Estimating Safety Effects of Cross-Section Design for Various Highway Types Using Negative Binomial Regression

MOHAMMED A. HADI, JACOB ARULDHAS, LEE-FANG CHOW, AND JOSEPH A. WATTLEWORTH

Improvements in cross-section design are expected to reduce crash rates. Previous studies on the subject have concentrated on two-way, two-lane highways and given less attention to other types of highways. In addition, several of those studies used conventional regression analyses that are not suitable to estimate a discrete non-negative variable like crash frequency. This study uses negative binomial regression analyses to estimate the effects of cross-section design elements on total, fatality, and injury crash rates for various types of rural and urban highways at different traffic levels. The results show that, depending on the highway type investigated, increasing lane width, median width, inside shoulder width, and/or outside shoulder width are effective in reducing crashes. The results also indicate that on four-lane urban highways, the raised median is safer than the two-way left-turn lane median and that the use of an open-graded friction course in lieu of a dense-graded friction course does not have any effect on crash rates.

Many studies have been conducted in the last four decades to investigate the effects of various highway designs on safety. The design elements that have been found to affect safety include cross-section design, horizontal alignment, vertical alignment, roadside features, intersection designs, interchange designs, narrow bridges, lighting, access control, pavement conditions, pavement edge drops, speed limit, marking, signing, delineation, railway crossings, pedestrian facility designs, and bicycle facility designs.

Previous studies indicated that improvements to highway design could produce significant reductions in the number of crashes. However, only a few quantitative relationships have been developed to relate various design elements to crash rates. The development of such relationships is necessary because they provide the information required to make the trade-off between the cost and the benefit of better highway designs. In addition, they permit better prioritization of safety improvement projects.

This study quantifies the effects of cross-section design elements on total, fatality, and injury crash rates for various types of rural and urban highways using data from Florida at different traffic levels. Poisson regression and negative binomial regression are investigated for possible use in deriving the estimation equations.

BACKGROUND

A large proportion of the available research on the safety effects of highway design elements has been devoted to two-way, two-lane rural highways; less attention has been given to other highway types. As discussed in a later section, regression analyses were used in some studies to develop quantitative relationships that estimate crashrate as a function of cross-section design elements. The majority of the studies that employed regression analyses used conventional regression to derive the required relationships.

Conventional regression analyses assume that the dependent variable is continuous and normally distributed with a constant variance. These assumptions are not correct when estimating crashes. Crash frequency is a discrete non-negative variable, and its variance depends on its mean. Thus, Poisson regression and negative binomial regression have been used in recent studies for crash data analysis (*1–3*).

Zegeer and Deacon (*4*) reviewed 30 studies performed until the mid-1980s and concluded that no satisfactory quantitative model relating crash rates to lane and shoulder width could be found. Therefore, they calibrated a regression model that estimates the most likely relationships between crashes and lane width, shoulder width, and shoulder type on two-lane rural highways. This model was derived using data obtained from four previous studies.

Later, Zegeer et al. (*5*) developed another regression model to quantify the benefits of shoulder and lane improvements based on data selected from seven states. Only two-lane roadway sites were investigated.

Several studies were conducted to investigate the effects on safety of median width and type. Few of them produced quantitative relationships. Among these was a study by Squires and Parsonson (*6*) in which regression equations were derived to predict crash rates on sections with raised medians and continuous two-way left-turn lanes (TWLTLs) for four- and six-lane highways. Based on these equations, the following points were concluded:

- For four-lane sections, raised medians had a lower crash rate over the range of data studied.
- For six-lane sections, the results were mixed. Raised medians were found to have lower crash rates for most conditions. However, TWLTLs had a lower crash rate where few concentrated areas of turns, such as signalized intersections and unsignalized approaches, existed.

Hoffman (*7*) examined four TWLTL sites with annual average daily traffic (AADT) ranging form 15,000 to 30,000 vehicles per day (vpd). Existing four-lane undivided highways were widened to five lanes to accommodate a center lane for left turn. Crash data for 1 year before and 1 year after the implementation of a TWLTL were examined. The results showed a 33 percent reduction in total crash frequency.

Transportation Research Center, University of Florida, 512 Weil Hall, Gainesville, Fla. 32611.

Bowman and Vecellio (*8*) compared arterial sections with different median types using statistical comparison tests and found that, in the central business district (CBD) areas, TWLTL medians had a lower vehicle crash rate than both raised curb median and undivided cross sections. Undivided arterials had the highest crash rate. In suburban areas, raised curb medians provided the lowest crash rates. In both CBD and suburban locations, raised medians had a lower injury crash rate than both the TWLTL median and undivided cross-sections.

In another paper, Bowman et al. (*2*), using the same data as those used by Bowman and Vecellio (*8*), developed equations to estimate crashes for urban and suburban arterial sections with different median types. Negative binomial regression was used in the analyses. Three different regression models were developed to estimate crashes for raised curb, TWLTL medians, and undivided cross sections. No attempt was made to compare the safety effects of the three median types using the derived models. The study concluded that the derived models provide better estimates of vehicle crash frequency than did the models from prior research. This conclusion, however, should be taken with caution, because the same data used in deriving the models in that study were used in the comparison.

