THE SAFETY EFFECTS OF SPEED LIMIT CHANGES: USE OF PANEL MODELS, INCLUDING SPEED, USE, AND DESIGN VARIABLES

Young-Jun Kweon
Post-Doc Research Assistant
Department of Civil Engineering
The University of Texas at Austin
ECJ 6.9, Austin, Texas 78712
Email: yngjnkweon@yahoo.com

Kara M. Kockelman
Clare Boothe Luce Associate Professor of Civil Engineering
The University of Texas at Austin
ECJ 6.9, Austin, Texas 78712,
Tel: (512) 471-4379
FAX: (512) 475-8744
Email: kkockelm@mail.utexas.edu
(Corresponding Author)

The following paper is a pre-print and the final publication can be found in
Transportation Research Record No. 1908: 148-158, 2005.
Presented at the 84th Annual Meeting of Transportation Research Board, January 2005

ABSTRACT
This work estimates the total safety effects of speed limit changes on high-speed roadways using traffic detector data and Highway Safety Information System (HSIS) data from 1993 to 1996. In order to gauge the total effects, this study applies a sequential modeling approach wherein average speed and speed variance models are first estimated, based on roadway design, use and speed limit information. Then, crash counts (of varying severity) are estimated, based on the speed estimates, design, and use variables. The four years of data come from 63,937 "homogeneous" roadway segments along 7 interstates and 143 state highways in Washington State. A random-effects negative binomial model was selected among several alternative panel and non-panel models for count data. Results indicate that the *average* road segment in the data set can be expected to exhibit lower non-fatal crash rates up to a 55 mph (88 km/h) speed limit. In contrast, fatality rates appear unresponsive to speed limit changes. Fatal and non-fatal rates fall for design reasons, including wider shoulders and more gradual curves, which appear to be key design variables. However, fatal and non-fatal rates move differently when traffic levels rise, with non-fatal rates remaining unchanged and fatal rates falling.

Key Words: Speed limit; Traffic Crash; Panel model; Highway Safety Information System (HSIS)

1. INTRODUCTION

Speed limit changes in the U.S. have been made at local and national levels without a thorough understanding of traffic safety consequences. These changes affect speed choices, which affect crash frequency and severity. However, speed choices are only barely recognized in studies aimed at evaluating the effects of speed limit changes on roadway safety. This lack of speed consideration may be attributable to a general lack of extensive speed data.

Some speed limit studies have attempted to control for chosen travel speeds by including coarse speed averages and/or variance values (e.g., Rodriguez 1990 and Lave 1985). They have relied on highly aggregate speed data (e.g., Rodriquez's [1990] data were at the national level, while Lave's [1985] were at the state level). Such aggregation obscures most distributional information about individual traveler speed choices. Golob and Recker (2003) assembled crash and traffic flow data at the 30-second aggregation level on California highways, yet did not include road design variables.

In addition to a lack of chosen speed information, roadway design features are rarely accounted for in speed limit safety studies, although, along with speed, they are recognized as a critical factor in traffic safety. Some speed limit studies (e.g. Lave 1985, Lave and Elias 1994, and Greenstone 2002) have attempted to control for overall design effects by separately analyzing crash counts on different road types (e.g., all interstate, arterial, and collector roads in each state). However, they have failed to include *detailed* geometric information, primarily due to aggregation of non-homogeneous roadway segments. Again, this obscures the true relations that may exist. In order to tightly control for geometric details, use of disaggregate, segment-based roadway data is indispensable. However, crash counts for such short segments are highly discrete and contain many zeros, complicating analysis.

This paper highlights the importance of travel speeds and roadway design, and attempts to quantify the effects of speed limit changes on high-speed roads using data from Washington State. This study also exploits several relatively sophisticated count models for panel data. Existing roadway safety studies employing such models have relied only on the fixed-effect negative binomial model (e.g., Noland 2003 and McCarthy 1999). In contrast, this study considers eight different models and determines the best among these for each of six crash/victim response variables, using a combination of statistics and intuition.

In the following sections, road safety studies associated with speed limits, and with speed and speed variance are summarized. Then, the data used here are described. The methodology section first describes speed models (for average speed and speed variance) based on loop detector data from the northwest region of Washington State, as a function of roadway design, traffic levels and speed limits. It then describes the heart of this research, the crash occurrence models. The results follow next, along with conclusions and suggestions for further research.

2. LITERATURE REVIEW

This section reviews literature emphasizing the safety impacts of speed limit changes as a function of the changes themselves, and as function of changing speed choices.

2.1 The Effects of Speed Limit Change

2.1.1 The Effects of Speed Limit Change on Traffic Safety

Most existing studies have concluded that higher speed limits result in higher crash rates and victim rates (e.g., fatalities per million vehicle miles traveled). However, Forester et al.'s (1984) cost-benefit analysis (of travel delays vs. crash costs) did not support the 1975 imposition of the 55 mph (88 km/h) National Maximum Speed Limit (NMSL). As for the 1985 NMSL relaxation, permitting speed limits of 65 mph (104 km/h) on rural interstates, Greenstone (2002), Ledolter and Chan (1996), Baum et al (1989, 1991), McKnight and Klein (1990), Wagennar et al. (1990), Gallaher et al. (1989), and Upchurch (1989) all found evidence of significant reductions in safety. However, conflicting results also exist for that speed policy. For example, Pant et al. (1992), Sidhu (1990), and Chang and Paniati (1990) found minimal or no change in traffic safety (due to the 1985 relaxation), while Lave and Ellias (1994, 1997), McCarthy (1991), and Lave (1985) concluded there to be safety benefits. Meanwhile, Garber and Graham (1990) argued that the effects vary across states, estimating that some states benefited while others' crash rates increased following a relaxation in their rural interstate speed limits. However, unlike the present paper's methods, none of the above studies has used disaggregate roadway data or panel models.

1995 saw the repeal of the NMSL, with many states choosing to raise speed limits. Relatively few studies have examined the impacts of this latest change in speed limit laws, yet conflicting results still exist. Moore (1999) reported reductions in the U.S.'s crash rates after this repeal, and Najjar et al. (2002) found evidence of no safety changes on Kansas's interstate highways. However, Farmer et al. (1999), Patterson et al. (2002), and Haselton et al. (2002) reported negative safety consequences.

Clearly, the effects of speed limit changes are a controversial topic, with much conflict in past results. Few studies have used models for panel data, though many have used crash data over several years. In fact, only three papers appear to use such models (Houston 1999; McCarthy 1999; and Greenstone 2002), and these all focused on the 1985 NMSL relaxation. Moreover, McCarthy's (1999) study is the only one utilizing a count data model, specifically a fixed-effects negative binomial model, while others used linear fixed-effect models (to model continuous response variables, using aggregate roadway segments). No study has yet applied a random-effects model and provided the results. Houston (1999) and Greenstone (2002) used state-level crash data, while McCarthy (1999) used regional-level data. This work distinguishes itself by applying a variety of count data models (including those permitting random effects) to disaggregate data in an examination of the impacts of the 1996 speed limit changes.

Noland's (2003) negative binomial models accounted for certain design variables at the aggregate level, such as average number of lanes and percentages of miles hosting specific lane widths. However, that level of detail does not compare to that used by Kweon and Kockelman (2004), who relied on segment-based panel data for thousands of roadway segments averaging just 0.1 mile in length. They predicted crash counts, based on a number of design variables (like degree of curvature and vertical curve length); however, their data set contained only five interstate highways and did not consider speed choices. This study overcomes those limitations by including all of Washington's 150 high-speed (50 mph [80 km/h] and over) routes and incorporating estimates of travel speed and speed variations as control variables. It also uses the latest speed limit data available for the State, after correcting the HSIS Washington State data files used in Kweon and Kockelman (2004). (Washington's HSIS speed limit information was up to a year behind the correct values.)

