONCLUSION

Overall, operating environment is much more important in determining the risk of accident involvement than vehicle configuration, except for bobtails. For singles and doubles, road type, area type, and time of day have much more impact on the probability of an accident than whether the vehicle is a single or double. Many questions remain, of course. The categories for singles and doubles combine distinct populations of trucks. Among doubles, a substantial fraction of the population consists of operations by the large interstate LTL freight haulers using twin 28-foot van trailers. Another segment of the doubles population comprises intrastate operations,

such as the doubles used by agriculture in California, logging doubles in the Northwest, and the heavy doubles that operate in Michigan. This latter population probably has a quite different accident experience from long haul van doubles. Clearly distinguishing the two groups would help clarify the issue. Similarly, the influence of size, weight, cargo body type, and many other factors would help narrow the focus to areas where a significant contribution to improving safety could occur.

Data remain a problem. The level of analysis offered here could not have been undertaken without special data collection. There are no existing, publicly available sources of truck accident and exposure data, other than the travel and fatal accident files maintained by UMTRI, that permit the calculation of accident rates at the level of cargo body style and gross vehicle weight. Nor will any such files be available in the foreseeable future. The program outlined by the Transportation Research Board offers hope of incremental improvement, though exposure data will apparently remain inadequate. The National Governor's Association-encouraged truck accident supplemental data form is a step in the right direction. In the meantime, data inadequacies are a substantial impediment to understanding the truck safety problem.

REFERENCES

- Aitkin, M.; Anderson, P.; Francis, B.; Hande, J. *Statistical modelling in GLIM*. New York: Oxford University Press; 1989.
- Blower, D. F.; Lyles, R. W.; Campbell, K. L.; Stamatiadis, P. *The Michigan heavy truck study*. Ann Arbor: Michigan State University and University of Michigan Transportation Research Institute; April 1990.
- Blower, D. F.; Pettis, L. C.; Sullivan, K. P. *Trucks involved in fatal accidents, 1987*. Ann Arbor: University of Michigan Transportation Research Institute; April 1990.
- Buyco, C.; Saccomanno, F. F. *Analysis of truck accident rates using loglinear models*. *Canadian Journal of Civil Engineering* 15:397-408; June 1988.
- Campbell, K. L.; Blower, D. F.; Gattis, R. G.; Wolfe, A. C. *Analysis of accident rates of heavy-duty vehicles*. Ann Arbor: University of Michigan Transportation Research Institute; April 1988.
- Carsten, O. *Safety implications of truck configuration*. *Transportation Research Record* 1111:17-26; 1987.
- Chirachavala, T.; O'Day, J. *A comparison of accident characteristics and rates for combination vehicles with one or two trailers*. Ann Arbor: Highway Safety Research Institute, University of Michigan, August 1981.
- Fritzler, M. L. *Are singles safer?* Unpublished paper. August 7, 1988.
- Joshua, S.; Garber, N. *Estimating truck accident rates and involvements using linear and Poisson regression models*. *Transportation Planning and Technology* 15(1):41-45, 1990.
- Jovanis, P. P.; Chang, H.-L.; Zabaneh, I. *Comparison of accident rates for two truck configurations*. *Transportation Research Record* 1249:18-29; 1989.
- Lee-Gosselin, M.; Richardson, A. J.; Taylor, G. *Truck accident involvement with and without front-axle brakes: Application for case-control methodology*. *Transportation Research Record* 1270:46-56; 1990.
- McCullagh, P.; Nelder, J. A. *Generalized linear models*. 2nd ed. Cambridge, England: Cambridge University Press, 1989.
- Saccomanno, F. F.; Buyco, C. *Generalized loglinear models of truck accident rates*. *Transportation Research Record* 1172:23-31; 1988.
- Sparks, G. A.; Bielka, J. *Large truck accident experience in western Canada: A case study of two large fleets*. *Symposium on the role of heavy freight vehicles in traffic accidents*. Vol. III. Ottawa, Ontario: Roads and Transportation Association of Canada; 1987.
- Stein, H. S.; Jones, I. S. *Crash involvement of large trucks by configuration: A case-control study*. *Am. J. Public Health* 78:491-498; May 1988.
- TRB (Transportation Research Board). *Twin trailer trucks: Effects on highways and highway safety*. Special report 211. Washington, DC: National Research Council; 1986.
- TRB. *Data requirements for monitoring truck safety*. Special report 228. Washington, DC: National Research Council; 1990.