APPENDIX **F**

# AN EMPIRICAL EXAMINATION OF THE RELATIONSHIP BETWEEN SPEED AND ROAD ACCIDENTS

## INTRODUCTION

At the expert panel meeting to develop AMFs for urban/suburban arterials (conducted at the UNC Highway Safety Research Center in July 2005), the findings of Elvik et al. about the relationship between speed and accidents were discussed. (*1*) The following questions were raised:

1. Does the power model hold for North American data?
2. Does the power model appropriately account for a variety of 'before' speed conditions?
3. Are there no conditions (variables) about which we have information other than before and after mean speed that significantly affect the speed-accidents relationship?

To answer these questions, consultants Dr. Ezra Hauer and Dr. James Bonneson, conducted a study:

1. To examine whether it is possible to provide a logical justification to the power model or whether a different model from is indicated and could be derived from 'first principles.'
2. To use Elvik's data in an attempt to answer questions 1, 2 and 3 above.

Both researchers used the same common data base to conduct their own analyses. Preliminary results and insights were frequently exchanged. This mode of research co-operation proved very fruitful by ensuring a certain amount of commonality and the ability to correct missteps and errors while still enabling each author to pursue directions they thought promising. The two approaches are described in separate sections and compared in the last section.

## DATA

The data used for the examination were provided by Elvik et al. A sample of these data is provided in the table below. It shows the headings and the data for three studies, as reported in the literature by the original authors. Column headings 1-30 and 41-51 identify the original data. Column headings 31 to 40 identify additional variables estimated for this examination.

Columns 31 to 40 (in italics) contain the variables that were added to the original data. The notations (*rc*, *rf*, etc.) and the formulae are based on Hauer (1997). (*2*) Using the authors estimated number of accidents expected in the after period, had there been no change in speed ($\pi$ in column 34). This quantity is later referred to as *N*. The expected change in accidents (*Delta* in column 36, later referred to as *dN*) and its standard error (*se(Delta)* in column 37) were also estimated. Similarly, the ratio *Theta* of expected with speed change/expected without speed change (column 38) and its standard error (column 39) were estimated.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| **Study rec no** | **Result rec no** | **Authors** | **Publ year** | **Data country** | **Publ type** | **Study design** | **Main measure** | **Accomp measure** | **Traffic environ** |
| **F1** | **F2** | **F3** | **F4** | **F5** | **F6** | **F7** | **F8** | **F9** | **F10** |
| | | | | | | | | | |
| 1 | 1 | Munden | 1966 | GBR | REP | BAC | POLIS | | URBAN |
| 1 | 2 | Munden | 1966 | GBR | REP | BAC | POLIS | | URBAN |
| 1 | 3 | Munden | 1966 | GBR | REP | BAC | POLIS | | URBAN |

| 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|
| **Veh/users Involved** | **Types of accident** | **Acc/inj severity** | **Accs or victims** | **Speed limit – b** | **Speed Limit - a** | **Mean speed - b** | **Mean speed - a** | **Veh km before (case)** | **Veh km after (case)** |
| **F11** | **F12** | **F13** | **F14** | **F15** | **F16** | **F17** | **F18** | **F19** | **F20** |
| | | | | | | **mph** | **mph** | **O** | **P** |
| ALL | ALL | SER | ACC | 0.0 | 0.0 | 31.5 | 30.0 | 1 | 1.0 |
| ALL | ALL | SLI | ACC | | | 31.5 | 30.0 | | |
| ALL | ALL | SER | ACC | | | 30.0 | 30.8 | | |

| 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|
| **Veh km before (contr)** | **Veh km after (contr)** | **Acc/Vic before** | **Acc/Vic after** | **Comp before** | **Comp after** | **Duration before** | **Duration after** | **Speed change** | **Acc/Vic change** |
| **F21** | **F22** | **F23** | **F24** | **F25** | **F26** | **F27** | **F28** | **F29** | **F30** |
| **Q** | **R** | **K** | **L** | **M** | **N** | | | | |
| 1.0 | 1.0 | 11 | 7 | 82 | 117 | 1 | 1 | 0.951 | 0.446 |
| | | 22 | 13 | 146 | 154 | 1 | 1 | 0.951 | 0.560 |
| | | 11 | 4 | 414 | 418 | 1 | 1 | 1.029 | 0.360 |

| *31* | *32* | *33* | *34* | *35* | *36* | *37* | *38* | *39* | *40* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |
| | | | | | | | | | |
| ***rc=*** | ***rt=*** | ***VAR(rt)/rt^2=*** | ***Pi*** | ***VAR{Pi}*** | ***Delta*** | ***se(Delta)*** | ***Theta*** | ***se{Theta}*** | ***MCF*** |
| ***(N*Q/R)/M*** | ***rc*P/O*** | ***1/M+1/N*** | | | | | | | |
| *1.427* | *1.427* | *0.021* | *15.7* | *27.5* | *8.7* | *5.9* | *0.45* | *0.225* | *3.0* |
| *1.055* | *1.055* | *0.013* | *23.2* | *31.7* | *10.2* | *6.7* | *0.56* | *0.206* | *3.0* |
| *1.010* | *1.010* | *0.005* | *11.1* | *11.8* | *7.1* | *4.0* | *0.36* | *0.212* | *3.0* |

| 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fnc power** | **Est of power** | RTM bias | **Trend bias** | **Volume bias** | **Factor bias** | **FE weight** | **Speed limit - b** | **Speed limit - a** | **Mean speed – b** | **Mean speed - a** |
| **F41** | **F42** | F43 | **F44** | **F45** | **F46** | **F47** | | | **km/h** | **km/h** |
| | | | | | | | | | | |
| 0.446 | 15.968 | YES | NO | NO | NO | 3.93 | | | 50.7 | 48.2 |
| 0.560 | 11.459 | YES | NO | NO | NO | 7.37 | | | 50.7 | 48.2 |
| 0.360 | -35.667 | YES | NO | NO | NO | 2.89 | | | 48.2 | 49.6 |

In some instances, the values of *Theta* (i.e., the computed AMF in column 38) for various studies did not agree with the equivalent value computed by Elvik (shown in column 30). Through correspondence with Elvik et al., many of these discrepancies were eliminated.

Elvik et al. (p.29) categorized each published study they identified in the literature in terms of its study design (see column 7). The seven study design designations they used are listed in Table F-1.

**Table F-1. Study Design Designations.**

| | Designation | Study Design |
|---|---|---|
| 1 | EXP | Randomized Controlled Trial |
| 2 | BAM | Before-After with Matched Comparison Group |
| 3 | BAC | Before-After with Non-Equivalent Comparison Group |
| 4 | BAS | Before-After, NO Comparison Group |
| 5 | CST | Cross-Section |
| 6 | CACO | Case-Control |
| 7 | TI-SE | Time Series Analysis |

A subsequent comparison of study design with the data provided for comparison sites revealed a few discrepancies. In some instances, comparison site data were provided yet the study was classified as a simple before-after study. In some other instances, no comparison site data were provided yet the study was classified as a before-after with comparison group. The design designation of the following studies (as identified in column 1) was modified to reflect the provision of comparison group data.

Study Designation Changed from BAC to BAS:
Study 12 Kemper, Byington, and
Study 48 Andersson (PDO records only).

Study Designation Changed from BAS to BAC:
Study 18 Christensen,
Study 72 Lamm, Psarianos, Mailaender, and
Study 77 Kronberg, Nilsson.

The standard errors *se(Delta)* and *se(Theta)* reflect mainly the number of accidents in the study. They are 'ideal' in the sense that they rely on the assumption that all confounding factors were appropriately accounted for and all functional forms used are the correct ones. Since this is never true, a multiplicative Method Correction Factor (*MCF* in column 40) was used to adjust the standard errors. The magnitude of the *MCF* depends on the study design. Judgement was used to establish the set of *MCF*s listed in **Table F-2**. These *MCF*s are an elaboration of the values used in the forthcoming *Highway Safety Manual*.

**Table F-2. Method Correction Factors.**

| | | MCF |
|---|---|---|
| EXP | B-A traffic correction for both treatment and comparison group | 1.2 |
| | B-A traffic correction only for treatment group | 1.5 |
| | No traffic correction | 1.8 |
| BAM | B-A traffic correction for both treatment and comparison group | 2.0 |
| | B-A traffic correction only for treatment group | 2.4 |
| | No traffic correction | 2.8 |
| BAC | B-A traffic correction for both treatment and comparison group | 2.2 |
| | B-A traffic correction only for treatment group | 2.6 |
| | No traffic correction | 3 |
| BAS | With traffic correction | 4.5 |
| | Without traffic correction | 5 |
| CST | With traffic correction | 5 |

The database thus developed was used by both authors in their modeling efforts. The findings from these efforts are described in sections 3 and 4. The potential for regression-to-the-mean bias was identified by Elvik et al. in the database on a study-by-study basis. The effect of this bias was examined in the context of the regression modeling, as described in subsequent sections of this report.

## MODELING APPROACH 1

### Model Examination

This section examines two models that relate speed to crash frequency. The first model examined is that described by Elvik et al. (2004) as the "power" model. This model relates crash frequency to speed, where the speed variable has an exponent of two or more. The second model is developed by the authors of this paper. It relates crash frequency to the probability of a crash, where crash probability is based on the travel time required for the crash avoidance maneuver.

*Power Model*

The power model developed by Elvik et Al.(2004) is defined as:

$$N = E c_0 v^{\alpha} \qquad \text{... 1}$$

where:

$N$ = crash frequency of specified severity (i.e., PDO, injury, fatal);
$E$ = exposure;
$v$ = mean speed, mph;
$\alpha$ = power term; and
$c_0$ = empirical constant.

Values of $c_0$ and $\alpha$ vary, depending on whether the model is used to estimate PDO, injury, or fatal crash frequency.

Equation 1 could also be restated as:

$$N = E\ P(crash)\ P(severe\ crash \mid crash) \qquad \text{... 2}$$

The rationale for this model is that crash occurrence is related to the distance required to stop, which is a function of the square of speed. If a vehicle is unable to stop prior to reaching a roadway hazard, then a collision will likely occur. From this relationship, it is postulated that crash frequency is proportional to the square of speed (i.e., $P(crash) \propto v^2$). Hence, when Equation 1 is calibrated to property-damage-only (PDO) crash data, the power term is theoretically equal to about 2.0.

Elvik et al. (2004) also rationalize that the kinetic energy involved in a collision is related to the square of speed and that the likelihood of a severe (i.e., injury or fatal) crash is related to the amount of kinetic energy in the collision (i.e., $P(severe\ crash|crash) \propto v^2$). Hence, when Equation 1 is calibrated to severe crash data, the power term α is theoretically equal to 4.0 (= 2.0 + 2.0).

The first derivative of the power model is:

$$\frac{dN}{dv} = \frac{a}{v} N \qquad \text{... 3}$$

In the context of a before-after study, the change in crash frequency $dN$ equals the crash frequency with (i.e., after) treatment $N_w$ minus the crash frequency without (i.e., before) treatment $N_{w/o}$. Similarly, the change in speed $dv$ represents the difference between the mean speed with, and without, treatment. Thus, the derivative in Equation 3 is used in its discrete (as opposed to "continuous") form when applied to before-after data. In recognition of this discrete nature, the value of $N$ on the right side of Equation 3 can be set to equal $N_{w/o}$ and $v$ can be set to equal $v_{w/o}$. Substitution of these two variables yields the following variation of Equation 3:

$$\frac{dN}{dv} = \frac{a}{v_{w/o}} N_{w/o} \qquad \text{... 4}$$

where:

$N_{w/o}$ = crash frequency without treatment to effect a change in speed; and
$v_{w/o}$ = mean speed without treatment, mph.

Equation 4 indicates that the change in crash frequency associated with a change in speed is proportional to the number of crashes before the change and inversely proportional to the speed before the change. The magnitude of the change in crash frequency is also proportional to the power term.

The "percent change ratio" can be defined as the ratio of the percent change in crashes to the percent change in speed. It is computed as:

$$R_c = \frac{dN}{dv}\left(\frac{v_{w/o}}{N_{w/o}}\right) \quad \ldots 5$$
$$= \alpha$$

where:

$R_c$ = percent change ratio.

Equation 5 indicates that the percent change ratio is a constant. The percent change in crash frequency does not depend on the percent change in speed (i.e., that the percent change in crashes is the same irrespective of whether the speed changes from 30 to 31 mph or from 60 to 62 mph).

*Exponential Model*

The exponential model is defined as:

$$\mathbf{N = E\,c_3\left(e^{-MT/\overline{MT}}\right)\left(\frac{1}{1+e^{(c_1-v)/c_2}}\right)} \quad \ldots 6$$

where:

| | |
|---|---|
| $c_3$ = | constant of proportionality; |
| $MT$ = | maneuver time needed to avoid a crash (= $Dc/v$), s; |
| $\overline{MT}$ = | average maneuver time (= $Dc/m$), s; |
| $m$ = | average maneuver speed based on facility design, mph; |
| $c_i$ = | empirical constants, $i$ = 1, 2, 3; and |
| $Dc$ = | maneuver distance needed to avoid a crash, miles. |

The first term in parentheses represents the probability of a crash *P(crash)*. It is based on maneuver time which, in turn, is based on speed and the "critical maneuver distance" *Dc*. This distance could be stopping distance or lane-change distance. It would include the distance traveled during perception-reaction time plus the distance traveled during the avoidance maneuver. With some cancellation, this term can be reduced to $e^{(m/v)}$ , where $m$ is the average maneuver speed. This speed is related to the design speed of the road and represents the speed at which the roadway design can accommodate the critical maneuver with reasonable safety.