2.1.2 The Effects of Speed Limit Change on Speed Choices

Measures of average speeds and speed variance are often used to describe speed conditions. They also depend, to some extent, on speed limits. Several studies researching the safety effects of speed limit changes have investigated the speed effects as well. As one might expect, it has been widely found that speed limit increases result in increased speeds, (though speed changes are somewhat less than the limit changes themselves). For example, Burritt (1976), Dart (1977) and Forester et al. (1984) found average speeds to fall by 5 to 10 mph following imposition of the 1974 NMSL (which mandated maximum speed limits of 55 mph). As for the 1985 NMSL relaxation, to 65 mph on rural interstate highways, Ossiander and Cummings (2002), Jernigan and Lynn (1991), Freedman and Esterlitz (1990), Brown et al. (1990), and Upchurch (1989) all found increases in average speeds – from 2 to 7 mph.

Results diverge on the topic of speed limit effects on traffic speed variations, among individual drivers. Burritt (1976), Forester et al. (1984), and Rama (1999) estimated reductions in speed variations following lowered speed limits, while Garber and Gadiraju (1992) found variance reductions when their speed limits were differentially raised. Mace and Heckard (1991) estimated increases in speed variance following raised speed limits, while, Ossiander and Cummings (2002), Pfefer et al. (1991), and Brown et al. (1990) found no such changes. Of course, higher speed variations are expected to mean more vehicle interactions, through overtaking and braking, and higher speeds are expected to mean more severe crashes. How these speed variables truly translate to crash rates and crash severity is discussed in the following section.

2.2 Effects of Speed Conditions on Safety

Speed is assumed to be one of the most critical factors affecting crash severity. The laws of physics (i.e., kinetic energy = $0.5\times$ mass $\times$ velocity^2) plainly support this (TRB, 1998). When this conventional knowledge was applied to speed limit studies, the "Speed kills" theory emerged, and has continued to prevail.

In some contrast, however, Lave (1985) raised a "Variance kills" theory, which has been supported by work by Rodriguez (1990) and Reed (2001) and first raised by Solomon (1964) (and replicated by Cerillo (1968)). Solomon's (1964) and Cerillo's (1968) rural road data appear to indicate that crash likelihood increases with an individuals' deviation from a roadway's average speed. Their conclusions were supported by later studies, including those West and Dunn (1971) and Fildes and Lee (1993).

Several researchers (Fowles and Loeb 1989; Levy and Asch 1989; and Snyder 1989) have attempted to refute Lave's variance kills theory, by enhancing his data and model specification. Interestingly, as Lave (1989) notes in his reply/rebuttal, their findings provide evidence for both the variance and speed kills theories. Garber and Ehrhart (2000), Forester et al. (1984), and Zlatoper (1991) also concluded that variance, as well as average speed, contribute to crash frequency. However, none of these studies may reflect correct driving speed behaviors since all use spatially and temporarally aggregated speeds, which are subject to an "ecological fallacy." As Robinson (1950) formally recognize (in his field of sociology), data correlation at aggregate levels can easily differ from that at the individual or disaggregate level.

Rodriquez (1990) and Davis (2002) illustrated how aggregation of speed and crash data invites an "ecological fallacy" in safety study results. Rodriquez (1990) provided empirical evidence for the variance kills theory while assuming a monotonic increase in a driver's likelihood of fatal crash involvement with speed. Davis (2002) demonstrated the fallacy resulting from aggregation of various forms of disaggregate behaviors, using theoretical examples. Therefore, researchers should strive to rely on disaggregate data, wherever possible. Moreover, both speed and speed variance should be controlled for, so that their effects are not confused.

In summary, the existing literature reveals several voids in safety research on the topic of speed limit changes. This work addresses the lack of panel data models for count data (particularly those allowing random-effects), while making speed choices (and speed variations) explicit and relying on detailed, highly disaggregate roadway design and use data. The following sections describe the data, the models, and their results.

3. DATA

3.1 Data Compilation

For safety research, it is useful to have data on crashes (their location, severity, involved vehicles and occupants), road design factors (such as shoulder widths and horizontal curvature), traffic levels, and speed conditions all in the same data set. Such data sets do not exist at raw data levels simply because these data come from totally different sources (e.g., police reports, design plans, and loop detectors). Nevertheless, it is possible to merge three sources into a single database. For example, Garber and Ehrhart (2000) matched Virginia crash data with lane and shoulder widths and traffic detector data. Golob and Recker (2003) matched individual crash records for Southern California freeways with 30-second loop detector data.

Fortunately, the Washington State's Highway Safety Information System (HSIS)[1] database contains crash data and posted speed limits, along with design details. The northwest region of the State has 122 speed trap sites, and the regional office of Washington DOT provided these data on 18 CD-ROMs. Speed and count data originally recorded at 20-second intervals were aggregated automatically by Washington DOT software to 5-minute speed averages. Unfortunately, the original 20-second data could not be procured.

Since 122 detector sites is far less than the 75,909 roadway segments in the HSIS data set, and was not sufficient (in number or in variation across control variables) to provide the desired statistical accuracy sought for this work, data from the detector sites was used to construct models of speed and speed variance. The results of these were used to estimate speed data for all other sites in the HSIS data set, and used in the crash count models. Forester et al. (1984) used a similar sequential approach, first estimating speed conditions, and then assessing the safety effects of the 1975 NMSL's introduction. However, they relied on spatially aggregated demographic data and did not include any geometric design variables.

Monthly speed data were constructed from the 5-minute traffic detector data, and the entire 1993-1996[2] HSIS database was tailored for this work. The speed choice models were constructed using the HSIS's speed limit and road design variables, along with the monthly speed measures (i.e., time-of-day dependent average speeds and speed "variances[3]" [based on 5-minute averages]). Data for the crash occurrence models combined the HSIS data for 75,909 segments over the 4-year period with the speed model estimates of average speeds and speed variance. Further details on these data sets are provided in the following sections.

3.2 Speed-Choice Data

Only 36 of the 122 speed trap detector sites contain a reasonable number of *valid* speed records and could be mapped to distinct road segments in the HSIS data set. Five monthly speed averages and variances for each month were computed for each of the five different times of day (entire day, AM peak, AM off-peak, PM peak, and PM off-peak)[4] at each of those 36 sites. This resulted in five models for average speed and five for speed variance.

The HSIS data were matched to the speed detector site data using route and milepost marker numbers existing in both data sets. The resulting speed-choice data contain roadway design variables, road classification and location indicators, year indicators, and speed limits.

3.3 Crash Occurrence Data

The speed choice models were estimated using the speed-choice data, and estimates of the five speed averages and five speed variances were appended to the HSIS-based data. After removing observations with unreasonable data values (for example, segments with AADTs and lane counts resulting in more than 24,000 vehicles/day/lane and those with vertical grades higher than 12%, the final data included 190,475 observations covering 63,937 segments from 7 interstates and 143 high-speed roadways posted at 50 mph (80 km/h) or higher in Washington State for four years (1993-1996). The data are temporally aggregated (i.e., yearly), but spatially disaggregate (i.e., homogenous in geometric details, with an average section length of just 0.1 mi). Table 1 provides descriptions and basic statistics of these data.

<Table 1 inserted here>

Four categories of variables were included in the empirical analysis. The speed-related variables include speed limit, average speed, speed variance (by time of day), and their squared terms. Road geometry characteristics included horizontal curve length, degree of curvature, vertical curve length and grade, median width, and shoulder width. Segment length, AADT and the number of lanes were used to create average daily vehicle miles traveled (VMT) and AADT per lane variables. VMT enters as a multiplicative exposure term with an exponent constrained to equal 1.0, so that the resulting estimates can be interpreted as crash rates. (Control for traffic intensity, via the AADT per lane variable, negates any need for VMT to serve in that role.)

Road classification and location variables indicate whether a roadway is an interstate or state route, for example, and whether it is rural or urban in nature. Indicators for the years 1994 through 1996 (with 1993 as the reference year) also are used.

As indicated in Table 1, the average segment length is about 0.1 mile (160 meters), permitting close control for geometric design characteristics. The data include rural and urban area roadways with speed limits ranging from 50 to 70 mph (80 to 112 km/h) and AADT (as estimated by the Washington DOT, based on count sampling) from 61 to 215,037 vehicles per day. Geometries range from 0 to 9.55 degree of curvature (i.e., straight to a radius of 600 ft), from -8.9 to 11.4 percent vertical grades, and from 0 to 40-foot total (two-way) shoulder widths. Only the PM peak average speed and speed variance values are displayed in Table 1, because they were determined to be the most appropriate among the 10 speed-related variables (based on goodness-of-fit measures and expectations of estimator signs).