Knuiman et al. (*3*) studied the effect of median width on crash rate using a negative binomial regression model. For a median without a barrier, it was found that crash rates declined rapidly when median width exceeded about 7.6 m (25 ft). The decreasing trend seemed to become level at median widths of approximately 18.9 to 24.4 m (60 to 80 ft).

DATA ACQUISITION

The largest possible number of roadway segments from the state of Florida roadway system were used in this study. For each segment, roadway, traffic, and crash data were needed for the regression analyses.

Roadway and traffic data were obtained from the Florida Department of Transportation's Roadway Characteristics Inventory (RCI) system. Crash data were obtained from the Department of Highway Safety and Motor Vehicles' computerized accident record system. Data from the two systems were linked through their common location data.

Roadway samples were stratified by location, access type, and number of lanes into nine categories. These categories and the AADT range used in analyzing each category were:

- Rural freeways (5,000–60,000 vpd),
- Four-lane rural divided roads (1,145–40,000 vpd),
- Two-way, two-lane rural roads (200–10,000 vpd),
- Four-lane urban freeways (4,260–136,800 vpd),
- Six-lane urban freeways (20,000–200,000 vpd),
- Two-way, two-lane urban collectors (904–38,680 vpd),
- Four-lane urban divided roads (10,000–50,000 vpd),
- Six-lane urban divided roads (10,000–100,000 vpd), and
- Four-lane urban undivided roads (5,000–40,000 vpd).

The selection of geometric design variables for use in this study was based on the completeness of data (proportion of missing data) and the degree to which those variables are expected to affect safety. The variables selected on the basis of these two factors for possible inclusion as independent variables in the derived models were section length, AADT, lane width, outside paved and unpaved shoulder widths, inside paved and unpaved shoulder widths, median width and type, presence of curb, speed limit, number of intersections (or interchanges in case of freeways) and the use of an open-graded versus a dense-graded friction course. The dependent variables were the total number of crashes, number of injury crashes, and number of fatal crashes in 4 years.

Roadway samples were divided into sections such that the geometric design remained constant within a given section. A minimum highway section length of 0.05 mile was set to exclude short sections that might be influenced by adjacent section characteristics. Variable section lengths were used in this study, but the section length was included as an independent variable in the analyses. A section boundary was formed when one or more of the geometric characteristics changed.

Four years (1988–1991) of crash data were used for analyses. This period is long enough for stable rates of crashes to be obtained. A longer analysis period was not used because of the possibility of significant changes in traffic patterns, highway design features, or land uses over the period.

POISSON AND NEGATIVE BINOMIAL REGRESSION MODELS

Poisson regression is based on the assumption that the response (the dependent) variable is Poisson-distributed. The Poisson distribution models the probability of discrete events such as crashes according to the Poisson process as follows:

$$P(Y) = \frac{e^{-\mu}\mu^{Y}}{Y!} \tag{1}$$

where Y is the number of events in a chosen period and μ is the mean number of events in the chosen period.

The Poisson regression model assumes that the mean number of events is a function of regressor variables. Thus, to estimate crash frequency using Poisson regression, the number of crashes is assumed to be Poisson distributed according to the following equation:

$$P(Y=Y_i) = \frac{e^{-\mu_i(X_i,\beta)}[\mu_i(X_i, \beta)]^{Y_i}}{Y_i!} \tag{2}$$

where

Y_i = the number of crashes observed at road section i for a chosen period of time,
β = a vector representing a set of parameters to be estimated,
$\mu_i(X_i, \beta)$ = the mean number of crashes on road section i, which is a function of a set of regressor variables X, and
X_i = a vector representing the value of the regressor variables for highway section i.

The regressor variables in the above equation can be selected as various design and traffic operational variables that could influence safety.

The function $\mu_i(X_i, \beta)$ in Equation 2, which relates the distribution mean to regressor variables, is called the "link function" (*9*). The function used in this study is as follows:

$$\mu_i(X_i, \beta) = e^{X_i\beta} \tag{3}$$

The regression coefficients (β) in Equations and 3 are estimated using the maximum likelihood estimation.

The model of Equation 2 assumes that crash frequency is Poisson-distributed. This implies that the variance is assumed to be equal to the mean of the process. Although in theory the Poisson model is suitable for count data such as crash frequencies, these have been reported to display extra variation or overdispersion relative to a Poisson model. That is, the variance observed was higher than the mean. The overdispersion in crash data could be because not all relevant variables are normally included in the model and because of the uncertainty in regressor variables (*1*). When using the Poisson regression in the presence of overdispersion, maximum likelihood parameter estimates are consistent, but the variances of these parameters are inconsistently estimated. As a result of this, hypothesis test results become invalid.

To deal with overdispersion, negative binomial regression has been suggested for use instead of Poisson regression (*1–3*). This type of regression allows the variance of the process to differ from the mean. In this study, tests for overdispersion were conducted to decide whether Poisson or negative binomial regression should be used for model development.

The maximum-likelihood estimation of model parameters was performed using the LIMDEP statistical package (*10*). This package estimates the Poisson regression parameters using the Newton's method and the negative binomial regression parameters using a modification to the Davidson, Fletcher, and Powell search procedure (*11*). When using the negative binomial regression, LIMDEP starts with parameter values achieved during a Poisson regression analysis. This is expected to produce better values when maximizing the log-likelihood function.