The probability of a crash predicted from the power model and the exponential model is shown in Figure F-1. The *P(crash)* term for the exponential model is equal to $c_3\, e^{(m/v)}$. It is shown using a dashed line. The *P(crash)* term for the power model is equal to $c_4\, v^2$ . The dashed trend line has a slight concave shape while the solid line has a convex shape. It should be noted that a concave relationship between speed and crash frequency was developed by Kockelman (2006, Table 4-20) using HSIS data for 3370 miles of interstate and highway in Washington State.

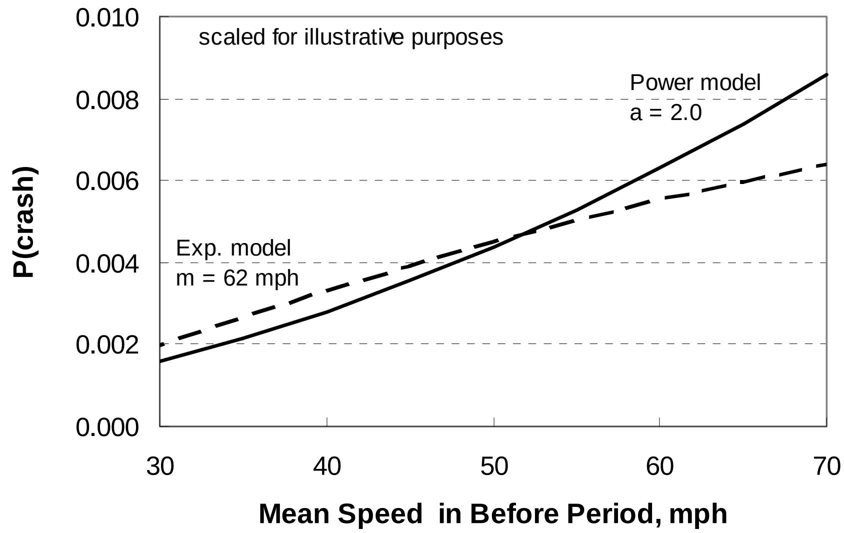


**Figure F-1. Comparison of P(crash) for two models.**

The second term in the parentheses in Equation 6 represents the probability of a severe crash, given that a crash has occurred *P(severe crash | crash)*. This term would not be included if Equation 6 were used to estimate the expected PDO crash frequency. For mathematical convenience, this probability is specified using the logistic function. This probability is shown in Figure 2 using a dashed line. It is compared to the *P(severe crash | crash)* term from the power model (i.e., $P(\text{severe crash} \mid \text{crash}) = c_5\, v^2$ ).

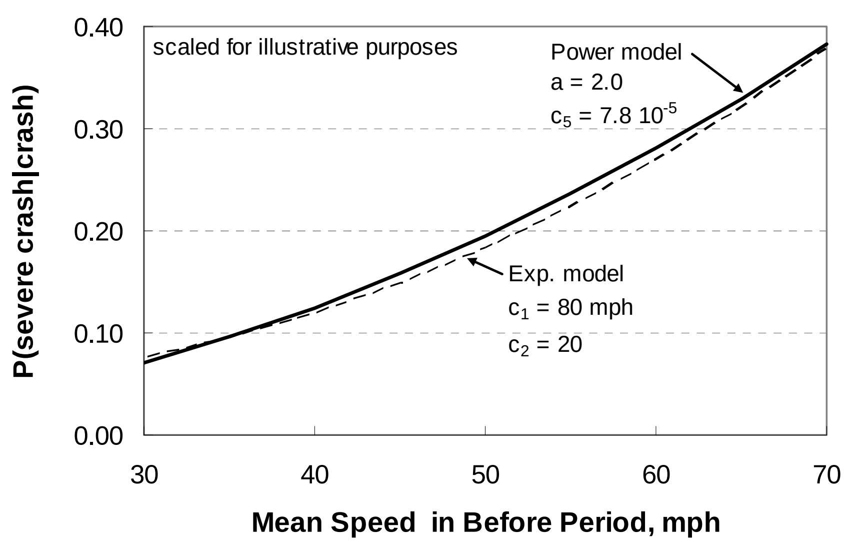


**Figure F-2. Comparison of P(severe crash |crash) for two models.**

The logistic formulation has the logical bounds of 0.0 and 1.0 at very low and high speeds, respectively. The formulation from the power model can yield values greater than 1.0 at exceptionally high speeds. However, over the range of typically encountered speeds, the two probability functions yield effectively equivalent values for any given speed. Therefore, the relationship *P(severe crash | crash)* = $c_5 v^2$ appears sufficiently accurate for the prediction of the probability of a severe crash, given that a crash occurred.

Based on the preceding discussion, the exponential model is revised to the following form:

$$N = E\,c_3\left(e^{-m/v}\right)v^{c_6 I} \qquad \ldots 7$$

where:

$I$ = indicator variable (0.0 when predicting PDO crash frequency, 1.0 when predicting severe crash frequency).

The revised exponential and power models are compared in Figure 3. The exponential model is shown using a dashed line. The upper pair of trend lines illustrates the use of each model to predict PDO crashes. The lower pair of trend lines illustrate the use of each model to predict severe crashes (i.e., injury + fatal). To predict PDO crashes, the *P(severe crash|crash)* term was excluded from each model.

The two models are shown in Figure 3 to have generally similar trends for the range of speeds shown. However, the PDO versions of each model are less similar than the severe crash versions. The PDO crash trend line for the exponential model has a concave shape and the power model has a convex shape. In contrast, the severe crash versions of each model have a fairly similar convex shape. The severe crash trend lines are similar for both models at speeds up to 60 mph. Above 60 mph, the two models diverge slightly with the exponential model having a smaller rate of increase in slope.

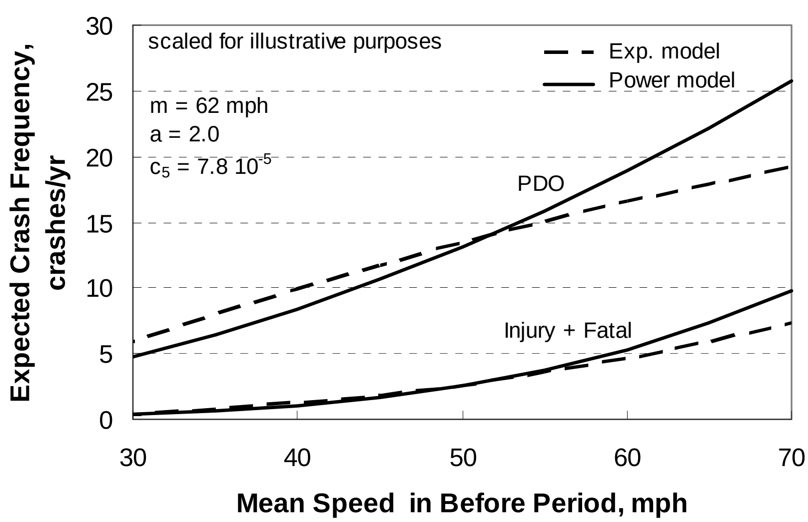


**Figure F-3. Comparison of expected crash frequency from two models.**

The first derivative of the exponential model is:

$$\frac{dN}{dv} = \left( \frac{m}{v^2} + \frac{c_6 I}{v} \right) N \qquad \ldots 8$$

In recognition of the discrete nature of before-after crash data, the value of *N* on the right side of Equation 8 can be set to equal $N_{w/o}$ and *v* can be set to equal $v_{w/o}$. Substitution of these two variables yields the following variation of Equation 8:

$$\frac{dN}{dv} = \left( \frac{m}{v_{w/o}^2} + \frac{c_6 I}{v_{w/o}} \right) N_{w/o} \qquad \ldots 9$$

Equation 9 indicates that the change in crash frequency associated with a change in speed is proportional to the number of crashes before the change and inversely proportional to the speed before the change. The magnitude of the change in crash frequency is also directly proportional to the value of *m* and $c_6$.

The "percent change ratio" can be computed as:

$$\begin{aligned} R_c &= \frac{dN}{dv} \left( \frac{v_{w/o}}{N_{w/o}} \right) \\ &= \left( \frac{m}{v_{w/o}} + c_6 I \right) \end{aligned} \qquad \ldots 10$$

Examination of Equation 10 indicates that the percent change ratio is a function of speed for the exponential model. This relationship suggests that the percent change in crash frequency for a given percent change in speed is larger at lower speeds than it is at higher speeds. This trend holds when the exponential model is applied to the PDO, injury, or fatal crash frequency. It should be noted that if $m = 0.0$, $I = 1.0$, and $c_6 = \alpha$, then Equation 10 yields the percent change ratio for the power model.

*Accident Modification Functions*

The following equation was used to estimate the AMF for both models:

$$AMF = \frac{N_w}{N_{w/o}} \qquad \ldots 11$$

where:

$N_{w/o}$ = crash frequency without treatment; and
$N_w$ = crash frequency with treatment to produce a change in speed.

The crash frequency without treatment $N_{w/o}$ was estimated using the following equation:

$$N_{w/o} = K\, r_d\, r_{tf}\, r_c \qquad \ldots 12$$

where:

$K$ = count of crashes in the before period;
$r_d$ = ratio of the after period duration to the before period duration;
$r_{tf}$ = ratio of the after period traffic volume to the before period traffic volume; and
$r_c$ = ratio of the after period crash frequency at the comparison sites to the before period crash frequency at the comparison sites.

It should be noted that some of the studies included in the database did not include data to compute the traffic volume ratio, in which case this ratio was assumed to equal 1.0. Also, some studies were simple before-after studies that did not include comparison sites, in which case the comparison site ratio was assumed to equal 1.0.

Combining Equation 1 with Equation 11 yields the following AMF for the power model:

$$AMF = \left( \frac{v_w}{v_{w/o}} \right)^{\alpha} \qquad \text{... 13}$$

where:

$v_w$ = mean speed without treatment, mph; and
$v_{w/o}$ = mean speed with treatment applied to change speed, mph.

Similarly, combining Equation 7 with Equation 11 yields the following AMF for the exponential model:

$$AMF = \left( \frac{v_w}{v_{w/o}} \right)^{c_6 I} e^{-m(1/v_w - 1/v_{v/o})} \qquad \text{... 14}$$

## Model Evaluation

The power and exponential models were evaluated using before-after data assembled by Elvik et al. (2004) from 98 studies conducted in 20 countries. One objective of this evaluation was to determine whether the supports the power or exponential model forms (or perhaps a third model form). A second objective was to determine if the trends in the data from U.S. studies were different from those in the data from other countries.

Initially, a qualitative evaluation was undertaken that focused on the percent change in crashes relative to the percent change in speed. Then, a quantitative evaluation was undertaken that focused on the change in crashes associated with a specified change in speed. For this evaluation, the AMF for speed change was related to the data using regression analysis. Finally, some alternative data subsets and model forms are discussed.

### *Qualitative Evaluation*

The qualitative evaluation focused on graphically exploring the relationship between percent change in crash frequency and percent change in speed. An examination of Equation 5 suggested that the two percentages are related by a constant. In contrast, an examination of Equation 10 suggests that the relationship between the two percentages is a function of speed.

To facilitate the graphical examination, the data were sorted to include only those studies that satisfied the following criteria:

1. Crash count in the "before" period of 50 crashes or more.
2. A change in mean speed from before to after period of ±1.0 mph or more.
3. Mean speed in the "before" period of 40 mph or more.

Criterion 1 was intended to provide some stability to the "percent change in crashes" variable used in the graphical examination. Similarly, Criterion 2 was established to provide some stability to the "percent change in speed" variable. The mathematics of one of the statistics being examined (i.e., percent change ratio) caused it to have a high variability when the speed change was small. Hence, to facilitate the graphical examination of trend, data associated with a very small speed change were excluded for convenience. Finally, Criterion 3 was based on a preliminary examination of the correlation between traffic environment and crash change variability. It was found that sites located in residential areas did not portray as well-defined a relationship between percent speed change and percent crash change as did the sites located in urban, rural, and freeway environments.

For the graphical examination, no distinction was made between studies that quantified the change in crash frequency or the change in victims. Also, studies that separately quantified "serious" injury crashes and "slight" injury crashes were combined with the studies that examined "all injury" crashes. Thus, in the following discussion of findings from the qualitative evaluation, the data include both crashes and victims. Also, reference to "injury" crashes includes studies that focused on "all," "serious," or "slight" injuries.