The average numbers of crashes and injured persons per segment per year are much less than 1.0, and counts of four or more are extremely rare, suggesting that count models are clearly needed. In addition, the variances of all dependent variables exceed their means, implying that simple overdispersion exists and may be present even after controlling for explanatory variables. Moreover, the presence of excessive zeros implies that zero-inflated models may be useful. The following section describes both model methodologies, using the speed and crash data sets.

4. METHODOLOGY

4.1 Model Specifications

The sequential modeling approach used here is illustrated in Figure 1. In order to guarantee positive predictions, a log-linear specification was used for the speed choice models. . And, to allow for any heteroskedasticity, White's consistent estimator (White 1980) was used to estimate standard errors of speed choice model parameter estimates.

Eight different models were evaluated for crash counts: the standard Poisson (PO) and negative binomial (NB) models, zero-inflated Poisson (ZIP) and negative binomial (ZINB) models, and fixed- and random-effects Poisson (FEPO/REPO) and negative binomial (FENB/RENB) models. (For statistical details on these models, see Cameron and Trivedi 1998 and/or StataCorp 2003. For applications of these models to crash data, please see, for example, Kweon and Kockelman 2004, Noland 2003, and Shankar et al. 1997.) Recognizing that crash counts do not equal crash victims, six count variables were used as dependent variables, as shown in Table 1: the number of fatalities, injuries, fatal crashes, injury crashes, property damage only (PDO) crashes, and total crashes. Therefore, a total of 48 model formulations were explicitly evaluated: 6 (dependent variables) × 8 (count models).

The PO approach is the simplest and has been popular in the past. A NB specification is more flexible in that it permits data overdispersion as well as an argument for random crash rates, after controlling for explanatory variables. The ZI extension to these models allows for excessive zero observations, by permitting the possibility of segments that never experience crashes. None of these four models exploits the panel data property of this work's data set, however. Therefore, fixed- and random-effects model specifications were formulated for the basic models, resulting in FEPO, REPO, FENB, and RENB models. These accommodate heterogeneity across individual segments as well as over time periods (within a segment) by introducing terms for individual effects, assuming that they take fixed values (FE) or vary randomly across individual segments (RE). The RENB model specification, which was found to perform best among the eight models used here, is presented in Eq. 1.

$$\Pr\left(y_{it} \middle| x_{it}, \delta_i\right) = \frac{e^{-\gamma_{it}} \gamma_{it}^{y_{it}}}{y_{it}!} \text{ where } \gamma_{it} \middle| \delta_i \sim Gamma\left(\lambda_{it}, \delta_i\right) \qquad (1)$$

Here, y_{it} is the number of crashes or victims in the year t at the segment i), x_{it} is a set of explanatory variables including speed limit, and geometric variables such as curve radius, $\lambda_{it} = \exp\left(\mathbf{x}'_{it}\boldsymbol{\beta}\right)$ and δ_i is a dispersion parameter. (StataCorp, 2003) The random-effects negative binomial models allows the dispersion parameter to vary such that $1/\left(1+\delta_i\right) \sim beta\left(p,q\right)$, implying that the RENB model permits dispersion to vary randomly by segment (δ_i). This approach yields the following joint probability over all time periods $(1,2,\ldots,T_i)$ for the segment i:

$$\Pr\left(y_{i1}, \ldots, y_{iT_i} \mid X_i\right) = \frac{\Gamma(p+q)\Gamma\left(p+\sum_t \lambda_{it}\right)\Gamma\left(q+\sum_t y_{it}\right)}{\Gamma(p)\Gamma(q)\Gamma\left(p+q+\sum_t \lambda_{it}+\sum_t y_{it}\right)} \prod_t \frac{\Gamma\left(\lambda_{it}+y_{it}\right)}{\Gamma\left(\lambda_{it}\right)\Gamma\left(y_{it}+1\right)} \qquad (2)$$

The following section describes comparison of these eight model specifications, as well as a method to facilitate interpretation of model estimates.

4.2 Model Comparisons

In order to select a final model for each of the six crash counts modeled here, graphical comparisons based on differences between observed and estimated crash frequencies were performed, along with statistical tests using likelihood ratios (LRs), Vuong's (1989) and Hausman's (1978) tests, and Aikaike and Bayesian Information Criteria (AIC and BIC). However, a graphical comparison provides a general though informal sense of how well a model fits the data at an aggregate level. A more rigorous comparison arises through application of the statistical tests. The LR test is useful in comparing models with and without restrictions; in this study these are the PO versus NB models, and cross-sectional versus panel data models. Vuong's test can be used for model selection in non-nested cases using log-likelihood values, particularly a basic count model versus its ZI counterpart. Moreover, in comparisons between NB and ZINB models, the test results can reveal whether overdispersion in the data is due only to a negative binomial data-generating process or to excessive zero outcomes in addition to a negative binomial process (Shankar et al. 1997). Note that ZINB and NB models are non-nested, so a LR test cannot be used.

Comparisons between FE and RE models can be made using Hausman's test, which examines whether a significant correlation exists between random effects and explanatory variables. In the presence of such correlation, the random-effects slope estimator is inconsistent; consequently, the FE model should be chosen over the RE model. However, this test is valid only under the assumption that both models are correctly specified so that their parameter estimates are consistent. The statistical tests adopted for model comparisons are depicted in Figure 2.

<Figure 2 inserted here>

However, as seen in Figure 2, the statistical tests are not exhaustive in comparing all possible pairs of the models. When a pair of models cannot be compared by a combination of the statistical tests adopted for this study, AIC and BIC values are used to determine a better one. For example, for fatal crash count models, the ZINB model was selected among four cross-sectional specifications (i.e., PO, NB, ZIP, and ZINB) and the RENB model was selected among four panel specifications and their cross-sectional counterparts (i.e., FEPO, REPO, FENB, RENB, PO, and NB). However, there is no statistical test to compare the ZINB and RENB models to authors' best knowledge. Therefore, AIC and BIC values were used for a comparison and the RENB model with AIC = 12,062 and BIC = 12,117 was chosen over the ZINB model with AIC = 12,037 and BIC = 12,091.

4.3 Model Interpretation

Due to the exponential transformation in the count data models used here, to ensure crash rate non-negativity, the effects of the model coefficients are not as obvious as those of an ordinary linear model. For such models, an incidence rate ratio (IRR), is useful to examine (Long, 1997):

$$IRR(x_{j\cdot}) = \frac{\exp\left[\beta_1 x_1 + \Lambda + \beta_j (x_j + 1) + \Lambda + \beta_J x_J\right]}{\exp\left[\beta_1 x_1 + \Lambda + \beta_j x_j + \Lambda + \beta_J x_J\right]} = \exp(\beta_j) \qquad (3)$$

Thus, if β_j = -0.1, the $IRR(x_j)$ = exp(-0.1) = 0.90, so a unit increase in $x_{j\cdot}$ is estimated to reduce the mean crash rate by 10%, assuming all other factors remain constant. This ratio is used in the following discussions of model results.