Because ordinary regression was not used, the selection of variables for inclusion in the final models and statistical tests to determine the significance of the derived relationships could not be done using conventional approaches. Rather, methods that do not assume normality of the dependent variable were used. These are reviewed in the following subsections.

Selection of Regressor Variables

In order to decide which subset of independent variables should be included in a crash estimation model, the Akaike's information criterion (AIC), was used. AIC was defined as follows (*12*):

$$\text{AIC} = -2 \cdot ML + 2 \cdot K \tag{4}$$

where K is the number of free parameters in the model and ML is the maximum log-likelihood.

The smaller the value of AIC, the better the mode. Starting with the full set of independent variables listed in the previous section, a stepwise procedure was used to select the best model based on minimizing the AIC value.

Testing Individual Coefficients

Individual parameters in the β vector of Equation 2 were tested to investiage the null hypothesis that a given parameter β_j is zero. The method used was based on the standard errors of coefficients, which is an analog to the *t*-test used in conventional regression analyses. In this method, the following term is computed (*9*):

$$\chi^2 = \frac{b_j^2}{(SE_j)^2} \tag{5}$$

where b_j is the estimate of β_j and SE_j is the standard error of the coefficient β_j.

A chi-square test with one degree of freedom was used to test the hypothesis that the parameter β_j is zero.

Testing for Overdispersion

Two tests for detecting overdispersion in the Poisson process were performed in this study to decide whether Poisson regression or negative binomial regression should be used in analyzing crash data. The first test was suggested by Cameron and Trivedi (*13*). This test involved simple least-squares regressions to test the significance of the overdispersion coefficient.

Another test for overdispersion was performed using outputs from the procedure that estimated the negative binomial regression in LIMDEP. These outputs (an overdispersion parameter and its standard error) were used to test the hypothesis that the overdispersion parameter was zero. Failure to reject this hypothesis would indicate insignificant dispersion and allow the use of Poisson regression.

Goodness of Fit

To measure the goodness of fit, the Pearsons chi-square statistic was used. This statistic was calculated as follows:

$$\chi^2 = \sum \frac{(Y_i - \mu_i)^2}{\mu_i} \tag{6}$$

The degree of freedom of this statistic equals the number of observations minus the total number of estimated parameters.

MODEL ESTIMATION AND TESTING

The Poisson regression and negative binomial regression analyses described in the previous section were considered for use in establishing the required relationships between cross-section design elements and total number of crashes, fatal crashes, and injury crashes. Separate sets of models were developed for each of the nine highway categories investigated in this study. In addition, the analyses were performed separately for nonintersection, or midblock, crashes and all crashes. The latter include, in addition to midblock crashes, intersection crashes, interchange crashes, and railway crossing crashes.

The independent variables in the models were selected based on the AIC value from the total set of geometric design and operational variables extracted from the RCI. The square, square root, and logarithmic terms of several of the variables were investigated for possible inclusion in the final models. In addition, interaction terms that include multiplications of variables that were thought to interact with each other were also tested for inclusion in the models.

Categorical variables (variables, each observation of which belongs to one of several distinct categories) such as median type, curb presence, and friction course type were represented as dummy variables. To represent K categories of a categorical variable, $K - 1$, binary (0 or 1) dummy variables were created.

The AADT, section length, number of intersections, number of interchanges, lane width, shoulder width, and median width were represented as continuous variables. However, there was some concern that models with continuous representation of cross-section design elements might not be able to identify thresholds above which improvements to these elements do not reduce crash rates. For example, if increasing lane width from 2.8 m to 4.0 m (9 ft to 13 ft) decreased crash rate, but increasing lane width from 4.0 m to 4.3 m (13 ft to 14 ft) did not, a continuous model might attempt to smooth the data according to the assumed transformation and thus might estimate incorrect reduction in crashes due to a widening a lane from 4.0 m to 4.3 m (13 ft to 14 ft).

For the above reason, additional Poisson and negative binomial regression analyses were performed in which lane width, shoulder width, and median width were represented as categorical variables. Based on the results of these analyses, the category of each variable that produced the minimum crash rate was identified and used to determine the threshold above which no reduction in crash rate was expected from improvements to that variable.

The LIMDEP package produced some of the information required to perform statistical inferences on the models developed in this study. Information not produced automatically by the package such as the AIC and the chi-square goodness-of-fit statistics was calculated using LIMDEP commands.

Tests for overdispersion indicated that overdispersion in crash data was significant for all types of highway investigated. Thus, it was decided to use negative binomial regression rather than Poisson regression to estimate model parameters in all cases investigated.

All derived Poisson and negative binomial models failed to pass the chi-square goodness-of-fit test at the 0.05 confidence level. Similar results were reported by other researchers who tried to estimate crashes based on geometric and operational design elements (*1,2*). Bowman et al. (*2*) pointed out that the chi-square test is not suitable for nonlinear problems such as the one under investigation. Miaou et al. (*1*) suggested that the lack of goodness of fit could also be related to the following:

- A large proportion of the roadway sections have very few or no crashes. The chi-square test is not appropriate for these conditions.
- There may have been uncertainties or omitted variables in the data.