A total of 166 studies satisfied the selection criteria. They represent 50 separate studies and a mix of estimates that focused exclusively on either PDO, injury, or fatal crashes. The percent change in crashes was computed as (AMF - 1) ×100. The relationship found in the data is shown in Figure F-4. The trend in this data confirms that there is a defined relationship between the percent change in speed and percent change in crashes.

The first step in the evaluation was to determine if the relationship between the two percentages shown in Figure 4 is influenced by mean speed, as suggested by the exponential model (see Equation 10). If this secondary influence is not found, then this finding would support the power model form, as suggested by Equation 5.

To facilitate this examination, the data shown in Figure 4 were subset into two speed categories containing an approximately equal number of estimates. One category included estimates where the speed in the "before" condition was between 40 and 60 mph. The other category included estimates where the speed in the "before" condition was between 60 and 75 mph. The relationship between percent change in speed and percent change in crashes for the low-to-moderate speed category is shown in Figure F-5. A best-fit trend line (with a forced intercept at 0.0) is shown in the figure. The slope of the trend line is 2.92.

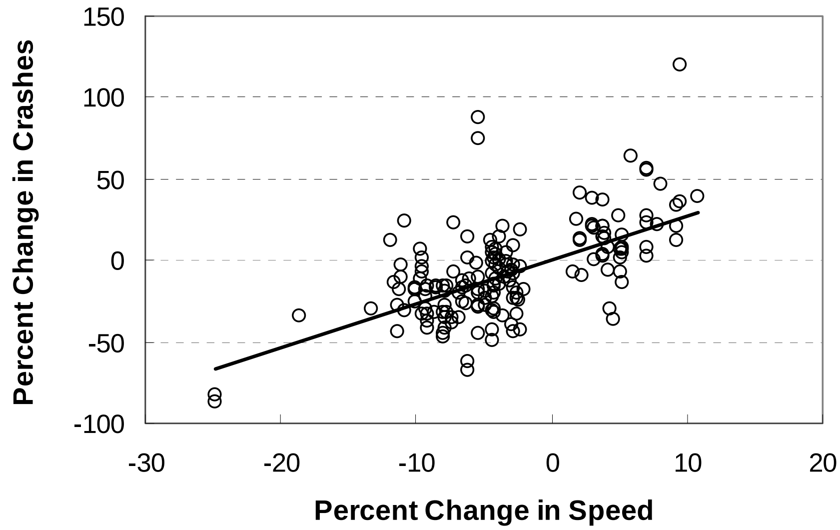


**Figure F-4. Relationship between percent change in speed and percent change in crashes.**

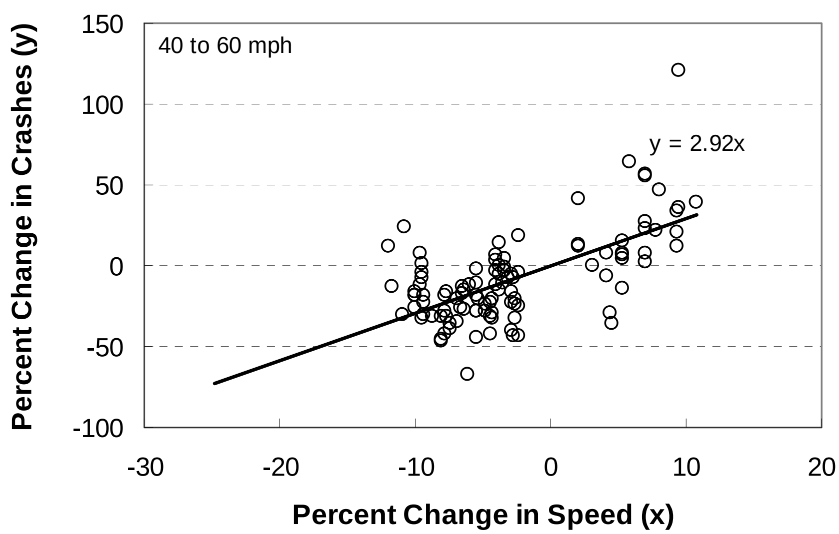


**Figure F-5. Percent change in crashes for sites with low-to-moderate speeds.**

Figure F-6 shows the relationship between the percent change in crashes and percent change in speed for the high-speed category. The slope of the best-fit trend line is 2.39. It is smaller than that for the low-to-moderate speed category and indicates that the relationship between the two percentages is influenced by speed in a manner consistent with the exponential model.

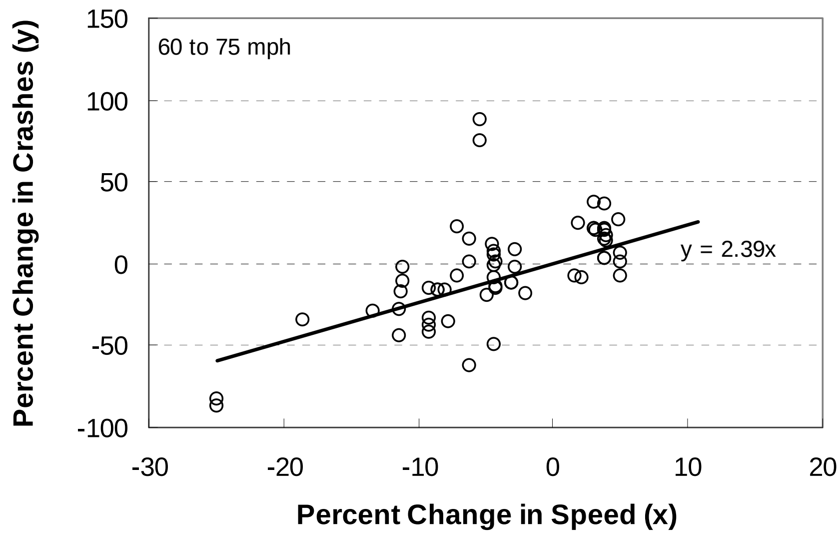


**Figure F-6. Percent change in crashes for sites with high speeds.**

Further exploration of the relationship between percent change in crashes and percent change in speed was conducted to determine if the effect was consistent for each road environment (i.e., freeway, rural highway, and urban street). From this examination, it was found that the pattern shown in Figures F-5 and F-6 was consistent for each road environment. Specifically, the slope of the line is consistently larger for the low-to-moderate-speed sites than it was for the high-speed sites for all three road environments.

The next step in the graphical evaluation was to compare the two models using Equations 13 and 14. For this comparison, the estimated AMF for PDO crashes was used because it facilitated an examination of whether the power or exponential model provided a better fit to the data. The forms of these two models are most distinct when applied to PDO crashes, as noted previously with respect to Figure F-3. The findings from this analysis are shown in Figures F-7 and F-8.

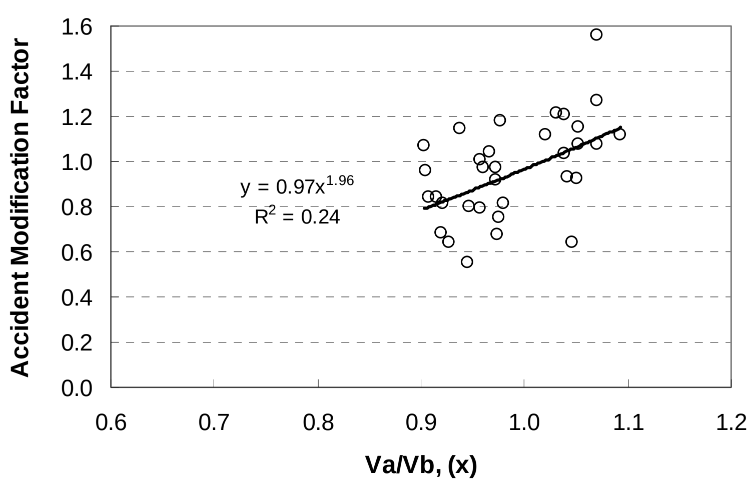


**Figure F-7. Power model AMF for PDO crashes.**

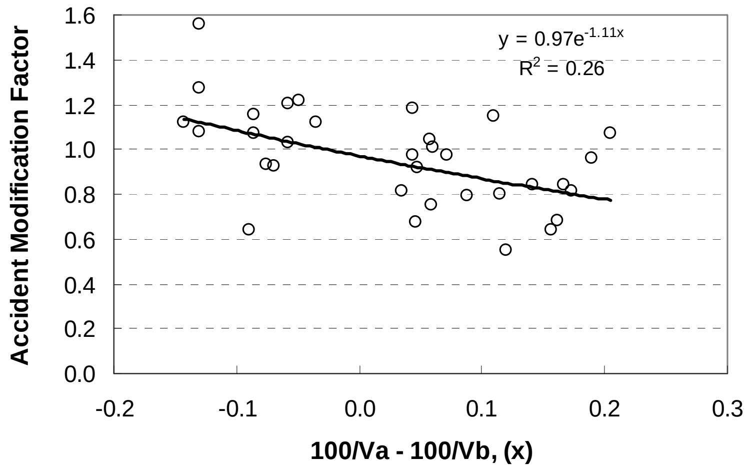


**Figure F-8. Exponential model AMF for PDO crashes.**

The trend line in Figure F-7 shows the best-fit trend line using the power model AMF. The power term is estimated as 1.96; however, the trend line is effectively linear over the range of the independent variable. In fact, a linear model provides a slightly better fit to the data (i.e., $R^2$ of 0.27).

Figure F-8 shows the best-fit trend line using the exponential model AMF. Based on the $R^2$ value shown, the fit of the data in Figure 8 is slightly better than that for the power model AMF. The trends in Figures F-7 and F-8 cast some doubt on whether PDO crash frequency, and the PDO AMF, has a convex shape over the range of typical speeds. In fact, the trends suggest that the concave shape of the exponential model may be more reasonable.

*Quantitative Evaluation*

The quantitative examination considered the relationship between AMF and speed. In recognition of the fact that most AMFs are used with crash frequency, the examination focused on data that reflected a change in crash frequency (as opposed to the number of victims). Also, to ensure the maximum utility of the AMF developed for this research, the analysis focused on records that were based on PDO, injury, or fatal crashes. No data were excluded due solely to their country-of-origin, presence of regression-to-the-mean bias, trend bias, or volume bias, as identified in the database prepared by Elvik et al. (2004). However, as is discussed in a subsequent paragraph, only AMFs based on crash data for urban streets, rural highways, and freeways were ultimately used to calibrate the regression models.

To facilitate the examination of AMFs, a regression modeling analysis was undertaken. It was rationalized that the distribution of the natural log of the AMF (i.e., *LN[AMF]*) is normally distributed. Thus, a least-squares regression model was used for this purpose. Equations 13 and 14 were converted to the following forms for this analysis:

<u>Power model:</u>

$$LN(AMF) = (b_0 + b_1 I_i + b_2 I_f)\, LN(v_w / v_{w/o}) \qquad ... 15$$

Exponential model:

$$LN(AMF) = (b_1 I_i + b_2 I_f)\, LN(v_w / v_{w/o}) - b_3\, (1/v_w - 1/v_{w/o}) \qquad \text{... 16}$$

where:

$I_i$ = indicator variable for injury crashes (= 1.0 if severity type is injury, 0.0 otherwise); and

$I_f$ = indicator variable for fatal crashes (= 1.0 if severity type is fatal, 0.0 otherwise).

In Equation 16, the $b_3$ regression coefficient represents the average critical maneuver speed (previously referred to as variable *m*).

Weighted regression was used because of the differences in variance associated with the AMF estimate for each record. The reciprocal of the variance of the natural log of each AMF estimate was used as the weight for each observation. This variance was estimated as:

$$Var[LN(AMF)] = \frac{Var[AMF]}{AMF^2} \qquad \text{... 17}$$

The variance of the AMF (i.e., *Var[AMF]*) was estimated using the methods described by Hauer (*3*). This variance was subsequently adjusted using a method correction factor to reflect the quality of the various studies used to estimate each of the AMF estimates.

After some preliminary regression analysis, it was found that the studies of residential street treatments yielded AMFs that were highly varied and had trends quite different from those for urban streets, rural highways, and freeways. A closer examination of the residential data indicated that seven of nine studies evaluated the effect of a traffic-calming, or self-enforcing-geometry, technique. These seven studies accounted for 34 of the 37 AMF estimates from residential street studies. It was rationalized that the difference between the two groups of data (i.e., AMFs from residential streets and AMFs from non-residential streets) may be related to the use of "active" versus "passive" methods of speed control. In this context, active methods use road humps, traffic circles, chicanes, etc. to control speed through ride discomfort or oscillatory path changes. Passive methods use heightened enforcement or a change in speed limit to control speed by encouraging drivers to be lawful and, thereby, avoid citation. Passive methods were typically applied to urban streets, rural highways, and freeways. Based on this finding, residential data were excluded from the analysis and the analysis focused on streets, highways, and freeways. The effect of speed change due to traffic calming could not be evaluated using this database.

The preliminary analysis also revealed that AMFs associated with urban street studies exhibited a different value for $b_3$ relative to that for studies of rural highways and freeways. From these preliminary findings, the exponential model's regression equation was revised to the following form:

$$LN(AMF) = (b_1 I_i + b_2 I_f) LN(v_w / v_{w/o}) - (b_3 + b_4 I_u)(1/v_w - 1/v_{w/o}) \quad \text{... 18}$$

where:

$I_u$ = indicator variable for urban records (= 1.0 if facility type is urban street, 0.0 otherwise).