5. RESULTS AND DISCUSSIONS

5.1 Speed Choice Models

10 speed choice models were estimated and the two PM peak period models (Eqs. 1 and 2) were selected for use in the crash count models. Inclusion of other predicted-speed measures

would result in a very high level of multicollinearity, obscuring results. The PM peak average speed and speed variance estimates were selected due to: (1) the intuitive sign of the speed limit variable's coefficient estimate (i.e., positive) in the average speed model, (2) the statistical significance of their speed limit coefficient (i.e., 0.1 p-value or lower), and their higher goodness of model fit (R-squared terms of 0.484 and 0.265, respectively).

$$AverageSpeedPMPeak = \exp\left\{\begin{array}{l} 0.3822 + 0.0608 \times \text{Speed limit} + 0.000138 \times \text{Horizontal curve length} \\ - 0.1747 \times \text{Degree of curve} - 0.0286 \times \text{Vertical grade} \\ + 0.0000342 \times \text{Vertical curve length} + 0.0223 \times \text{Shoulder width} \\ - 0.0385 \times \text{Number of lanes} - 0.0000196 \times \text{AADT per lane} \\ - 0.0783 \times \text{Indicator for } 50\text{k} \le \text{population} < 100\text{k} \\ + 0.3534 \times \text{Indicator for } 100\text{k} \le \text{population} < 250\text{k} \\ + 0.2816 \times \text{Indicator for year 1994} + 0.0451 \times \text{Indicator for year 1996} \end{array}\right\} \quad (4)$$

(Adjusted R-squared=0.469 and N_{obs} = 437)

$$VarianceSpeedPMPeak = \exp\left\{\begin{array}{l} - 0.4314 + 0.0896 \times \text{Speed limit} + 0.1177 \times \text{Degree of curve} \\ - 0.0743 \times \text{Vertical grade} - 0.00632 \times \text{Median width} \\ + 0.0428 \times \text{Shoulder width} - 0.4153 \times \text{Number of lanes} \\ + 0.3986 \times \text{Indicator for } 50\text{k} \le \text{population} < 100\text{k} \\ + 0.6076 \times \text{Indicator for } 100\text{k} \le \text{population} < 250\text{k} \\ - 0.2613 \times \text{Indicator for year 1996} \end{array}\right\} \quad (5)$$

(Adjusted R-squared=0.250 and N_{obs} = 437)

All estimates in Eqs. 4 and 5 are statistically significant at the 0.05 level; all others were removed via a process of stepwise deletion. As can be inferred from these exponential equations, the effects of speed limit changes on average speed and speed variance measures are estimated to be quite large, perhaps because speed limits proxy for a great many safety features that are unobserved in the data and thus uncontrolled for in the models. For example, while speed limits increase with horizontal curve radius and shoulder width, which are included in the models, they also go up with sight distance, median barrier strength, and pavement condition – all variables that are unobserved. Thus, Eq. 4's coefficient on speed limit is expected to be biased high.

Past studies suggest average speed changes less than speed limit changes (e.g., Ossiander and Cummings 2002; Jernigan and Lynn 1991; Upchurch 1989). These studies only looked at speed changes on roadways whose limits had changed, comparing before and after conditions (assuming all other variables to remain constant), after correcting for any speed changes noted on roadways whose limits had not changed during the same time period. Thus, the approach is more straightforward than the multiple regression models pursued here.

5.2 Crash Occurrence Models

In estimating the crash occurrence models, variables were chosen through an exhaustive search of the data described in Section 3.3. Only statistically significant variables were selected to remain in the final models.

Estimates of final specifications for all 48 models were obtained and one final (best) model for each of the six dependent variables was selected (for a total six final models) using the comparison methods described in Section 4.2.

For all six dependent variables, Vuong's test suggested that ZIP and ZINB models perform much better than the standard PO and NB models. LR tests between pooled-data models and their panel counterparts determined that the panel count models perform better, implying that heterogeneity over time within a segment exists.

Unfortunately, sample sizes for estimation of the FEPO and FENB models were reduced by 65% (in the case of total crashes) to 99% (in the case of fatalities) either because the segment is observed for just one year or exhibits only zero counts over all data periods. The use of conditional maximum likelihood estimation (MLE) for FE count models results in such data reductions (Powers and Xie, 2000).

Owing to the downsizing of the available data for FE models, Hausman's test was not valid although the test statistics could be calculated in some cases. In addition, some final FE models, based on such a small sample, produced unreasonable results including excessive coefficient estimates for variables like shoulder width. Therefore, the FE models were removed from further consideration in the model selection processes.

In cases where the above test statistics could not determine a better model, the two information criteria (AIC and BIC) were used to compare models, along with intuition regarding estimators' signs and magnitudes, and their consistency across dependent variables. The RENB model proved the most effective for all six dependent variables, suggesting that it is a robust form to use in future crash models of panel count data. Table 2 presents the results of the final models.

<Table 2 inserted here>

All variables in Table 2's models are statistically significant at the 0.1 significance level, and most at the 0.01 level. Incorporation of the average daily VMT as an *exposure* variable enables one to view the crash/victim count (i.e., dependent variable) as a rate.

It is worth noting that the effects of the speed limit variables used in the final models need special interpretation, since speed limit also affects the average speed and speed variance variables. Thus, in estimating the effects of a speed limit change, one must also evaluate those effects on speed conditions, to appreciate the total effect.

Noticeable differences were found between the fatal and non-fatal models. All speed-related variables turned out to be statistically insignificant in the fatal models, and only 7 control variables remained in those models – versus 17 to 20 control variables in the non-fatal models. This is probably due to the fact that the vast majority (99%) of observations exhibited zero fatal crashes, so variation in the fatal crash and fatalities counts was extremely limited and dependence on control variables was difficult to distinguish. Figure 3 and Table 3 illustrate the predicted crash rate changes due to changes in speed limits and other variables. These effects are computed using the incident rate ratio (IRR) and are discussed below.

In order to appreciate the total estimated safety effects of speed limit changes, average speed estimates are computed first, then used for estimating of crash rates. Assuming average values for all other control variables (e.g., 2.49 lanes per segment), the total safety effects of a 5 mph (8 km/h) speed limit increase were computed at different base speed limits; these are shown in Figure 3 for easy comparison.

<Figure 3 inserted here>

This unresponsiveness of speed limit variables for fatality and fatal crash rates is counterintuitive and is thought to be due to (1) a lack of variation in fatality counts, due to their relative rarity, and (2) a positive correlation between speed limits and unobserved safety features, such as sight distances and pavement quality (thus biasing the speed limit variable's coefficients towards zero). The second of these two issues may also be at play in biasing speed limit effects downward for other crash rate estimates. One way to address this issue is to only consider crash rates on facilities whose speed limits changed during the study period, and compare before and after crash counts.

For nonfatal rates, as the base speed limit rises to 55 mph (so the new limit rises to 60 mph), the "average" roadway segment in the data set (which happens to have a speed limit of 55 mph) enjoys reductions in nonfatal crash rates. The rate estimates increase with limits higher than 60 mph (base speed limit, or 65 mph new speed limit). This suggests that optimal speed limits, in terms of (non-fatal) crash rates, may be only somewhat higher than those already established, on average, in the State of Washington. However, with more extensive data sets, involving more variation in fatal crash counts, new models of fatalities may recommend lower speed limits. At present, those models are "silent" on the issue of speed limits as studied in this work.

Table 3 provides predicted rate changes for non-speed variables. For example, the effects of shoulder width are quite consistent across the six crash count models. An added 5 ft of shoulder in each direction is estimated to result in a 24 to 27 percent reduction in fatal and nonfatal rates. For other design variables, however, the effects between fatal and nonfatal models do differ. As noted, many were not found to have a statistically significant effect on fatal crash rates, for the data sets and modeling techniques employed here. Yet degree of horizontal curve seems only related to fatal crash occurrence: a 1 degree shaper curve (i.e., 1 degree more subtended by 100 ft of arc) is associated with almost a 10 percent increase in fatal crash rate estimates. Vertical grade, median width, and number of lanes were not found to be statistically significant for any crash counts. However, this does not mean that they are not practically significant. There may not be enough variation in crash counts for these effects to register, even though the sample size is substantial.

<Table 3 inserted here>

Interstate highways are associated with much lower rates in all severity levels than non-interstate highways. In particular, they are estimated to exhibit a 46 percent lower fatal crash rate 22 percent lower injury crash rate, and a 13 percent lower PDO crash rate. Since all interstate segments in the State of Washington also qualify as limited-access facilities, their rates drop even further, by an additional 20 percent for both injury crash rates, 14 percent for PDO crashes, and 18 percent for total crashes. Evidently, the special design features of interstate highways that are not controlled for here (such as pavement quality and clear zone width) are roughly as significant in reducing (non-fatal) crash rates as the removal of at-grade crossings (through use of interchanges and limited access ramps).