Thus, they suggested using the following criteria for model acceptance:

- The signs of all model parameters are as expected.
- AIC is the lowest possible.
- Each individual parameter is accepted when tested using appropriate statistical testing.

TABLE 1 Models Derived To Estimate Total Crash Frequency in 4-Year Period

Highway Type	Crash Location	Model[a]
two-lane rural	mid-block	exp(-10.26 +.8249Llen +.8783Ladt -.0857Lw -.0130Sp +.0589Is -.0150Ts)
	total	exp(-9.053 +.7212Llen +.8869Ladt -.0435Lw -.0262Sp +.1145Is -.0123Ts)
four-lane rural divided	mid-block	exp(-9.545 +.6706Llen +.7205Ladt -.0524Su +.1746Is -.0458Sm)
	total	exp(-7.908 +.4140Llen +.7672Ladt -.0129Su -.3503Is -.0688Sm)
four-and six lane rural freeways	mid-block	exp(-12.89 +.9020Llen +.9156Ladt -.0272Ip +.2164Ic-.0252Sm)
	total	exp(-12.14 +.8533Llen +.9032Ladt -.0252Ip +.4679Ic -.0472Sm)
two-lane urban undivided	mid-block	exp(-10.62 +.8966Llen +.9008Ladt -.0355Lp -.0234Sp +.1707Co +.0603Is - .0323Up)
	total	exp(-8.263 +.7212Llen +.8560Ladt -.0246Lp -.0307Sp +.3652Co +.1111Is - .0387Up)
four-lane urban undivided	mid-block	exp(-8.275 +.8646Llen +.8318Ladt -.1127Lw -.0301Sp -.2831Co +.0427Is)
	total	exp(-4.251 +.6914Llen +.6950Ladt -.1056Lw -.0536Sp -.3101Co +.8251Is - .0309Ps)
four-lane urban divided	mid-block	exp(-13.88 +.7009Llen +1.195Ladt -.0299Ps +.1131Is -.0588Sm +.0982D1 -.2008D2 - .0871D3)
	total	exp(-9.996 +.4890Llen +1.026Ladt -.0367Ps +.2053Is -.1060Sm +.1115Ci)
six-lane urban divided	mid-block	exp(-12.04 +.8223Llen +1.072Ladt -.0270Sp +.0631Is -.0412Sm +.1671Co)
	total	exp(-8.766 +.6335Llen +.8152Ladt -.0026Mw +.1309Is +.2819Co)
four-lane urban freeways	mid-block	exp(-8.837 +.7848Llen +1.213Ladt -.3909Lw -.0263Up -.0225Sp +.2786Ic - .0801Sm)
	total	exp(-8.972 +.0.7292Llen+1.171Ladt -.2585Lw -.0268Sp +.3674Ic - .0926Sm)
six lane urban freeways	mid-block	exp(-13.56 +.8753Llen +1.454Ladt -.3504Lw -.0667Ps +.1787Ic - .0345Sm)
	total	exp(-8.163 +.8049Llen +1.178Ladt -.3740Lw -.0445Ps -.0310Sp +.2935Ic)

[a] *Lw = lane width (ft), Lp = pavement width (ft), Ps = paved shoulder width (ft), Up = unpaved shoulder width (ft), Ts = total shoulder width (ft), Mw = median width (ft), Sm = (median width)$^{.5}$, Su = (unpaved shoulder width)$^{.5}$, Ip = inside paved shoulder, Sp = posted speed limit (mph), Is = number of intersections, Ic = number of interchanges, Co = presence or absence of outside curb (1,0), Ci = presence or absence of inside curb (1,0), Llen = log of (1000 • section length in miles), Ladt = log (AADT), D1 = TWLTL median, D2 = grass median, D3 = raised curb median and median type is crossover resistance when D1=D2=D3= 0*

TABLE 2 Models Derived To Estimate Injury Crash Frequency in 4-Year Period

Highway Type	Crash Location	Model[a]
two-lane rural	mid-block	exp(-10.72 +.8157Llen +.8681Ladt -.0787Lw -.0108Sp +.0601Is -.021Ts)
	total	exp(-9.478 +.7064Llen +.8806Ladt -.0426Lw -.0236Sp +.1155Is -.013Ts)
four-lane rural divided	mid-block	exp(-9.91 +.6288Llen +.6919Ladt +.1973Is)
	total	exp(-8.36 +.3849Llen +.76Ladt +.3617Is -.0455Sm -.0223Sp)
four-and six lane rural freeways	mid-block	exp(-14.032 +.9107Llen +.9599Ladt -.0407Ip +.2127Ic)
	total	exp(-13.19 +.88667Llen +.9527Ladt -.0307Ip +.43Ic -.0463Sm)
two-lane urban undivided	mid-block	exp(-11.415 +.933Llen +.9137Ladt -.0489Lp -.0201Sp +.056Is -.0342Up)
	total	exp(-9.065 +.7451Llen +.864Ladt -.0337Lp -.0253Sp +.108Is -.043Up + 1.48Co)
four-lane urban undivided	mid-block	exp(-9.584 +.8831Llen +.8317Ladt -.1037Lw -.015Sp -.3318Co +.0395Is)
	total	exp(-5.285 +.699Llen +.6993Ladt -.128Lw -.0371Sp -.3407Co +.08Is)
four-lane urban divided	mid-block	exp(-14.023 +.7979Llen +1.216Ladt -.0303Ps +.0839Is -.0325Sm -.0295Sp)
	total	exp(-11.2 +.5254Llen +1.0625Ladt -.0353Ps +.3617Is -.0833Sm +.1191Ci)
six-lane urban divided	mid-block	exp(-14.0 +.8164Llen +1.0934Ladt +.0701Is -.0501Sm +.2202Co)
	total	exp(-8.536 +.7022Llen +.8491Ladt +.113Is +.1311Co -.05Sm -.0278Sp)
four-lane urban freeways	mid-block	exp(-10.61 +.7733Llen +1.1832Ladt -.307Lw -.0232Up -.0154Sp +.24Ic -.06Sm)
	total	exp(-12.6 +.712Llen +1.1373Ladt -.0223Sp +.3512Ic -.0706Sm)
six lane urban freeways	mid-block	exp(-14.04 +.93Llen +1.405Ladt -.339Lw -.0594Ps -.031Sm)
	total	exp(-8.507 +.8418Llen +1.14Ladt -.3845Lw -.0370Ps -.0302Sp +.2433Ic)