The results of the regression analysis are listed in Tables F-3 and F-4 for the power and exponential models, respectively. The coefficient of determination $R_w^2$ in both tables is computed using the weighted average of the squared error in AMF estimates, where the weight used is the reciprocal of the variance of the AMF. The variance of the natural log of the AMF (from Equation 17) was not used to estimate this coefficient of determination. The same weight used to compute the coefficient of determination was used to compute the weighted standard error. It should be noted that three AMF estimates were determined to be outliers (i.e., records 69, 292, and 103). These three estimates were excluded from the database.

The regression coefficients for the power model are listed in the last three rows of Table F-3 and F-4. The values listed coincide with the power term α for each severity type. The power term for PDO crashes is 0.949. The power term for injury crashes is 2.513 (= 0.949 + 1.564). The power term for fatal crashes is 3.884 (= 0.949 + 2.935). These terms are consistent with the power estimates reported by Elvik (*1*) of 0.73, 2.61 and 3.65 for PDO, injury, and fatal crashes, respectively, based on data from "well-controlled" studies.

The results of the regression modeling for the exponential model are summarized in Table F-4. Comparison of the model statistics in this table with the statistics in Table F-3 suggest that the exponential model provides a slightly better fit to the data. However, it is noted that the exponential model has one more term and thus, would be expected to obtain a lower coefficient of determination. An F-test of the full (exponential) versus partial (power) models indicates that the fourth variable in the exponential model makes a statistically significant improvement in model fit, relative to the power model.

**Table F-3. Calibrated Power Model Statistical Description.**

| Model Statistics | | Value | | |
|---|---|---|---|---|
| | $R_w^2$: | 0.54 (0.45 based on natural log of AMF) | | |
| | Observations: | 323 AMF observations | | |
| | Weighted Standard Error: | ±0.13 | | |
| **Range of Model Variables** | | | | |
| **Variable** | **Variable Name** | **Units** | **Minimum** | **Maximum** |
| $v_w$ | Mean speed with treatment | mph | 16 | 74 |
| $v_{w/o}$ | Mean speed without treatment | mph | 19 | 74 |
| **Calibrated Coefficient Values** | | | | |
| **Variable** | **Definition** | **Value** | **Std. Dev.** | **t-statistic** |
| $b_0$ | PDO crash power term | 0.949 | 0.435 | 2.2 |
| $b_1$ | Incremental injury crash power term | 1.564 | 0.535 | 2.9 |
| $b_2$ | Incremental fatal crash power term | 2.935 | 1.182 | 2.5 |

**Table F-4. Calibrated Exponential Model Statistical Description.**

| Model Statistics | | Value | | |
|---|---|---|---|---|
| | $R_w^2$: | 0.55 (0.47 based on natural log of AMF) | | |
| | Observations: | 323 AMF observations | | |
| | Weighted Standard Error: | ±0.13 | | |
| **Range of Model Variables** | | | | |
| **Variable** | **Variable Name** | **Units** | **Minimum** | **Maximum** |
| $v_w$ | Mean speed with treatment | mph | 16 | 74 |
| $v_{w/o}$ | Mean speed without treatment | mph | 19 | 74 |
| **Calibrated Coefficient Values** | | | | |
| **Variable** | **Definition** | **Value** | **Std. Dev.** | **t-statistic** |
| $b_1$ | Injury crash power term | 1.368 | 0.552 | 2.5 |
| $b_2$ | Fatal crash power term | 2.742 | 1.172 | 2.3 |
| $b_3$ | Average maneuver speed for rural highways and freeways, mph | 70.9 | 25.3 | 2.8 |
| $b_4$ | Incremental average maneuver speed for urban streets, mph | -51.2 | 25.2 | -2.0 |

The power term coefficients listed in Table F-4 for the exponential model have magnitudes that are similar to the incremental power terms in Table F-3. However, the effect of speed on the probability of a severe crash appears to be lower in the exponential model. For example, the term in the power model for a fatal crash is 3.884. In the exponential model, the power term for a fatal crash is 2.742. A similar trend exists for the injury power term in the two models (i.e., 2.513 vs. 1.368).

The coefficient associated with the $b_3$ term in the exponential model suggests that the average maneuver speed for rural highways and freeways is 70.9 mph. The sum of the $b_3$ and $b_4$ coefficients is 19.7. This value indicates that the average maneuver speed for urban streets is 19.7 mph. These two speeds are empirically derived from the regression analysis and are not equal to the design speeds of the respective facility types. Nevertheless, the fact that they are likely to be near to these design speeds is evidence that the theoretical constructs that underlie the exponential model have some validity.

## Alternative AMF Models

Variations of the power and exponential models were examined to determine if additional correlations existed in the AMF data. For this analysis, the data were subset to include only AMFs for which regression-to-the-mean (RTM) bias was determined by Elvik not to exist in the AMF. Also, an additional term was added to the regression model to determine if there is a difference between AMFs derived from data collected in the U.S.A. versus other countries. The alternative model, that coincides with the investigation of country effects is:

Alternative Power Model Regression Equation:

$$LN(AMF) = (b_0 + b_1 I_i + b_2 I_f)\, LN(v_w / v_{w/o}) + b_5 I_{usa} \qquad \ldots 19$$

Alternative Exponential Model Regression Equation:

$$LN(AMF) = (b_1 I_i + b_2 I_f)\, LN(v_w / v_{w/o}) - (b_3 + b_4 I_u)(1/v_w - 1/v_{w/o}) + b_5 I_{usa} \qquad \ldots 20$$

where:

$I_{usa}$ = indicator variable for country (= 1.0 if country is U.S.A., 0.0 otherwise.

The results of the regression analysis are listed in Table F-5. A comparison of the AMFs yielded by the base and "no RTM bias" coefficients for the power model indicates that the AMF for injury crashes may be overestimated by about one percent. In contrast, the AMF for fatal crashes may be underestimated by one percent. A similar trend is noted for the exponential model. These differences were determined to be too small to suggest that only the data without RTM bias should be used for model calibration.

The examination of "Country" indicated that the coefficient for the U.S.A. indicator variable was very small and highly varied. This finding suggests that the trends in the data from other countries cannot be determined to be different from U.S.A. data with any degree of certainty. Thus, the corresponding U.S.A. term is excluded from the two models.

**Table F-5. Examination of Alternative Models.**

| Model | Variable [1] | Base Model Statistics [2] | No RTM Bias | U.S.A vs. Other Countries |
|---|---|---|---|---|
| Power | Observations | 323 | 131 | 323 |
| | $R_w^2$ | 0.54 | 0.53 | 0.54 |
| | $b_0$ | <u>0.949</u> | 0.755 | <u>1.104</u> |
| | $b_1$ | <u>1.564</u> | <u>1.893</u> | <u>1.438</u> |
| | $b_2$ | <u>2.935</u> | <u>2.995</u> | <u>3.029</u> |
| | $b_5$ | -- | -- | -0.016 |
| Exponential | Observations | 323 | 131 | 323 |
| | $R_w^2$ | 0.55 | 0.57 | 0.56 |
| | $b_1$ | <u>1.368</u> | <u>1.774</u> | <u>1.255</u> |
| | $b_2$ | <u>2.742</u> | <u>2.768</u> | <u>2.888</u> |
| | $b_3$ | <u>70.9</u> | <u>61.7</u> | <u>80.0</u> |
| | $b_4$ | <u>-51.2</u> | -60.0 | <u>-57.6</u> |
| | $b_5$ | -- | -- | -0.019 |

Notes:
1 - Variable $b_5$ was added to the regression equations (i.e., Equations 15 and 18) to explore U.S.A crash trends.
2 - Underlined coefficients are statistically significant at a 95 percent level of confidence.

*Sensitivity Analysis*

The AMFs predicted by the calibrated models were compared to determine if the trends were reasonable. For this analysis, a 10 percent reduction in speed was assumed. The results are shown in Figures F-9, F-10, and F-11 for PDO, injury, and fatal crashes, respectively.

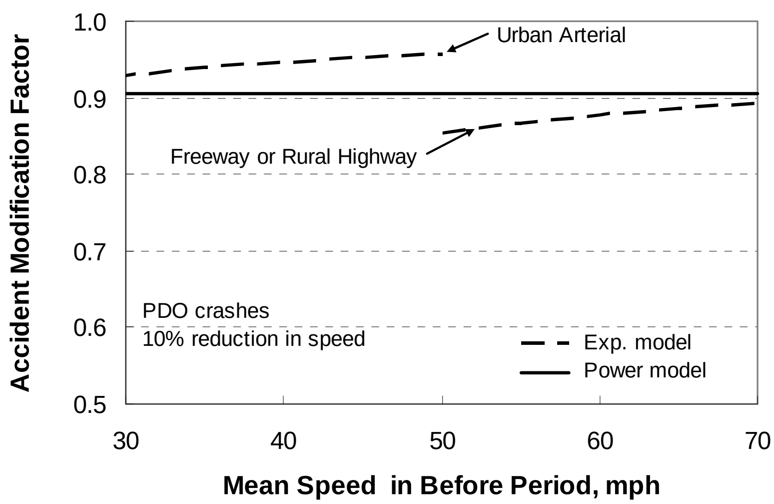


**Figure F-9. AMF for PDO crashes given 10 percent reduction in speed.**

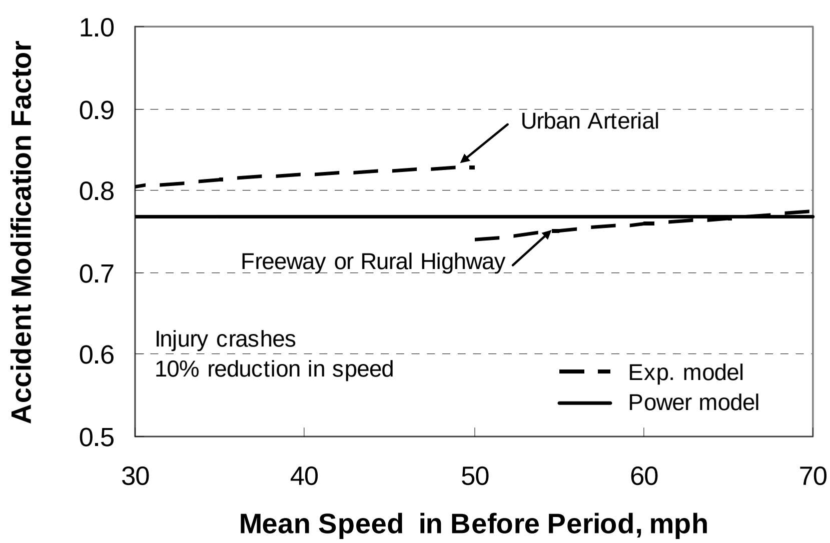


**Figure F-10. AMF for injury crashes given 10 percent reduction in speed.**

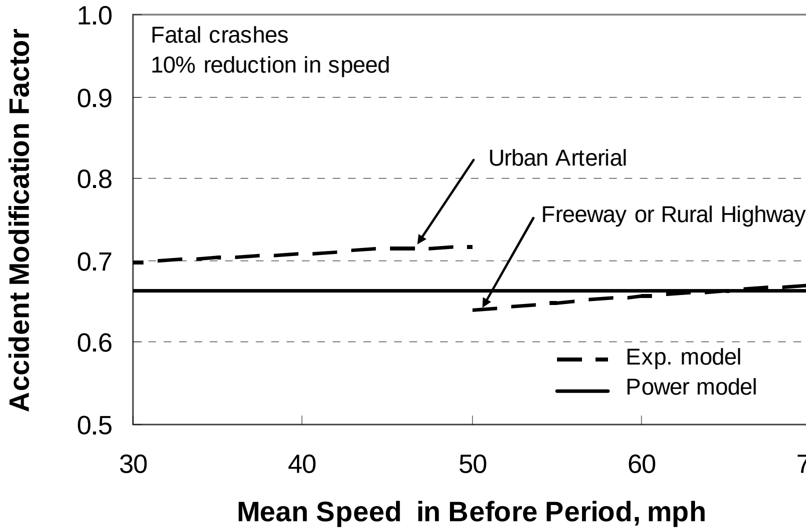


**Figure F-11. AMF for fatal crashes given 10 percent reduction in speed.**

The trends in Figures F-9, F-10, and F-11 show the two models' sensitivity to speed and environment (i.e., urban versus freeway or rural highway). In fact, the AMF value predicted by the power model is not sensitive to speed or environment. It predicts a constant AMF value for each crash severity, regardless of speed or environment. The AMF value predicted by the exponential model tends to increase with increasing speed. This difference between the two models can be illustrated using the example calculation provided in Table F-6.