Mountainous terrain is expected to have a major effect: almost 50 percent higher injury rates than on level terrain. Driving on rolling terrain is associated with somewhat higher PDO and total crash rates. Roadways in rural areas also are associated with much higher fatal crash rates: 32 to 33 percent higher than urban area highways. This may be due to any number of factors, including longer distances to hospitals, less street lighting, and higher speeds (though the area-type variable was controlled for [via a log-linear specification] in the average speed models). Area types defined by population levels are estimated to affect non-fatal crash rates, with highways in more populated regions exhibiting lower non-fatal rates (perhaps due to better lighting and more barrier controls).

1000 more vehicles per lane per day is estimated to result in a 6 percent reduction in fatal rates (probably due to lower speeds not picked up in the average speed estimate), and no change in nonfatal rates. For reasons that are unobserved/uncontrolled for here (such as weather conditions, vehicle design, seat belt use, and average driver age), 1994 is estimated to be 15 percent less fatal a year (per VMT) than 1993. However, 1995 and 1996 appear to have resulted in significantly more non-fatal crashes (per VMT) than 1993.

6. CONCLUSIONS AND RECOMMENDATIONS

Three of the most important elements of this study are: (1) controls for speed conditions in models of crash counts, (2) use of disaggregate roadway data permitting tight control of design factors, and (3) specification and evaluation of various count models for panel data.

The analysis is based on 5-minute traffic detector data and Highway Safety Information System (HSIS) data for Washington State from 1993 through 1996. Speed conditions (i.e., average speed and speed variance) were estimated using monthly values based on 5-minute detector data, coupled with roadway design and speed limit data. The segment-based panel data contains 190,475 observations stemming from 63,937 segments on 7 interstates and 143 state routes.

Recognizing various crash severities and distinguishing crashes from victims, six different crash/victim counts were modeled; these are the number of fatalities, injuries, fatal crashes, injury crashes, property-damage-only (PDO) crashes, and total crashes on each segment each year. For each of these six counts, eight different count data models were estimated: Poisson, negative binomial, zero-inflated Poisson and negative binomial, fixed-effects and random-effects Poisson, and fixed-effects and random-effects negative binomial models. Average daily VMT served as a proportionality factor, or exposure variable, and traffic intensity (AADT per lane) as a control variable.

Among the eight count model specifications, final models were chosen using a combination of statistical tests, information criteria, and intuition. The random-effects negative binomial (RENB) model was selected for all six crash responses, suggesting that intra-segment heterogeneity over time as well as inter-segment heterogeneity (across segments) contribute to overdispersion in all crash and victim counts and that unobserved factors affecting crash occurrence tend to be distributed randomly across roadway segments. The elimination of the zero-inflated models suggests that a two-state data generating process (where one state is a crash-free state) does not exist in these data.

Based on the final model estimates and incident rate ratios, the safety effects of speed limit changes, geometric factors and other control variables were evaluated (using average values for all variables. Responding to a 5 mph hypothetical uniform increase in speed limits, a road segment with average characteristics (including a 55 mph speed limit) was estimated to experience minimum (non-fatal) crashes at 60 mph. Speed limits were not statistically significant in fatal crash count models, suggesting that the data do not offer sufficient variation in fatal counts and/or that unobserved safety factors positively correlated with speed limits may counteract speed limit effects, thus biasing the associated parameter toward zero. One way to perhaps address this issue, for all crash rate models, is to model only *changes* in crash counts following (secular) changes in speed limits in a way that eliminates/cancels any unobserved effects (which may be correlated with the level of speed limits). However, it is not obvious how this may be done with discrete distributions in such a way that those effects cancel. If counts were normally distributed, the difference in independent normal variables (having conditioned on any unobserved fixed effects) would also be normal with a mean that eliminates the fixed effects. This presents a critical area for future research. The correlation of speed limits with unobserved factors in *any* of the crash count models examined here may be biasing speed limit effects (most likely towards zero, thus understating speed limit effects).

A 5 ft wider shoulder in each direction is estimated to result in a 24 to 27 percent decrease in fatal and nonfatal rates. A 1 degree shaper horizontal curve is predicted to result in 10 percent higher fatal crash and fatality, but was not found to be statistically significant for nonfatal crash rates. Vertical grade, median width, and the number of lanes also lacked statistical significance, though the sample size was extremely large. 1000 more vehicles per lane per day is linked to 6 percent lower fatal rates, but no statistically significant change in nonfatal rates.

Much lower crash rates were estimated to occur on interstate highways, while much *higher* (non-fatal) crash rates are expected in mountainous terrain and much higher fatal crash rates are expected in rural areas.

This work offers new methods, data and results in the areas of crash analysis and speed limit safety impacts. However, improvements can and should be pursued. For example, over-predictions of average speed dependence on speed limits can be resolved by collecting better speed data and by focusing on models of speed changes before and after speed limits change. In addition, there are omitted variables that influence roadway safety and may be correlated with speed limits and other control variables. Such correlations result in biased parameter estimates on variables like speed limit. Weather information, presence of driveways and interchanges, and design speeds also would be useful to have. (For example, Garber and Gadiraju (1990) found that differences between posted speed limit and design speed affect traffic safety and speed variance.) T. In addition, the data used for this study could not distinguish direction of traffic, so all variables involve two directions. Using one-way directional data may permit more precision in certain variables, while doubling data set size. However, it also would make fatal cases scarcer, resulting in less dependent variable variation, which is needed for parameter prediction.

Speed limit decisions represent a major policy action, with serious repercussions for public safety. It would be best to enhance such decisions, using rigorous research. This study aims to assist in this key policy effort.

Acknowledgements

This study was sponsored by the American Association of State Highway and Transportation Officials (AASHTO), in cooperation with the Federal Highway Administration (FHWA), and was conducted in the National Cooperative Highway Research Program (NCHRP) 17-23. The results of this work are preliminary and have not yet been approved by the NCHRP 17-23 project panel for publication as an NCHRP report. Authors would like to thank Mohamedshah Yusuf at the Highway Safety Research Center for providing HSIS data, Jianming Ma at the University of

Texas at Austin for his construction of the speed data, Dr. Daniel Powers at the University of Texas at Austin for his helpful comments, and Annette Perrone for her excellent editing assistance.

Endnotes

[1] Of the nine states currently providing HSIS data, only Illinois, Utah, and Washington states contain curve and grade files (FHWA 2004). And Illinois and Utah data are believed to have some accuracy problems. (Personal communication with an HSIS staff member.)

[2] 1996 data were the most recent set available at the time of this study. More recent data should be available soon. (Personal communication with an HSIS staff member.)

[3] Without individual vehicle speed data, true variances could not be estimated. The count-weighted variance of speed averages, however, are expanded here, to provide estimates of the true variances, since $n\mathrm{V}(X_{avg}) = \mathrm{V}(X)$ if the X_i's are iid during the time periods of interest. Of course, over the course of a day, it is unlikely that the speed distribution does not change, particularly on roadways that congest during certain periods. Thus, the peak- and off-peak variances are expected to be better estimates of true speed variances.

[4] The AM peak is assumed to be from 7:30AM to 8:30AM, AM off-peak from 10:00AM to noon) PM peak from 4:00PM to 6:00PM, and PM off-peak from 9:00PM to 11:00PM.

REFERENCES

Baum, H.M., A.K. Lund, and J.K. Wells. (1989). The mortality consequences of raising the speed limit to 65 mph on rural interstates. *American Journal of Public Health*, 79(10), pp. 1392-1395.

Baum, H.M., J.K. Wells, and A.K. Lund. (1991). The fatality consequences of the 65 mph speed limits, 1989, *Journal of Safety Research*, 22(4), pp. 171-177.

Brown, D.B., S. Maghsoodloo, and M.E. McArdle. (1990). The safety impact of the 65 mph speed limit: a case study using Alabama accident records. *Journal of Safety Research*, 21, pp. 125-139.

Burritt, B.E. (1976). Analysis of the relation of accidents and the 88-km/h (55-mph) speed limit on Arizona highways. *Transportation Research Record*, 609, pp. 34-35.