[a] *Lw = lane width (ft), Lp = pavement width (ft), Ps = paved shoulder width (ft), Up = unpaved shoulder width (ft), Ts = total shoulder width (ft), Mw = median width (ft), Sm = (median width)$^{.5}$, Su = (unpaved shoulder width)$^{.5}$, Ip= inside paved shoulder, Sp = posted speed limit (mph), Is = number of intersections, Ic = number of interchanges, Co = presence or absence of outside curb (1,0), Ci = presence or absence of inside curb (1,0), Llen = log of (1000 • section length in miles), Ladt = log (AADT).*

These three criteria were used in this study. All of the models derived and discussed in the following section satisfied the above three criteria. Tables 1–3 present the models derived to estimate total, injury, and fatal crash frequencies for a 4-year period. Table 3 indicates that in most cases, the best model found to estimate fatal crashes included only the section length and AADT as explanatory variables. This might be because the number of fatal crashes were too low to allow for correct estimation of the effect of other variables.

MODEL IMPLEMENTATION

This section presents results obtained from the regression analyses for each of the nine highway types investigated. These results were obtained using examples in which the effects of each significant variable on safety was investigated by changing its value while keeping the other significant variables constant at values representing typical roadway sections for the investigated highway type. The models with continuous representation of lane, shoulder, and median width were used in the analyses. However, the thresholds obtained based on categorical representation, as explained in the previous section, were also considered.

The values presented in the following discussion are those for midblock crashes. All-location crashes generally follow similar trends. In the following presentations, crash frequency is expressed in number of crashes per year (C/Y) and crash rate is expressed in number of crashes per million vehicle-kilometers (C/MVKM).

Annual Average Daily Traffic

The derived models presented in Tables 1–3 indicate that crash frequency increases with higher AADT for all highway types investigated. Figure 1 indicates that when comparing different rural highways with the same AADT, two-way, two-lane highways have the highest crash rates, followed by freeways, then by four-lane divided highways. For all three rural highway types, increasing the AADT decreased crash rates. The rate of this decrease, however, is lower for rural freeways than for the other two highway types; this may be due to the higher volumes on freeways.

For urban highways, Figure 2 suggests that crash rates decreased in the following order: four-lane undivided, two-way two-lane, six-lane divided, six-lane freeways, four-lane freeways, and four-lane divided. The crash rates for the last two highway types were very close. Higher AADT levels were associated with lower crash rates for two-way two-lane and four-lane undivided highways. However, higher AADT levels resulted in higher crash rates for urban freeways and other urban divided highways.

Lane Width

Significant relationships could be found between lane width and crashes for undivided highways and urban freeways. For other highway types, no such relationship could be identified. Table 4 shows that, based on categorical representation of lane width, for two-lane

TABLE 3 Models Derived To Estimate Fatal Crash Frequency in 4-Year Period

Highway Type	Crash Location	Model[a]
two-lane rural	mid-block	exp(-15.47 +1.025Llen +.9624Ladt -.1428Lw)
	total	exp(-14.401 +.875Llen +.9362Ladt -.097Lw)
four-lane rural divided	mid-block	exp(-12.644 +.7904Llen +.6036Ladt)
	total	exp(-10.526 +.6404Llen +.541Ladt)
four- and six-lane rural freeways	mid-block	exp(-14.758 +.9714Llen +.7057Ladt)
	total	exp(-14.054 +.947Llen +.6673Ladt)
two-lane urban undivided	mid-block	exp(-12.504 +.8872Llen +.6675Ladt -.11Lp)
	total	exp(-10.93 +.9793Llen +.467Ladt -.0777Lp)
four-lane urban undivided	mid-block	exp(-17.8 +1.281Llen +.854Ladt)
	total	exp(-14.839 +1.0812Llen +.7735Ladt +.4849Co)
four-lane urban divided	mid-block	exp(-14.321 +1.0237Llen +.6193Ladt)
	total	exp(-13.59 +.9514Llen +.6765Ladt)
six-lane urban divided	mid-block	exp(-14.251 +.945Llen +.676Ladt)
	total	exp(-10.88 +.73Llen +.5376Ladt +.0754Is)
four-lane urban freeways	mid-block	exp(-13.861 +.9116Llen +.6326Ladt)
	total	exp(-13.723 +.789Llen +.727Ladt)
six-lane urban freeways	mid-block	exp(-19.835 +1.2169Llen +1.01Ladt)
	total	exp(-12.41 +1.242Llen +1.152Ladt)