**Table F-6. Example Calculation Using the Power and Exponential Models.**

| Model | Mean Speed in the Before Period, mph | AMF for PDO Crashes [1] | Before PDO Crash Frequency [2] | After PDO Crash Frequency | PDO Crash Reduction |
|---|---|---|---|---|---|
| Power | 50 | 0.90 | 264 | 238 | 26 |
| | 55 | 0.90 | 289 | 260 | 29 |
| | 60 | 0.90 | 314 | 283 | 31 |
| | 65 | 0.90 | 338 | 304 | 34 |
| | 70 | 0.90 | 363 | 327 | 36 |
| Exponential | 50 | 0.85 | 242 | 206 | 36 |
| | 55 | 0.87 | 276 | 240 | 36 |
| | 60 | 0.88 | 307 | 270 | 37 |
| | 65 | 0.89 | 336 | 299 | 37 |
| | 70 | 0.89 | 363 | 323 | 40 |

Notes:
1 - AMFs obtained from the calibrated power and exponential models for a specified 10 percent speed reduction and freeway or rural highway environment.
2 - Equations 1 and 7 were calibrated to yield 363 crashes for a 70 mph mean speed in the before period.

The AMFs in column 3 of Table F-6 are based on the trends shown in Figure F-9 for the freeway or rural highway environment. Equations 1 and 7 were calibrated to yield 363 crashes in the before period for a 70 mph speed using the power and exponential models, respectively. The trends in column 4 of the table indicate that the exponential model indicates a more rapid rate of decrease in crash frequency with decreasing speed in the before period. This, at 50 mph, the exponential model estimates an expected before crash frequency of 242 crashes and the power model estimates an expected 264 crashes.

As indicated in column 3 of Table F-6, the exponential model and power models have about the same AMF for 70 mph (i.e., 0.89 vs. 0.90) and a similar reduction in crash frequency (i.e, 40 vs. 36). However, at 50 mph, the exponential model predicts an AMF value of 0.85 while the power model predicts an AMF of 0.90. These AMFs translate into a reduction of 36 crashes based on the exponential model and a reduction of 26 crashes for the power model. The estimated reduction in crashes based on the exponential model is more nearly constant over the range of speeds (i.e., 36 to 40 crashes) evaluated than the power model (i.e., 26 to 36 crashes).

The information in Figures F-9, F-10, and F-11 also indicate that speed change has less influence on urban street crash frequency that it has on freeway or rural highway crash frequency. This trend is likely due to the busier, more complicated environment of the urban street, relative to the freeway or rural highway. There are likely many factors that influence crash frequency on urban streets (e.g., significant driveway density, frequent turn movements, etc.) such that the isolated effect of a change in speed limit is moderated by these other influences.

### Summary

Two models are examined. The trends in Figures F-5 and F-6 suggest that the percent change ratio (i.e., percent change in crashes to percent change in speed) is a function of speed. This finding supports the exponential model formulation. An analysis of the AMFs predicted by each model provides evidence that the exponential model is slightly more accurate than the power model. There does not appear to be any difference in AMF values based on data collected in the U.S.A. versus data collected outside the U.S.A.

## MODELING APPROACH 2

This section consists of four parts. The first part provides some theoretical consideration which may shape the direction of the analysis. This is followed by examining what can be gleaned from two-dimensional exploratory representations of the data. The central section will be devoted to modeling, and is followed by a summary.

### Some Theoretical Considerations

The main task is to suggest a model that makes the number of accidents (N) a function of the mean speed (v). Elvik et al. suggested the relationship in equation 21. I begin (section 4.1.1) by asking what function N of v does this imply. The next subsection exploits the results of 4.1.1 to suggest directions for more general approached to the modeling of N(v). The last section is an attempt to use some logical conditions to identify the kind of function that may be chosen.

*What function is implied by the 'power model'*

One of the concerns about the power model was that it appears to be too simple and that the percent change in crashes does not depend on the speed (i.e., that the percent change in crashes is the same irrespective of whether speed changes from 30 to 31 mph or from 60 to 62 mph). The purpose here is to establish what functional relationship between the number of crashes and speed is implied by the power model.

Let 'N' be the number of accidents of a given severity for a set of facilities, 'v' the mean speed on this set and 'α' a parameter. The power model is then

$$\frac{N(v+\Delta v)}{N(v)} = \left(\frac{v+\Delta v}{v}\right)^{\alpha} \qquad \textbf{... 21}$$

When Δv is small compared to v,

$$N(v+\Delta v) \cong N(v) + N'(v)\Delta v \cong N(v)\left(1+\frac{\Delta v}{v}\right)^{\alpha} \cong N(v)(1+\alpha\frac{\Delta v}{v})$$

from which

$$N'(v) = \frac{\alpha}{v}N(v) \qquad \textbf{... 22}$$

or, in differentials,

$$\frac{dN}{N(v)} = \alpha\frac{dv}{v}$$

The solution of this differential equation is

$$N(v) \propto v^{\alpha} \qquad \textbf{... 23}$$

Obviously, for a specific set of facilities (e.g., rural interstates in Colorado), N(v) is a function of the length of the system, of the traffic it serves and of many other variables. The influence of all these variables would have to be represented in equation 23 within the coefficient of proportionality that links N(v) and $v^{\alpha}$. This coefficient of proportionality can be a function of any variable but it must not be a function of v. Thus, if for the set of facilities for which equation 23 holds N(v*) is the count of accidents when v=v*, the coefficient of proportionality in equation 23 must be $N(v^*)/(v^*)^{\alpha}$. Using this coefficient of proportionality, equation 23 can be written as

$$N(v) = \frac{N(v^*)}{(v^*)^{\alpha}}v^{\alpha} \qquad \textbf{... 24}$$

It follows that the power model featuring only the ratios of accident counts and speeds implies the specific functional relationship in equation 24. (Conversely, the function given by equation 24 implies the power model in equation 21).

Some Conclusions

Equation 24 is not much different from the starting point in equation 21. It might seem that the entire argument led, predictably, back to its origin, achieving no new insight. It is therefore useful to state what insight was gained and what guidance for modeling can be stated.

The first conclusion is that while the power model in equation 21 is written in terms of increments, it logically implies the function in equation 24 that links N and v. That is, the absence of functional dependence on v originally thought to be disturbing is a matter of representation, not of substance; when the power model is written as in equation 24, N and the change in N depend functionally on v and the change in v. How N depends v when the power model holds is shown in Figure F-12 (different coefficients of proportionality are used for exponents $\alpha$=2, 3 and 4).

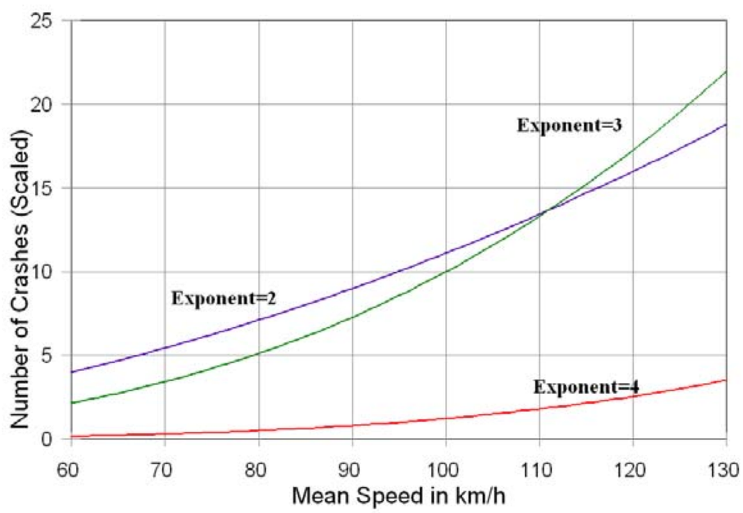


**Figure F-12. How N depends on v in the 'power model'.**

The second insight comes from the fact that the data are in the form $N_i$, $\Delta N_i$, $v_i$ , $\Delta v_i$, where i=1, 2, 3 ... refers to different sets of facilities. Because of the presence of the $\Delta N$ and $\Delta v$ terms it might be natural for modeling to proceed in two steps. The first step is to estimate the parameters of a differential equation based on the data and the seconds step is to integrate it thereby establishing the sought functional relationship between N and v. Thus, e.g., if data are arranged in pairs $\{\Delta N_i/N_i, \Delta v_i/v_i\}$ one can plot $\Delta N_i/N_i$ as the ordinate and $\Delta v_i/v_i$ on the abscissa for i=1, 2, 3,...and check whether the straight-line-through-origin relationship (with a slope $\alpha$) is reasonable choice for this data. If it is, then there is empirical support for the differential equation 22, one can estimate $\alpha$ by an appropriate statistical method and, integrating equation 22, one obtains equation 24.

The third conclusion is that the differential equation embodied in equation 22 (which implies the power model and is being implied by it) is a very specific choice of representation for the available data. Thus, e.g., if instead of using separately the variables v and Δv the modeller elected to represent their influence using one combined variable that is a function of Δv and v. Furthermore of all possible functions of v and Δv [e.g., as $v^{\alpha}\Delta v^{\beta}$, $\ln v^{\alpha}$ arctan(Δv)] the modeller chose the specific abridgment Δv/ v. Had different choices been made, a different model would have been obtained. Thus, e.g., one could start generally thinking that ΔN=f(N, v, Δv) and searching the data for patterns to hint at the proper form of the three-dimensional function f. Alternatively, one could reduce generality and assume that $\Delta N=N^{\beta}\times g(v, \Delta v)$ and thereby narrowing exploration to two-dimensional functions g(v, Δv) and so on. An added advantage of modeling ΔN as a function of N, v and Δv is that the result is easily written in AMF form since AMF=(N+ΔN)/N=1+ΔN/N=1+f(N,v,Δv)/N.

The question is whether the power model (written either as equation 21 or equation 24) is the most sensible choice in view of the data. Of course, the data points $\Delta N_i/N_i$ , $\Delta v_i/v_i$ may not indicate straight line fit. Perhaps a different relationship fit the data better. To illustrate, suppose, e.g., that $dN/N=(\beta_0+\beta_1 v)dv$ is a reasonable representation of the $\{N_i, \Delta N_i, v_i, \Delta v_i, i=1, 2, 3 ...\}$ data. Then, we would estimate (presumably by weighed least squares) the parameters $\beta_0$, $\beta_1$ and use these in the solution of the differential equation which is:

$$\mathbf{N(v) = Ce^{\beta_0 v+\frac{\beta_1}{2}v^2}} \quad \text{... 25}$$

Alternatively, if $dN/N=dv/(\beta_0+\beta_1 v)$ represents the data then $\mathbf{N(v) = Ce^{\frac{1}{\beta_1}\ln(\beta_0+\beta_1 v)}}$ and so on.

*Possible forms for f(v)*

Both equations 23 and 25 imply that the number of accidents of given severity increases exponentially with v. This is perhaps a good approximation for the range of v for which we have data. Nevertheless, one can speculate that such an increase might not continue indefinitely; that at very high speeds the increase will begin to taper off just as the probability to die in a crash tapers off at very high speeds. To take this into consideration, rather than choosing a function by only examining the ΔN/Δv versus v data plot, one might begin by restricting the choice of function (and its derivative) to the family of S-shaped functions of v. Thus, e.g., one might assume that N(v) is similar to a Gamma probability distribution function for which the derivative is:

$$\mathbf{\frac{dN}{dv} = \beta_0(\beta_1 v)^{\beta_2} e^{-\beta_1 v}} \quad \text{... 26}$$

In fact, any function of v the derivative of which is initialing increasing reaches a peak and then decreases, such one of those shown in Figure F-13, could be chosen.

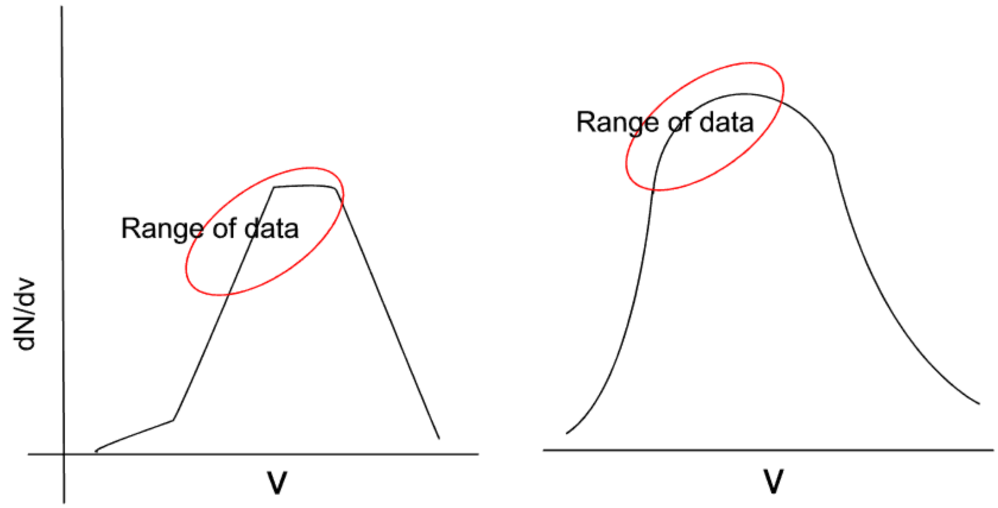


**Figure F-13. Probability Distribution Functions.**

These considerations will guide our analysis of the data.