Cameron, A.C., and P.K. Trivedi. (1998) *Regression Analysis of Count Data*. Cambridge University Press, New York.

Cerillo, J.A. (1968). Interstate system accident research study II. Interim Report II. *Public Road*, 35(5).

Chang, G.-L., and J.F. Paniati. (1990). Effects of 65-mph speed limit on traffic safety. *Journal of Transportation Engineering*, 116(2), pp. 213-226.

Dart, O.K. (1977). Effects of the 88.5-km/h (55-mph) speed limit and its enforcement on traffic speeds and accidents. *Transportation Research Record*, 643, pp. 23-32.

Davis, G.A. (2002). Is the claim that 'variance kills' an ecological fallacy? *Accident Analysis and Prevention*, 34(3), pp. 343-346.

Farmer, C.M., R.A. Retting, and A.K. Lund. (1999). Changes in motor vehicle occupant fatalities after repeal of national maximum speed limit. *Accident Analysis and Prevention*, 31(5), pp. 537-543

Federal Highway Administration (FHWA), Highway Safety Information System. Available at http://www.hsisinfo.org/. Accessed on July 18, 2004.

Fildes, B.N., and S.J. Lee. (1993). *The Speed Review: Road Environment, Behavior, Speed Limits, Enforcement and Crashes*. Report CR 127 (FORS) CR 3/93 (RSB), Federal Office of Road Safety and the Road Safety Bureau, Roads and Traffic Authority of New South Wales.

Forester, T.H., R.F. McNown, and L.D. Singell (1984). A cost-benefit analysis of the 55 MPH speed limit. *Southern Economic Journal*, 50(3), pp. 631-641.

Fowles, R., and P.D. Loeb. (1989). Speeding, coordination and the 55 MPH limit: comment. *The American Economic Review*, 79(4), pp. 916-921.

Freedman, M., and J.R. Esterlitz. (1990) Effects of the 65-mph speed limit on speeds in three states. *Transportation Research Record*, 1281, pp. 52-61.

Fridstrom, L, and S. Ingebrigtsen. (1991). An aggregate accident model based on pooled, regional time-series data. *Accident Analysis and Prevention*, 23(5), pp. 363-378.

Gallaher, M.M., C.M. Sewell, S. Flint, J.L. Herndon, H. Graff, J. Fenner, and H.F. Hull. (1989). Effects of the 65-mph speed limit on rural interstate fatalities in New Mexico. *Journal of the American Medical Association*, 262(16), pp. 2243-2245.

Garber, N.J., and A.A. Ehrhart. (2000). Effects of speed, flow, and geometric characteristics on crash frequency for two-lane highways. *Transportation Research Record*, 1717, pp. 76–83.

Garber, N.J., and R. Gadiraju. (1990). Factors influencing speed variance and its influence on accidents. *Transportation Research Record*, 1213, pp. 64–71.

Garber, N.J., and R. Gadiraju. (1992). Impact of differential speed limits on the speed of traffic and the rate of accidens. *Transportation Research Record*, 1375, pp. 44-52.

Garber, S., and J.D. Graham. (1990). The effects of the new 65 mile-per-hour speed limit on rural highway fatalities: A state-by-state analysis. *Accident Analysis and Prevention*, 22(2), pp. 137-149.

Golob, T.F., and W.W. Recker. (2003). Relationships among urban freeway accidents, traffic flow, weather and lighting conditions. *Journal of Transportation Engineering*, 129, pp. 342–353.

Greene, W.H. (2000). *Econometric Analysis*. 4th Edition, Prentice-Hall, Inc., Upper Saddle River, NJ.

Greenstone, M. (2002). A reexamination of resource allocation responses to the 65-mph speed limit. *Economic Inquiry*, 40(2), pp. 271-278.

Haselton, C.B., A.R. Gibby, and T.C. Ferrara. (2002). Methodologies used to analyze collision experience associated with speed limit changes on selected California highways. *Transportation Research Record*, 1784, pp. 65-72.

Hausman, J. (1978) Specification tests in econometrics. *Econometrica*, 46(6), pp. 1251-1271.

Houston, D.J. (1999). Implications of the 65 mph speed limits for traffic safety. *Evaluation Review*, 23(3), pp. 304-315.

Jernigan, J.D., and C.W. Lynn. (1991). Impact of 65-mph speed limit on Virginia's rural interstate highways through 1989. *Transportation Research Record*, 1318, pp. 14-21.

Kweon, Y.-J., and K. M. Kockelman. (2004). Spatially disaggregate panel models of crash and injury counts: the effect of speed limits and design. *Proceeding of the 83rd Transportation Research Board Annual Meeting*, National Research Council, Washington, D.C., January 2004.

Lave, C. (1985). Speeding, coordination and the 55 MPH limit. *The American Economic Review*, 75(5), pp. 1159-1164.

Lave, C. (1989) Speeding, Coordination and the 55 MPH limit: reply. *The American Economic Review*, 79(4), pp. 926-931.

Lave, C., and P. Elias. (1994). Did the 65 mph speed limit save lives? *Accident Analysis and Prevention*, 26(1), pp. 49-62.

Lave, C., and P. Elias. (1997). Resource allocation in public policy: The effects of the 65-mph speed limit. *Economic Inquiry*, 35, pp. 614-620.

Ledolter, J., and K.S. Chan. (1996). Evaluating the impact of the 65 mph maximum speed limit on Iowa rural interstates. *The American Statistician*, 50(1), pp. 79-85.

Levy, D.T., and P. Asch. (1989). Speeding, coordination and the 55 MPH limit: comment." *The American Economic Review*, 79(4), pp. 913-915.

Long, J.S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. SAGE Publication, London, UK.

Mace, D.J., and R. Heckard. (1991). *Effect of the 65 mph Speed Limit on Travel Speeds and Related Crashes*. Final Report. Last Resource, Inc., Bellefonte, PA.

McCarthy, P.S. (1999). Public policy and highway safety: a city-wide perspective. *Regional Science and Urban Economics*, 29, pp. 231-244.

McKnight, A.J., and T.M. Klein. (1990). Relationship of 65-MPH Limit to Speed and fatal accidents. *Transportation Research Record*, 1281, pp. 71-77.

Moore, S. (1999). Speed doesn't kill: the repeal of the 55-MPH speed limit. *Policy Analysis*, 346, pp. 1-23.

Najjar, Y.M., E.R. Russell, R.W. Stokes, and G. Abu-Lebdeh. (2002). New speed limits on Kansas highways: impact on crashes and fatalities. *Journal of the Transportation Research Forum*, 56(4), pp. 119-147.
Noland, R.B. (2003). Traffic fatalities and injuries: the effect of changes in infrastructure and other trends. *Accident Analysis and Prevention*, 35(4), pp. 599-611.
Ossiander, E.M., and P. Cummings. (2002). Freeway speed limits and traffic fatalities in Washington state. *Accident Analysis and Prevention*, 34(1), pp. 13-18.
Pant, P.D., J.A. Adhami, and J.C. Niehaus. (1992). Effects of the 65-mph speed limit on traffic accidents in Ohio. *Transportation Research Record*, 1375, pp. 53-60.
Patterson, T.L., W.J. Frith, L.J. Povey, and M.D. Keall. (2002). The effect of increasing rural interstate speed limits in the United States. *Traffic Injury Prevention*, 3, pp. 316-320.
Pfefer, R.C., W.W. Stenzel, B.D. Lee. (1991). Safety impact of the 65-mph speed limit: a time series analysis. *Transportation Research Record*, 1318, pp. 22-33.
Powers, D.A., and Y. Xie. (2000). *Statistical Methods for Categorical Data Analysis*. Academic Press, San Diego, CA.
Rama, P (1999). Effects of weather-controlled variable speed limits and warning signs on driver behavior. *Transportation Research Record*, 1689, pp 53-59.
Reed, R. (2001). *Impact of the Speeding Fine Function on Driver Coordination on State Highway*. Maryland State Highway Administration Report, National Transportation Center, Morgan State University.
Robinson, W.S. (1950). Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, 15(3), pp. 351-357.
Rodriguez, R.J. (1990). Speed, speed dispersion, and the highway fatality rate. *Southern Economics Journal*, 57, pp. 349-356.
Shankar, V., J. Milton, and F. Mannering. (1997). Modeling accident frequencies as zero-altered probability processes: an empirical inquiry. *Accident Analysis and Prevention*, 29(6), pp. 829-837.
Sidhu, C.S. (1990). Preliminary assessment of the increased speed limit on rural interstate highways in Illinois. *Transportation Research Record*, 1281, pp. 78-83.
Snyder, D. (1989). Speeding, coordination and the 55 MPH limit: comment. *The American Economic Review*, 79(4), pp. 922-925.
Solomon, D. (1964). *Accidents on Main Rural Highways Related to Speed, Driver, and Vehicle*. Federal Highway Administration, Washington, D.C., July.
StataCorp. (2003). *STATA References: Release 8*. Stata Press, College Station, Texas.
Steven M. S.M. Casey, A.K. Lund. (1992). Changes in speed and speed adaptation following increase in national maximum speed limit. *Journal of Safety Research*, 23, pp. 135-146.
Upchurch, J. (1989). Arizona's experience with the 65-mph speed limit. *Transportation Research Record*, 1244, pp. 1-6.
Vuong, Q.H. (1989). Likelihood ratio test for model selection and non-nested hypothesis. Econometrica, 57, pp. 307-333.
Wagenaar, A.C., F.M. Streff, and R.H. Schultz. (1990). Effects of the 65 mph speed limit on injury morbidity and mortality. *Accident Analysis and Prevention*, 22(6), pp. 571-585.
West, L.B., and J.W. Dunn. (1971). Accidents, speed deviations, and speed limits. *Traffic Engineering*, 41(10).
White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica*, 48, pp. 817-838.