[a] *Lw = lane width (ft), Lp = pavement width (ft), Ps = paved shoulder width (ft), Up = unpaved shoulder width (ft), Ts = total shoulder width (ft), Mw = median width (ft), Sm = (median width)$^{.5}$, Su = (unpaved shoulder width)$^{.5}$, Ip = inside paved shoulder, Sp = posted speed limit (mph), Is = number of intersections, Ic = number of interchanges, Co = presence or absence of outside curb (1,0), Ci = presence or absence of inside curb (1,0), Llen = log of (1000 • section length in miles), Ladt = log (AADT).*

rural, two-lane urban, four lane urban undivided, and urban freeways, widening lane width up to 4.0 m, 3.7 m, 4.0 m, and 4.0 m (13 ft, 12 ft, 13 ft, and 13 ft), respectively, could be expected to decrease crash rates.

Figure 3 shows that the highest benefits of lane widening were estimated for urban freeways, followed by four-lane undivided urban highways, then by two-lane rural highways. For two-lane urban highways, there was a significant relationship between pavement width (lane width plus paved shoulder width) and crash frequency, rather than between lane width and crash frequency, when continuous representations of variables were used. The effect of lane width on crash rate for this highway type was lower than for other highway types.

Shoulder Width

Figure 4 indicates that the safety significance of outside total shoulder width, paved shoulder, and unpaved shoulder widths depends on the highway type investigated. Increasing paved shoulder width

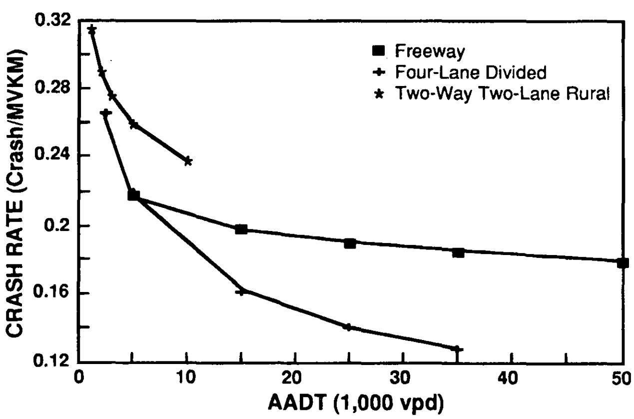

FIGURE 1 Effect of AADT on midblock crash rates of rural highways.

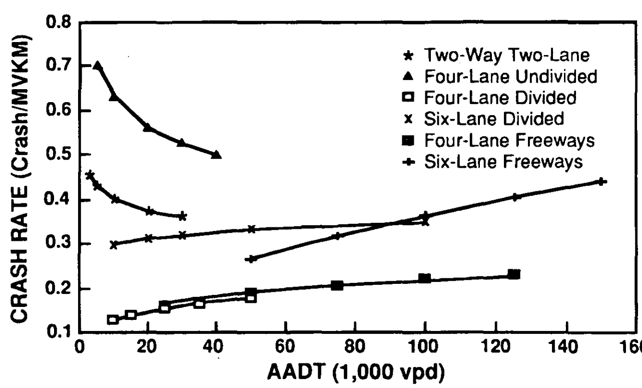

FIGURE 2 Effect of AADT on midblock crash rates of urban highways.

TABLE 4 Values of Lane and Shoulder Widths That Produced Lowest Crash Rates Based on Models Derived Using Categorical Representations of These Variables

Highway Type	Significant Variable	Variable Range[a] (m)[b]	Category with the Minimum Crashes (m)[b]
two-lane rural	lane width	2.8-4.6	4.0
	total shoulder width	0.6-3.7	3.0-3.7
four-lane rural divided	unpaved shoulder width	0.0-3.7	3.0-3.7
rural freeway	inside paved shoulder	0.0-1.8	-[c]
two-lane urban	lane width	3.0-4.6	3.7
	paved shoulder width	0.0-2.4	1.5-2.1
	unpaved shoulder width	0.0-3.0	2.4-3.0
four-lane urban undivided	lane width	2.7-4.0	4.0
four-lane urban divided	paved shoulder width	0.0-3.0	3.0
four-lane urban freeway	lane width	3.7-4.0	4.0
	unpaved shoulder width	0.0-3.0	3.0
six-lane urban freeways	lane width	3.4-3.7	-[c]
	paved shoulder width	0.0-3.7	-[c]

[a] *the range in the RCI data base*
[b] *1 m = 3.28 ft*
[c] *'-' indicates that no model could be developed based on categorical representation of this variable. Only a continuous model could be developed.*

was estimated to produce lower crash rates on six-lane urban freeways, two-lane urban highways, and four-lane urban divided highways. Increasing unpaved shoulder width was estimated to decrease crash rates on four-lane rural highways, two-lane urban highways, and four-lane urban freeways. Greater total shoulder width (paved plus unpaved) was associated with lower crash rates on two-lane rural highways.