## EXPLORATORY ANALYSIS

It was previously concluded that considering the form of the data (as $N_i$, $\Delta N_i$, $v_i$, $\Delta v_i$), it is natural to examine differential relationships such as $dN/dv=N^{\beta}\times f(v)$.

### *The Relationship between dN/dv and N*

In Figure F-14 (0<Before Accidents<2000) and Figure F-15 (0<Before Accidents<15,000) I show estimates of dN/dv and against the count N of fatal accidents and fatalities. The question is whether dN/dv seems proportional to $N^{\beta}$ and what β might be. The figures are in two dimensions and f(v) is not represented. It is therefore possible that, if there is some systematic relationship between N and f(v), the visual impression are misleading.

It appears that for fatal accidents and fatalities the relationship is one of proportionality (i.e., that β=1). There are a few data points for which the dN/dv is below 0. This occurs only for studies in which N and dN are small and can be ascribed to randomness inherent in small dN counts.

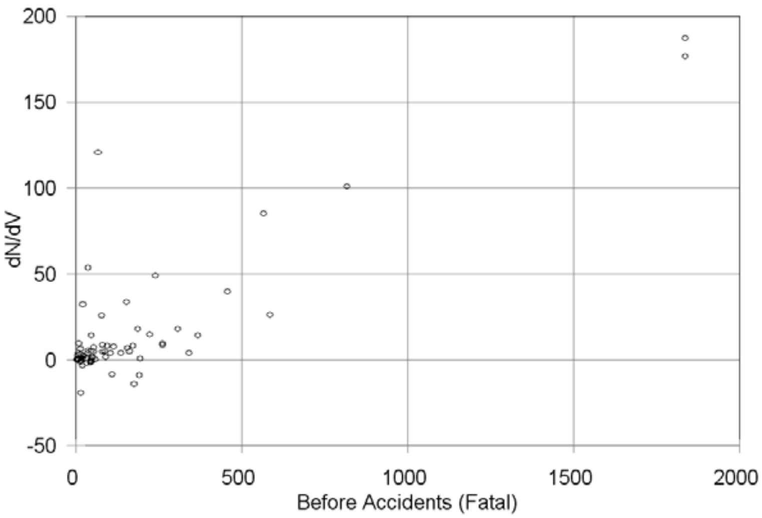


**Figure F-14. dN/dv against the count of 'before' fatal accidents or fatalities if<2000.**

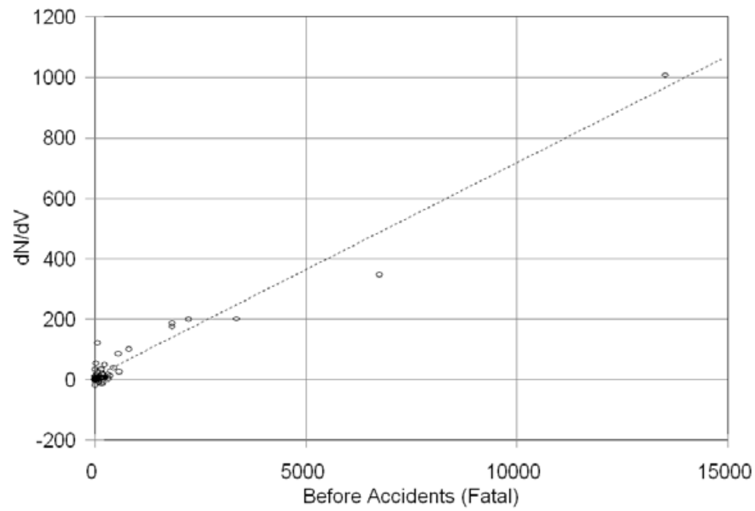


**Figure F-15. dN/dv against the count of fatal accidents or fatalities.**

There is a difference between the count of fatal accidents and the count of fatalities. To keep analysis free of such ambiguities and to retain the ability to compare results with those by Elvik et al., modeling will be based only on data about the count of fatal accidents; studies the results of which are given in fatalities (number of victims) will not be used here. In addition, in Figure F-14 and Figure F-15 the abscissa is in terms of the number of fatal accidents during the 'before' period. Modeling will be based on the estimated number of accidents that would be expected in the after period had there been no change in mean speed (denoted either as $\pi$ or as N). The representation of dN/dv against N for fatal accidents is in Figure F-16. Here too one can imagine a rising straight line through the origin as fitting the data. Was one to include only data points from USA studies, the figure would be a bit sparser but very similar.

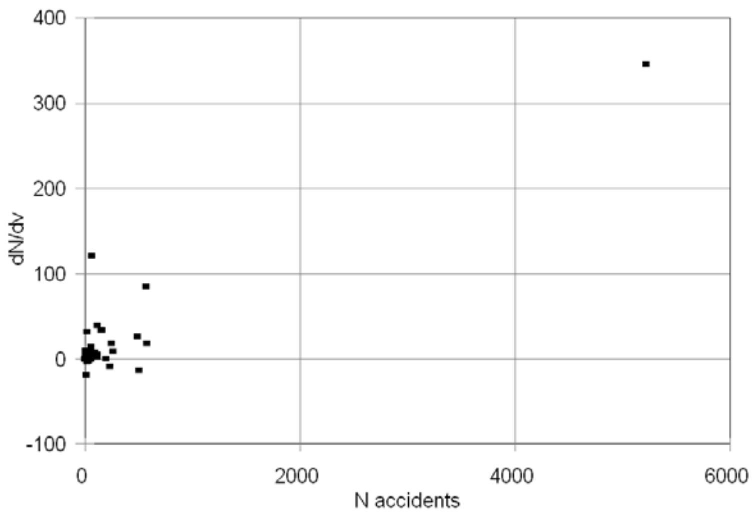


**Figure F-16. How dN/dv depends on N, Fatal Accidents**

For Injury & Serious accidents and victims the relationship between estimates of dN/dv and the count before is in Figure F-17 (0 to 10,000) and Figure F-18 (0 to 400,000). As for fatal accidents, the relationship seems to be one of proportionality.

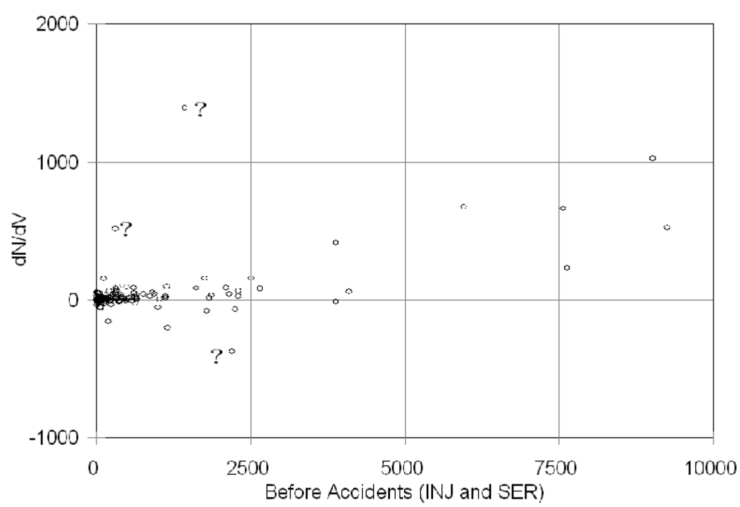


**Figure F-17. How dN/dv depends on Before Injury & Serious Accidents and Victims when <10,000.**

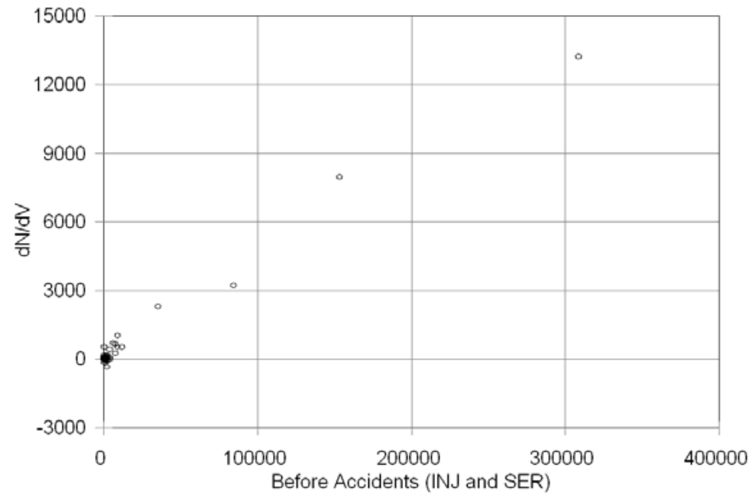


**Figure F-18. How dN/dv depends on all Injury & Serious Accidents and Victims.**

In addition to a few data points for which the dN/dv<0 due to randomness, there are in Figure F-17 a few apparent outliers noted by question marks. The results of outlier investigation are in TableF-7.

**TableF-7. Outliers (Injury & Serious).**

| Study | Result | Reasoning and action |
|---|---|---|
| 13 | 69 | Probably error in data. Fatal and PDO results of the same study are reasonable. Delete Injury data |
| 15 | 88 | Small Δv (0.12 mph) makes ΔN/Δv large. Retain. |
| 38 | 196 | Small Δv (0.12 mph) makes ΔN/Δv large. Retain. |

As noted before, modeling here will be based on the count of fatal accidents. Studies the results of which are given in fatalities (i.e., number of victims) will not be used. Second, modeling will be based on the estimated number of accidents that would be expected in the after period had there been no change in mean speed (denoted either as π or as N), not on the count of before accidents used in Figure F-17 and Figure F-18. Third, to ensure that our results can be compared to those by Elvik et al., analysis will be confined to the category of injury accidents, not the joint category of injury & serious accidents. The result is in Figure F-19.

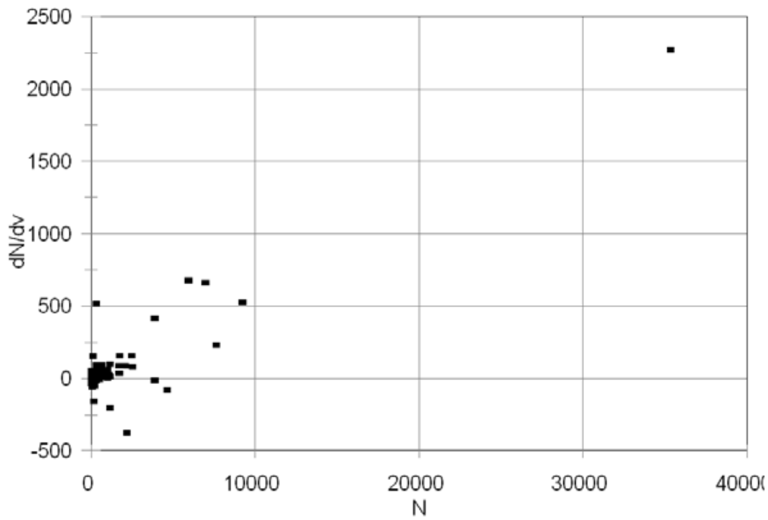


**Figure F-19. How dN/dv depends on N for Injury Accidents in USA & OTHER Countries.**

In Figure F-20 are shown USA data only. The relationship, although more tenuous, is still consistent with proportionality.

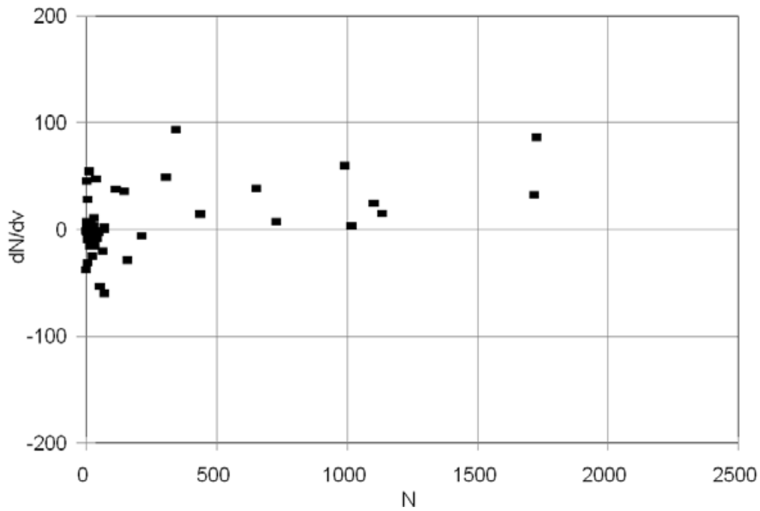


**Figure F-20. How dN/dv depends on N for Injury Accidents for USA data.**

For Slight & PDO accidents the relationship between estimates of dN/dv and before accidents is shown in Figure F-21 and Figure F-22. Unlike fatal accidents and injury & serious accidents, here the relationship is ambiguous. Without the rightmost point in Figure F-22 the increasing tendency would be very weak. (This data point, if retained, would an 'influential observation'.)