Zlatoper, T.J. (1991). Determinants of motor vehicle deaths in the United States: a cross-sectional analysis. *Accident Analysis and Prevention*, 23(5), pp. 431-36.

List of Tables and Figures

TABLE 1. Variable Definitions and Summary Statistics

Variables	Mean	Std. Dev.	Min	Max
Dependent Variables				
Number of fatalities	0.0040	0.0772	0	11
Number of injuries	0.2224	1.0436	0	52
Number of fatal crashes	0.0033	0.0581	0	2
Number of injury crashes	0.1342	0.5751	0	26
Number of PDO* crashes	0.1509	0.6222	0	29
Number of total crashes	0.2885	1.0676	0	47
Independent Variables				
Speed-Related Variables				
Speed limit (mile/h)	54.512	2.8858	50	70
Average speed PM Peak (mile/h)	42.877	11.931	20	80
Variance speed PM Peak ($mile^2/h^2$)	45.780	21.280	2.53	687.5
Roadway Design Variables				
Segment length (mile)	0.0953	0.1173	0.001	2.550
Horizontal curve length (ft)	290.83	633.68	0	6,358
Degree of curvature (°/100ft)	0.6326	1.2587	0	9.550
Vertical curve length (ft)	419.67	502.55	0	6,000
Vertical grade (%)	0.0577	2.4484	-8.890	11.43
Median width (ft)	8.230	36.461	0	999
Shoulder width (ft)	6.7921	5.8992	0	40
Number of lanes	2.4875	1.1662	1	9
Roadway Classification & Location Variables				
Indicator for interstate highway	0.0694	0.2542	0	1
Indicator for limited access	0.3034	0.4597	0	1
Indicator for principal arterial	0.4366	0.4960	0	1
Indicator for rolling terrain	0.7374	0.4400	0	1
Indicator for mountainous terrain	0.0639	0.2445	0	1
Indicator for rural area	0.8332	0.3728	0	1
Indicator for population < 50k	0.0191	0.1369	0	1
Indicator for 50k≤population<100k	0.0096	0.0973	0	1
Indicator for 100k≤population<250k	0.0061	0.0777	0	1
Indicator for northwest region	0.1120	0.3154	0	1
Indicator for northeast region	0.3343	0.4717	0	1
Indicator for southwest region	0.1713	0.3768	0	1
Indicator for southeast region	0.2072	0.4053	0	1
Traffic Volume & Yearly Indicator Variables				
Annual Average Daily Traffic (AADT) (veh/day)	11,821	25,538	61	215,037
AADT per lane (veh/day/lane)	3,199	4,406	31	23,893
Average daily VMT (veh-mile/day)	861	2,673	0.142	81,687
Indicator for year 1994	0.2544	0.4355	0	1
Indicator for year 1995	0.2468	0.4312	0	1
Indicator for year 1996	0.2581	0.4376	0	1

* Property Damage Only

Note: The number of observations is 190,475 and the number of road segments is 63,937.

TABLE 2. Final Model Results for Six Crash/Victim Counts (Random-Effects Negative Binomial Models)

Dependent Variables	Fatality		Fatal Crash		Injury		Injury Crash		PDO Crash		Total Crash	
Independent Variables	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value
Constant	-9.7450	0.000	-5.8895	0.000	4.7937	0.001	7.0587	0.000	2.9036	0.043	4.3043	0.000
Speed limit	--	--	--	--	-0.4239	0.000	-0.4314	0.000	-0.3022	0.000	-0.3346	0.000
Speed limit squared	--	--	--	--	3.60E-03	0.000	3.67E-03	0.000	2.63E-03	0.000	2.88E-03	0.000
Average speed PM peak squared	--	--	--	--	--	--	--	--	-2.79E-05	0.002	--	--
Variance speed PM peak	--	--	--	--	1.98E-03	0.000	1.79E-03	0.002	--	--	--	--
Variance speed PM peak squared	--	--	--	--	--	--	--	--	6.31E-06	0.003	--	--
Horizontal curve length	--	--	--	--	-6.69E-05	0.000	-8.13E-05	0.000	-5.59E-05	0.000	-9.78E-05	0.000
Degree of curve	0.0917	0.007	0.0927	0.006	--	--	--	--	--	--	0.0462	0.000
Vertical curve length	--	--	--	--	-7.66E-05	0.000	-8.19E-05	0.000	-5.25E-05	0.003	-7.19E-05	0.000
Shoulder width	-0.0314	0.000	-0.0318	0.000	-0.0292	0.000	-0.0303	0.000	-0.0267	0.000	-0.0282	0.000
Indicator for interstate highway	-0.6134	0.000	-0.6180	0.000	-0.2129	0.000	-0.2515	0.000	-0.1343	0.000	-0.1850	0.000
Indicator for limited access	--	--	--	--	-0.2281	0.000	-0.2190	0.000	-0.1498	0.000	-0.1963	0.000
Indicator for principal arterial	--	--	--	--	-0.0720	0.002	-0.0774	0.002	--	--	--	--
Indicator for rolling terrain	--	--	--	--	--	--	--	--	0.0680	0.003	0.0386	0.051
Indicator for mountainous terrain	--	--	--	--	0.3866	0.000	0.3700	0.000	0.6174	0.000	0.4588	0.000
Indicator for rural area	0.2835	0.030	0.2806	0.031	--	--	-0.0790	0.010	--	--	--	--
Indicator for population<50k	--	--	--	--	0.1711	0.000	0.1358	0.002	0.1567	0.000	0.1513	0.000
Indicator for 50k≤population<100k	--	--	--	--	0.3646	0.000	0.2860	0.000	0.3255	0.000	0.3005	0.000
Indicator for 100k≤population<250k	--	--	--	--	0.4377	0.000	0.3985	0.000	0.5702	0.000	0.4568	0.000
Indicator for northwest region	--	--	--	--	-0.1560	0.000	-0.1584	0.000	--	--	-0.0762	0.004
Indicator for northeast region	--	--	--	--	-0.2149	0.000	-0.2367	0.000	-0.1350	0.000	-0.1856	0.000
Indicator for southwest region	--	--	--	--	-0.1365	0.000	-0.1430	0.000	--	--	-0.0503	0.034
Indicator for southeast region	0.2731	0.018	0.2687	0.020	-0.1975	0.000	-0.2074	0.000	0.0963	0.001	-0.0496	0.063
AADT per lane	-6.15E-05	0.000	-6.13E-05	0.000	--	--	--	--	--	--	--	--
Indicator for year 1994	-0.1672	0.081	-0.1663	0.082	--	--	--	--	--	--	--	--
Indicator for year 1995	--	--	--	--	0.0531	0.003	0.0595	0.001	--	--	0.0425	0.001
Indicator for year 1996	--	--	--	--	0.1264	0.000	0.1353	0.000	0.1902	0.000	0.1618	0.000
Ln(average daily VMT)*	1.0000	--	1.0000	--	1.0000	--	1.0000	--	1.0000	--	1.0000	--
*p***	5.9931		271.05		2.5049		12.198		15.032		10.328	
*q***	0.8065		0.8573		1.6540		0.9988		1.2956		1.2631	

*Coefficient was constrained to 1.0 for proportionality (i.e., average daily VMT is modeled as an *exposure* variable).