It appears from Figure 4 that greater outside shoulder width was particularly effective in reducing crashes on six-lane urban freeways and two-lane urban highways. Table 4 shows that, based on categorical representations of variables, widening outside shoulder width to between 3.0 and 3.7 m (10 and 12 ft) produced the best results, in most cases. In urban two-lane highways, unpaved shoulder width of 2.4–3.0 m (8–10 ft) and a paved shoulder of about 1.5–2.1 m (5–7 ft) produced the best results.

Use of an inside paved shoulder 1.2–1.8 m (4–6 ft) wide was found to be very effective in decreasing crashes on rural freeways. It was found that using a 1.8-m (6-ft) shoulder width could decrease crash rate by 15.7 percent.

Median Width

In spite of several attempts, useful models that use categorical representation of median width could not be derived in this study. Therefore, the results presented for median width are based on continuous models.

Figure 5 shows the effects of median width on mid-block crash rates. It appears that significant reduction in crash rates could be expected from greater median width for all highway types. In all cases, the square root of median width, rather than median width,

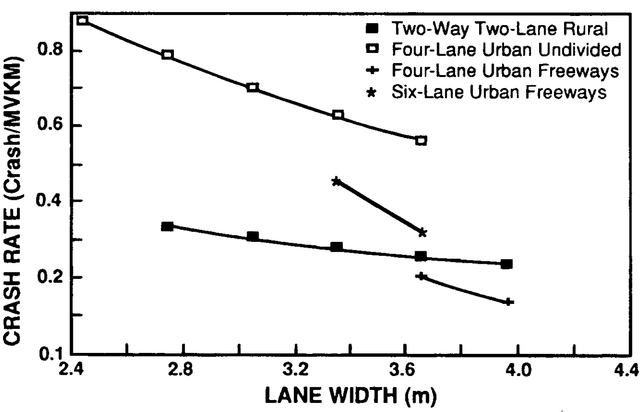

FIGURE 3 Effect of lane width on midblock crash rates.

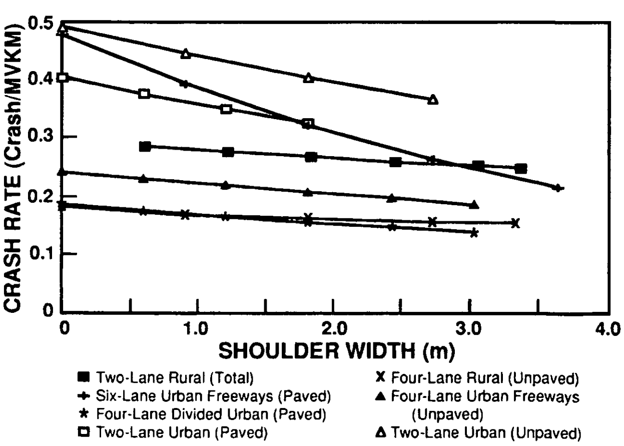

FIGURE 4 Effect of shoulder width on midblock crash rates.

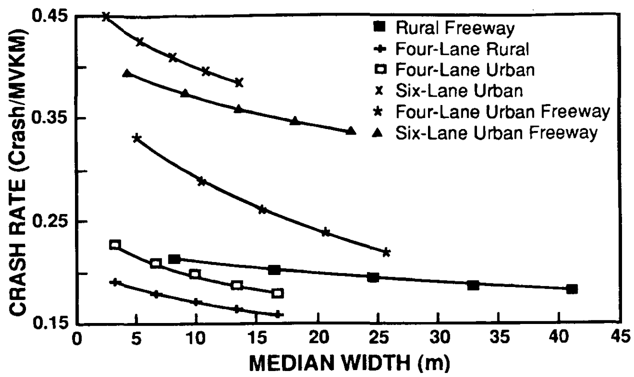

FIGURE 5 Effect of median width on midblock crash rates.

was the significant variable included in the models. This indicated that, as expected, the benefits of greater median width decreased as the median width increased. The safety benefit of increasing median width seems to be the highest in four-lane urban freeways and six-lane urban highways followed by six-lane urban freeways and four-lane urban highways followed by four-lane rural highways and rural freeways.

Median Type

A significant relationship was found between median type and crash experience only in four-lane divided urban highways. Sections with four median types were examined for this highway type. It was found that the safety of median type decreased in the following order: flush unpaved median (grass), raised curb, crossover resistance, and TWLTL. However, the differences in the median width between median types should be noted. For these four median types, the median width ranges based on the RCI database were 4.9–19.5 m, 0.9–11.3 m, 1.8–12.2 m, and 3–4.9 m (16–64ft, 3–37 ft, 6–40 ft, and 10–16 ft), respectively.

Other Elements

Longer section lengths and higher speed limits were associated with lower crash rates. This could be because longer section lengths indicate more uniform cross-section design and sections with higher speed limits generally have higher design speeds. The type of friction course (open-graded versus dense-graded) was not found to be significant in affecting safety.

The presence of more intersections on non-freeway highways and more interchanges on freeways increased crash rates significantly. The effect of intersection/interchange presence in rural highways decreased in the following order: freeways, four-lane divided, and two-lane. For urban highways, the effect decreased in the following order: four-lane freeways, six-lane freeways, four-lane divided, six-lane divided, two-lane undivided, and four-lane undivided.