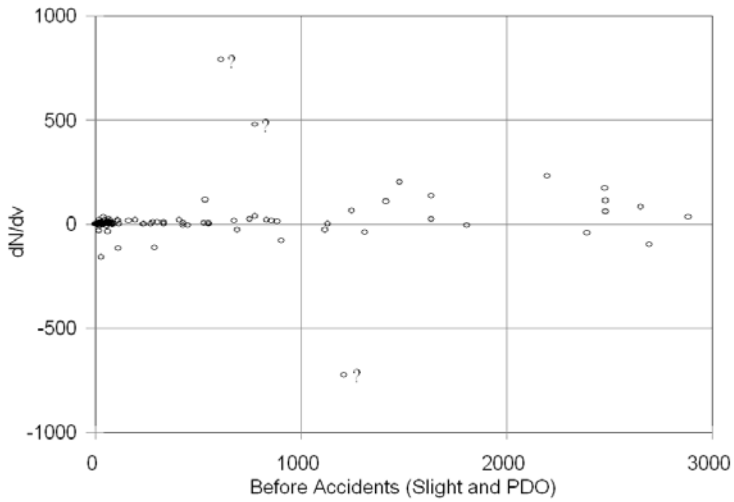


**Figure F-21. How dN/dv depends on Before Accidents for Slight & PDO Accidents when N<3,000.**

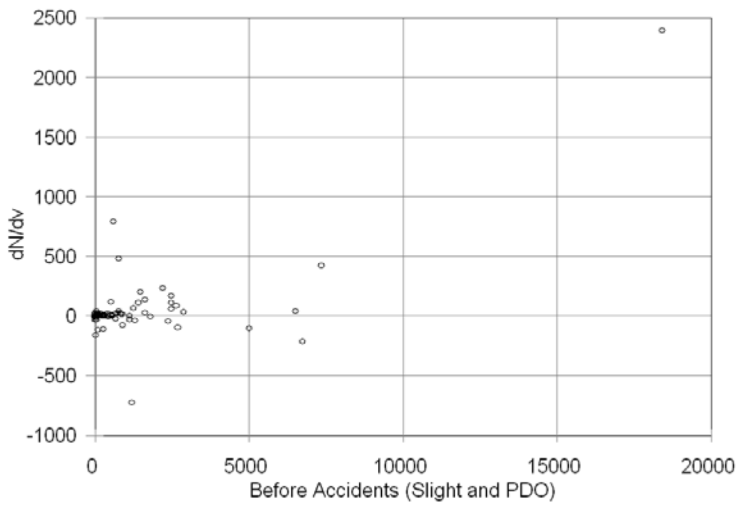


**Figure F-22. How dN/dv depends on Before Accidents for Slight & PDO Accidents, all N**

In addition, there are in there are a few apparent outliers noted in Figure F-21 by question marks.

**Table F-8. Outliers (Slight & PDO).**

| Study | Result | Reasoning and action |
|---|---|---|
| 15 | 89 | Small Δv (0.12 mph) makes ΔN/Δv large. Retain. |
| 38 | 197 | Small Δv (0.12 mph) makes ΔN/Δv large. Retain. |
| 39 | 201 | Small Δv (0.12 mph) makes ΔN/Δv large. Retain. |

With the realization that the noted outliers can be visually disregarded, there may be a small upward drift in the remaining data. Thus, even for the 'Slight & PDO' accidents when data from all countries are used, the assumption that dN/dv is proportional to N is perhaps tenable.

When only PDO accidents are used (without the Slight category), when N is used instead of the count of before accidents, and when only USA data are considered the relationship in Figure F-23 was obtained. Without the one point in the top right corner, no relationship with N would be indicated. Since that single point is very uncertain, it would be a stretch to assume that for PDO accidents a relationship of proportionality exists.

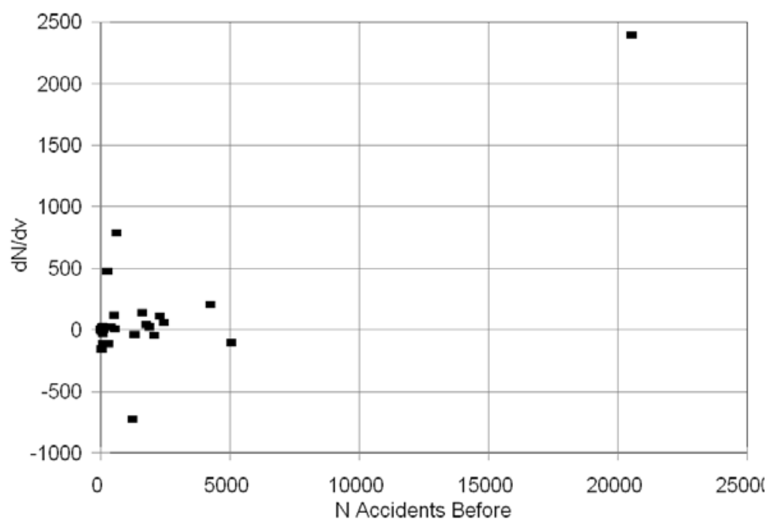


**Figure F-23. How dN/dv depends on N, PDO Accidents**

In summary, there is an indication in the data that dN/dv is proportional to N. The indication is strong for fatal accidents and gets progressively weaker as accident severity diminishes.

*The Relationship between (dN/dv)/N and 'v'*

As noted earlier, dN/dv is likely to be proportional to N. Therefore, using dN/dv=N×f(v), one can examine how dN/dv depends on v (the functional form f(v)) by plotting (dN/dv)/N against v. These plots are in Figure F-24, F-25 and F-26. Many of these estimates are rather inaccurate. Even so, neither individually nor collectively can one discern evidence of a systematic dependence on v. In particular, I can see no evidence that (dN/dv)/N diminishes as v increases as is implied by the functional form of the power model (see equation 22). Perhaps when, in modeling, the relative accuracy of the data points will be accounted for, some trend might emerge. In any case, at this point, nothing more elaborate than $f(v)=\beta_1+\beta_2 v$ is indicated. Whether $\beta_2$ will end up being used will depend on its magnitude (practical significance) and on the consistency of its sign.

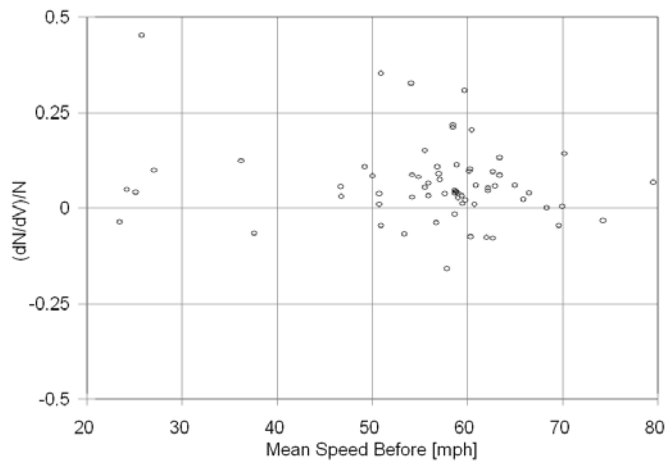


**Figure F-24. Relationship between (dN/dv)/N and Mean Speed Before (v) for FATAL accidents**

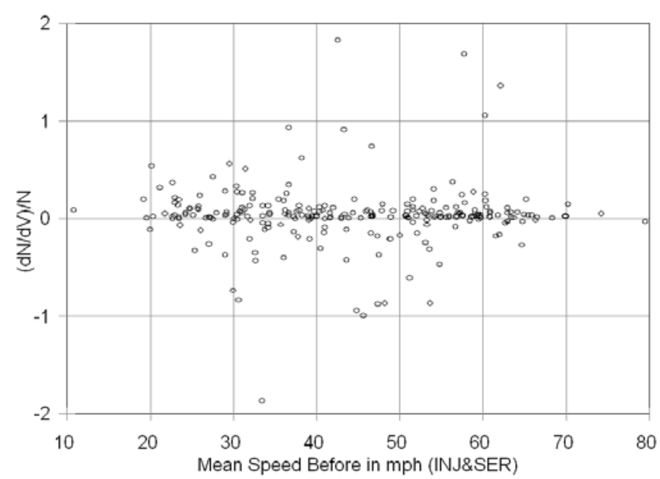


**Figure F-25. Relationship between (dN/dv)/N and Mean Speed Before (v) for INJURY & SERIOUS accidents**

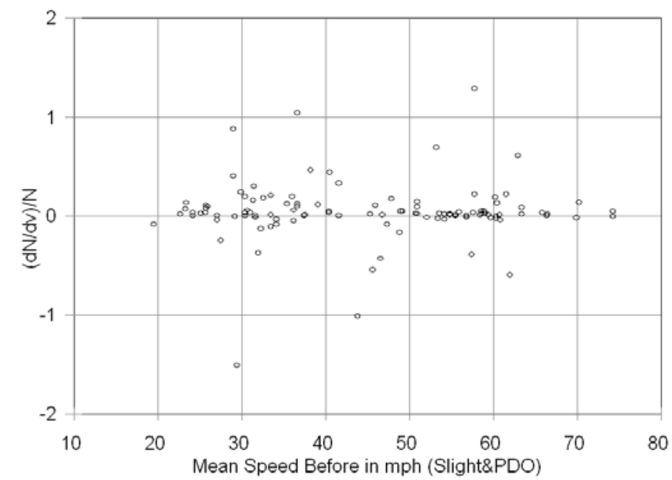


**Figure F-26. Relationship between (dN/dv)/N and Mean Speed Before (v) for SLIGHT & PDO accidents**

*The Relationship between dN and dv*

In this section, we examine only USA data. The relationship between dN and dv for Fatal accidents is shown in Figure F-27.

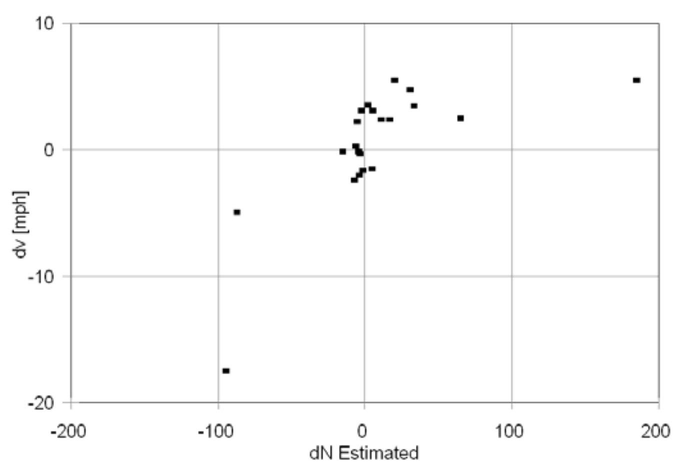


**Figure F-27. dN versus dv, USA Fatal Accidents**

The relationship between dN and dv for injury accidents is shown in Figure F-28. In spite of the non-specific cluster near the origin, the correlation is quite strong (0.85).

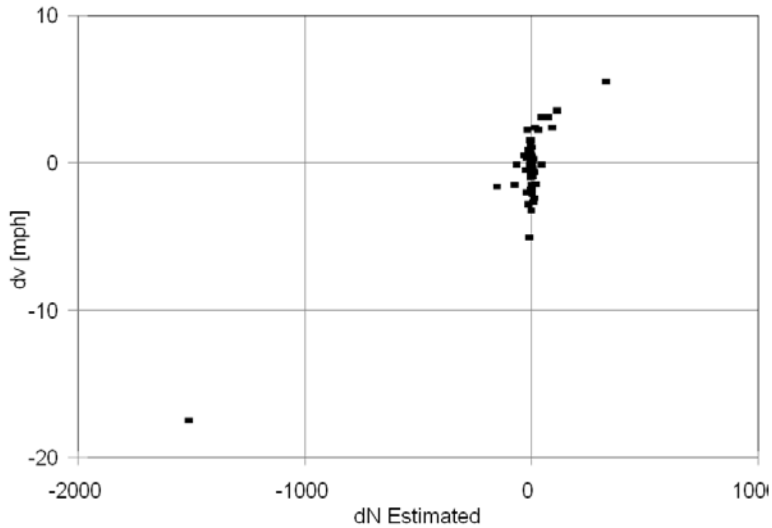


**Figure F-28. dN versus dv, Injury Accidents**

The relationship between dN and dv for PDO accidents is shown in Figure F-29. There is a non-specific cluster near the origin and two outlying points which create a semblance of correlation (0.6).

Taken together these three figures show that most of the time an increase in speed is associated with an increase in N and vice versa. However, in many cases, usually when dv is small, dN is very small.