**Two parameters of Beta distribution, *Beta*(*p*,*q*).

Note: The number of observations is 190,475, and the number of road segments is 63,937.

TABLE 3. Expected Percentage Changes in Crash Rates Responding to Changes in Variables

Explanatory Variables	Change in Variable	Expected Percentage Changes in Crash Rates					
		Fatality	Fatal Crash	Injury	Injury Crash	PDO Crash	Total Crash
Roadway Design Variables							
Horizontal curve length	100 ft	--	--	-0.7%	-0.8%	-0.6%	-1.0%
Degree of curve	1 °/100ft	9.6%	9.7%	--	--	--	4.7%
Vertical curve length	100 ft	--	--	-0.8%	-0.8%	-0.5%	-0.7%
Shoulder width	10 ft	-27.0%	-27.3%	-25.3%	-26.1%	-23.5%	-24.6%
Roadway Classification & Location Variables							
Indicator for interstate highway	Yes	-45.8%	-46.1%	-19.2%	-22.2%	-12.6%	-16.9%
Indicator for limited access	Yes	--	--	-20.4%	-19.7%	-13.9%	-17.8%
Indicator for principal arterial	Yes	--	--	-7.0%	-7.5%	--	--
Indicator for rolling terrain	Yes	--	--	--	--	7.0%	3.9%
Indicator for mountainous terrain	Yes	--	--	47.2%	44.8%	85.4%	58.2%
Indicator for rural area	Yes	32.8%	32.4%	--	-7.6%	--	--
Indicator for population<50k	Yes	--	--	18.7%	14.6%	17.0%	16.3%
Indicator for 50k≤population<100k	Yes	--	--	44.0%	33.1%	38.5%	35.1%
Indicator for 100k≤population<250k	Yes	--	--	54.9%	49.0%	76.9%	57.9%
Indicator for northwest region	Yes	--	--	-14.4%	-14.7%	--	-7.3%
Indicator for northeast region	Yes	--	--	-19.3%	-21.1%	-12.6%	-16.9%
Indicator for southwest region	Yes	--	--	-12.8%	-13.3%	--	-4.9%
Indicator for southeast region	Yes	31.4%	30.8%	-17.9%	-18.7%	10.1%	-4.8%
Traffic Volume & Yearly Indicator Variables							
AADT per lane	1000 veh/day/ln	-6.0%	-5.9%	--	--	--	--
Indicator for year 1994	Yes	-15.4%	-15.3%	--	--	--	--
Indicator for year 1995	Yes	--	--	5.5%	6.1%	--	4.3%
Indicator for year 1996	Yes	--	--	13.5%	14.5%	21.0%	17.6%

Note: Rate percentage changes are based on the incident rate ratio (IRR).

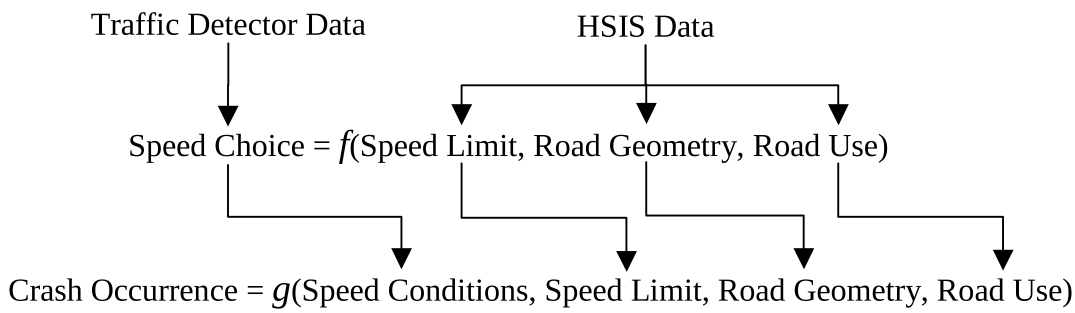

FIGURE 1. Overall Sequential Modeling Approach

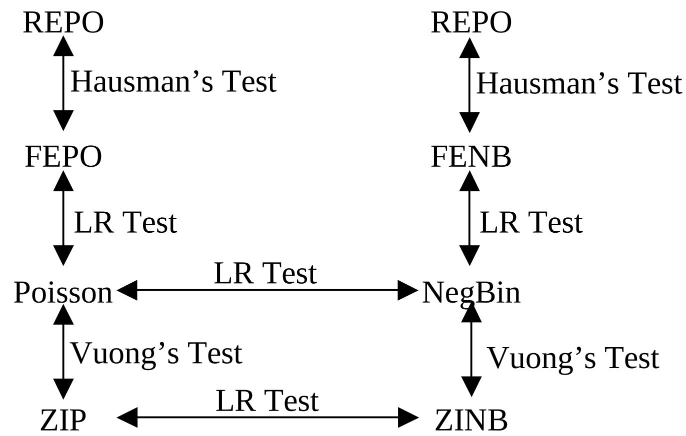

FIGURE 2. Statistical Tests for Model Comparisons

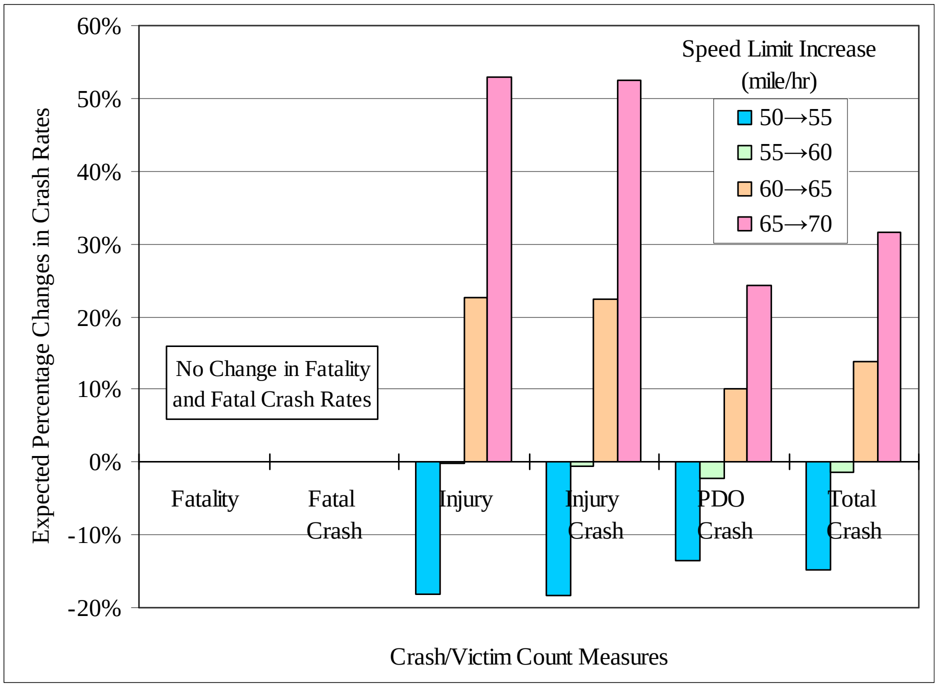

FIGURE 3. Expected Percentage Changes in Crash Rates Responding to Speed Limit Increases
(For a roadway segment having average data characteristics, including a 55 mph speed limit.)