In all urban highways except four-lane urban undivided highways, the presence of a curb had an adverse affect on vehicular safety.

POSSIBLE MODEL APPLICATIONS

The models derived in this study, along with any relevant models found in the literature, are being used to assess the cost-effectiveness of cross-section design standard requirements. In this analysis, an exhaustive search optimization procedure is being performed to determine the optimal cross-section designs that maximize the cost-effectiveness for each highway type, taking traffic level into consideration. The optimized cost includes four components: crash, construction, travel time, and vehicle operation costs.

The models derived in this study could also be used to examine the cost-effectiveness of improving cross-section design at a specific location. The four cost components mentioned previously could also be used.

CONCLUSIONS

On the basis of the findings of this study it can be concluded that several cross-section design elements affect safety. Increasing lane width to 3.7–4.0 m (12–13 ft), depending on highway type, is estimated to reduce crash rates for urban freeways and undivided highways. In general, increasing outside shoulder width to 3–3.7 m (10–12 ft) also decreases crash rates. An inside paved shoulder 1.2–1.8 m (4–6 ft) wide is effective in decreasing crashes on rural freeways. Median width is significant in affecting safety for all types of divided highway investigated. Median type affects crash rates: four-lane urban divided highways with flush unpaved (grass) medians are the safest, followed by raised curb medians, then crossover resistance, then TWLTL.

Sections with higher AADT levels are associated with higher crash frequencies for all highway types. Higher AADT levels result in higher crash rates on urban divided highways but in lower crash rates on rural highways and urban divided highways.

Other elements that adversely affect safety include nonuniform cross-section design, lower speed limit and thus lower design speed, the presence of a curb (in most cases), and the presence of interchanges or intersections. There is no significant difference between the safety of sections with an open-graded and a dense-graded friction course.

The models developed in this study could be used in conjunction with other relevant models from the literature to assess the cost-effectiveness of geometric design standards and to assign priority to geometric design improvement projects.

ACKNOWLEDGMENTS

This paper presents partial results of a project sponsored by the Florida Department of Transportation. The authors thank Cheng-Tin Gan, a graduate research assistant in the University of Florida Transportation Research Center, for his help in extracting the data from the computerized data bases.

REFERENCES

1. Miaou, S. P., H. S. Patricia, T. Wright, A. K. Rathi, and S. C. Davis. Relationship between Truck Accidents and Highway Geometric Design: A Poisson Regression Approach. In *Transportation Research Record 1376*, TRB, National Research Council, Washington, D.C., 1992, pp. 10–18.

2. Bowman, B. L., R. L. Vecellio, and J. Miao. *Estimating Vehicle and Pedestrian Accidents Resulting from Different Median Types.* Presented at the 73rd Annual Meeting of the TRB, Washington, D.C., January 1994.
3. Knuiman, M. W., F. M. Council, and D. W. Reinfurt. The Effect of Median Width on Highway Accident Rates. In *Transportation Research Record 1401,* TRB, National Research Council, Washington, D.C., 1993, pp. 70–80.
4. Zegeer, C. V. and J. Deacon. Effect of Lane Width, Shoulder Width, and Shoulder Type on Highway Safety. In *Relationship Between Safety and Key Highway Features, A Synthesis of Prior Research,* TRB, National Research Council, Washington, D.C., 1987, pp. 1–21.
5. Zegeer, C. V., D. W. Reinfurt, J. Hummer, L. Herf, and W. Hunter. Safety Effects of Cross-Section Design for Two-Lane Roads. In *Transportation Research Record 1195,* TRB, National Research Council, Washington, D.C., 1988, pp. 30–32.
6. Squires, C. A., and P. S. Parsonson. Accident Comparison of Raised Median and Two-Way Left-Turn Lane Median Treatments. In *Transportation Research Record 1239,* TRB, National Research Council, Washington, D.C., 1989, pp. 30–40.
7. Hoffman, M. R. Two-Way, Left-Turn Lanes Work! In *Traffic Engineering,* Vol. 44, No. 11, Institute of Transportation Engineers, Washington, D.C., 1974, pp. 24–27.
8. Bowman, B. L., and R. L. Vecellio. *The Effect of Urban/Suburban Median Types on Both Vehicular and Pedestrian Safety.* Presented at the 73rd Annual Meeting of the TRB, Washington, D.C., January 1994.
9. Myers, R. H. *Classical and Modern Regression with Applications.* PWS-KENT Publishing Company, Boston, Mass., 1992.
10. Green, W. H. *LIMDEP Version 6.0 User's Manual and Reference Guide.* Econometric Software, Inc., Bellport, N.Y., 1991.
11. Fletcher, R. *Practical Methods of Optimization.* John Wiley and Sons, New York, N.Y., 1980.
12. Gilchrist, W. *Statistical Modelling.* John Wiley and Sons, Inc., New York, N.Y., 1985.
13. Cameron, A., and P. Trivedi. Regression Based Tests for Overdispersion in the Poisson Model. *Journal of Econometrics,* Vol. 46, 1990, pp. 347–364.

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the Florida Department of Transportation or the U.S. Department of Transportation.

Publication of this paper sponsored by Committee on Operational Effects of Geometrics.