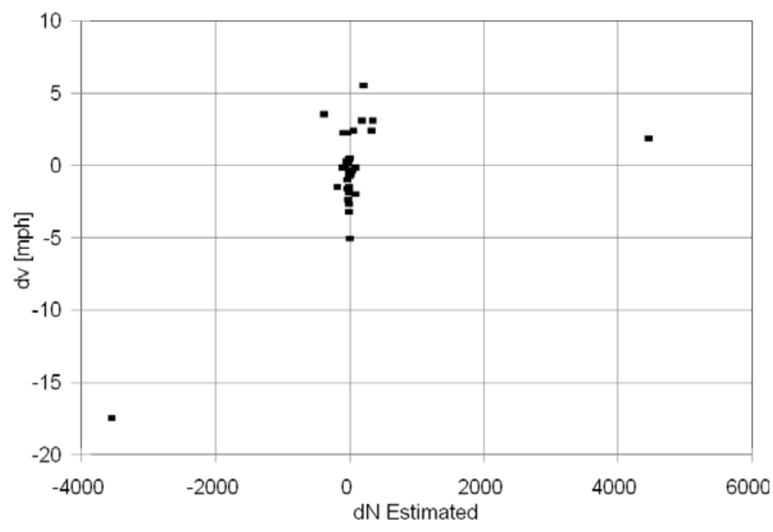


**Figure F-29. dN versus dv, PDO Accidents**

## Modeling

This is the main section of the report. In the first section I discuss matters of approach to modeling and the remaining three section deal with the modeling of fatal, Injury and PDO accidents.

*Approach and Assumptions*

Two issues are discussed in this section. The first describes the weight given to data of differing precision; the second specifies the way the scale parameter is computed.

Weight given to observed values differing in accuracy

Parameters will be estimated by minimizing the weighted sum of squared residuals. When the estimates of the dependent variable [be it dN, dN/dv or (dN/dv)/N] have different variances, a weighted least squared (WLS) criterion needs to be minimized. The weight of each residual follows from the argument below.

Let $y_1$ and $y_2$ be observed values of two dependent variables with variances $V_1$ and $V_2$. The subscripts 1 and 2 are so chosen that $V_1<V_2$. Think of $y_1$ as the average of n imagined independent observed values each of which has a variance $V_2$. That is,

$$y_1 = \frac{\text{Sum of n independent observed values each with a variance of } V_2}{n} \qquad \text{... 27}$$

If so,

$$\begin{aligned} VAR\{y_1\} &= \frac{nV_2}{n^2} = V_1 \\ n &= \frac{V_2}{V_1} \end{aligned} \qquad \text{... 28}$$

It follows that one observed value of $y_1$ can be regarded as the mean of $V_2/V_1$ independent observed values each with a variance of $V_2$. Thus, if the observed value $y_2$ is one data point with variance $V_2$, the observed value $y_1$ equivalent to $V_2/V_1$ data points with variance $V_2$. When it comes to summing squared residuals, if the residual of $y_2$ enters the sum once, the residual of $y_1$ must enter the sum $V_2/V_1$ times. This can be generalized to many observations of the dependent variable as:

$$\text{Weight of observation } i = \frac{\text{Largest Variance amongst all observed values}}{\text{Variance of observed value } i} \qquad \text{... 29}$$

Since the weighed sum of squares is the maximand when it comes to parameter estimation, and the maximum does not depend on the common factor of 'Largest Variance amongst all observed Values', the same parameter estimates will be obtained using:

$$\text{Weight of observation } i = \frac{1}{\text{Variance of observed value } i} \quad \text{... 30}$$

Sum Constraint

In general the sum of dependent variable values predicted by a regression model differs from the sum of the observed values. This makes the comparison of different models awkward since they will differ not only in the weighted sum of squares but also in the sum of predicted values. To eliminate this complication I will always minimize the weighed sum of squares subject to the constraint that the sum of predicted values be equal to the sum of observed values.

*Modeling Fatal Accidents*

The sequence of models and their evolution is described below. All estimated parameters are in Table F-12.

Model Estimation

Models 1. $dN = \alpha N^{\beta_N} dv$

In model 1a $\beta_N$ is set to 1. Therefore, $\alpha$ is estimated not by WLS but by equating the sum of estimated and predicted values of dN. The corresponding sum of weighted squared residuals is 133.5. In model 1b $\beta_N$ is estimated to be1.14 and the sum of weighted squared residuals is 8.0.

Models 2. $dN = \alpha \times N^{\beta_N} \times f(v) \times dv$

In model 2a I use $f(v)=1+\beta_v v$ and $\beta_N=1$. The estimate of $\beta_v$ is -0.0098and the sum of weighted squared residuals 12.7. This is a small improvement from model 1a indicating that the inclusion of the mean speed adds little to prediction. If it was true that $\beta_v<0$ it would mean that the larger the speed the smaller the dN due to a fixed change dv. When $\beta_N$ is not set to 1 but is estimated (1.14 in model 2b) $\beta_v$ is estimated to be -0.0137. The sum of weighed square residuals in this case is 6.7.

I tried $f(v)=1/(1+\beta_v \times v)$ is used instead of $f(v)=(1+\beta_v \times v)$ and $f(v) = e^{\beta_v \times v}$ but these functional forms had a larger sum of squared residuals than models 2a and 2b. Thus, the simple linear form $f(v)=1+\beta_v v$ is to be preferred. (Note that the functions f(v) in models 3 are all very similar when $\beta_v \times v$ is small compared to 1 because $e^{\beta_v v} = 1 + \frac{\beta_v v}{1!} + \frac{(\beta_v v)^2}{2!} + \ldots$ and $(1+\beta_v v)^{-1} = 1 - \beta_v v + (\beta_v v)^2 - \ldots$ .).

Model Examination and Modification

A. The relationship between the values of dN estimated directly from the data and the values predicted by model 2b is in Figure F-30.

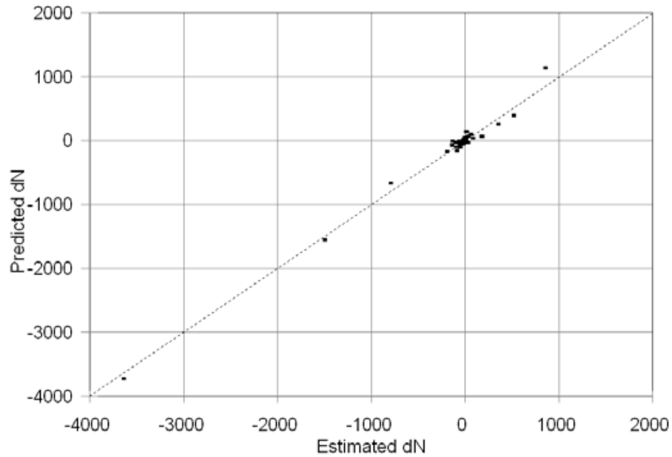


**Figure F-30. Comparison of all dN in the data with those predicted by model 2b.**

The correspondence between the dN estimated from the data and that predicted by the model is quite strong. There is no indication that the model prediction quality depends on whether dv was positive (to the right of 0) or negative, or on whether dN was large or small.

B. One of the questions to be answered was whether a model based on experience in all countries applies to USA. How model 2b predicts in the USA and in Other Countries is shown in Table F-9.

**Table F-9. Comparison of predicted and estimated dN (Fatal) for USA and Other Countries using model 2b**

| | | |
|---|---|---|
| Other Countries | Sum of Estimated dN<br>Sum of Predicted dN | -296.4<br>-433.0 |
| USA | Sum of Estimated dN<br>Sum of Predicted dN | 1008.6<br>1145.2 |
| Total Sum of Estimated | | 712.2 |
| Total Sum of Predicted | | 712.2 |

Because much of the evidence about dN comes from the USA, the sum of predictions for the USA is close to the sum of estimated dN.

C. The Cumulative Residuals plot against the Mean Speed Before for model 2b is in Figure F-31. It shows that the model fits well in all ranges of v.

D. The next question examined is whether the use of all data, including those studies in which the possibility of RTM bias was noted as present (YES) does not introduce bias into the results.

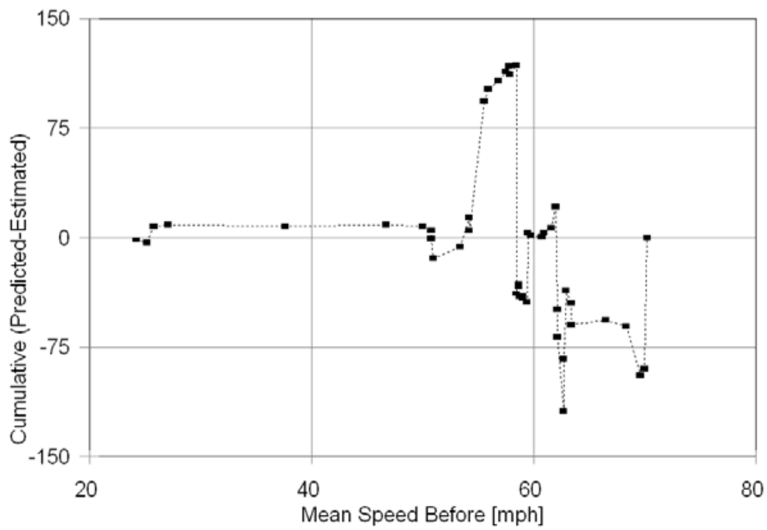


**Figure F-31. Cumulative Residuals Plot for v for model 2b**

**Table F-10. Comparison of predicted and estimated dN for RTM using model 2b**

| | | |
|---|---|---|
| RTM not present | Sum of Estimated dN | 661.1 |
| | Sum of Predicted dN | 705.4 |
| RTM may be present | Sum of Estimated dN | 51.1 |
| | Sum of Predicted dN | 6.8 |
| Total Sum of Predicted dN | | 712.2 |

Were RTM present (in those cases when the before accidents were unusually large) one might expect that the estimated dN would be smaller than the predicted (the biased reductions are larger than they should be). Since the opposite is the case it is unlikely that the results are biased by using data for which the possibility of the presence of RTM is shown as 'YES'. In any case, the influence of data where RTM is possible is very small.

E. The last question examined is whether model 2b performs well in all 'environments'. The comparison is in Table F-11.

**Table F-11. Comparison of predicted and estimated dN for Environment using model 2b**

| | | |
|---|---|---|
| ALL | Sum of Estimated dN | -8.4 |
| | Sum of Predicted dN | -1.6 |
| FREEWAY | Sum of Estimated dN | 68.1 |
| | Sum of Predicted dN | 28.1 |
| RESIDENTIAL | Sum of Estimated dN | -22.0 |
| | Sum of Predicted dN | -22.8 |
| RURAL | Sum of Estimated dN | 673.9 |
| | Sum of Predicted dN | 708.6 |
| URBAN | Sum of Estimated dN | 0.5 |
| | Sum of Predicted dN | -0.2 |
| Sum of Predicted dN | | 712.2 |

It appears that the model predicts well without adding a variable for Environment.

**Table F-12. Summary of Model Evolution and Parameters**

| | | $\alpha$ | $\beta_N$ | $\beta_{dv}$ | $\beta_v$ | Weighted ssq |
|---|---|---|---|---|---|---|
| 1a | $dN=\alpha\times N\times dv$ | 0.1046 | Set to 1 | Set to 1 | | 13.50 |
| 1b | $dN=\alpha\times N^{\beta N}\times dv$ | 0.0247 | 1.14 | Set to 1 | | 8 |
| 2a | $dN=\alpha\times N\times(1+\beta_v*v)\times dv$ | 0.2666 | Set to 1 | Set to 1 | -0.0098 | 12.7 |
| 2b | $dN=\alpha\times N^{\beta N}*(1+\beta_v*v)*dv$ | 0.1582 | 1.14 | Set to 1 | -0.0137 | 6.70 |

Results for Fatal Accidents

I will discuss results for models 2a and 2b:

$$dN = 0.2666N(1-0.0098v)dv$$

and ... 31

$$dN = 0.1582N^{1.14}(1-0.0137v)dv$$

Thus, e.g., if the mean speed 'before' is 68 mph, N 'before' is 500, and a measure is expected to reduce it by 1.2 mph then model 2a predicts 0.2666×500×(1-0.0098×68)×(-1.2)=-53.4 and model 2b predicts $0.1582\times500^{1.14}\times(1-0.0137\times68)\times(-1.2)$=-15.5.

Models 2a and 2b are compared with Elvik's model for fatal accidents in Figure F-32.

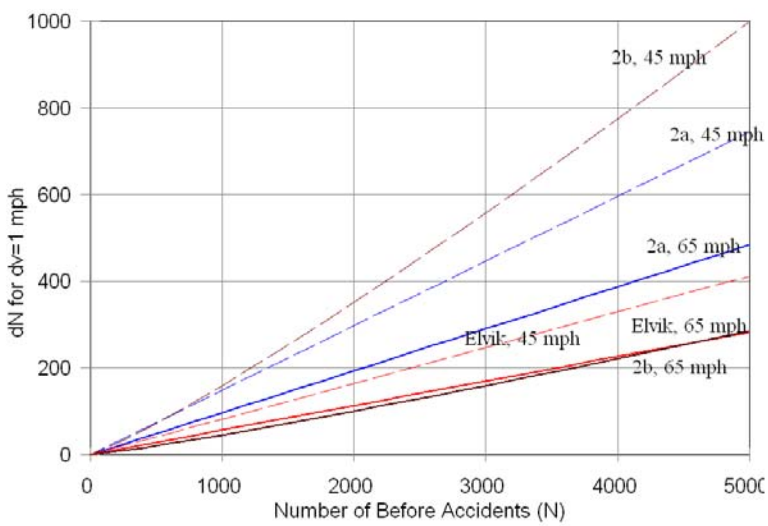


**Figure F-32. Comparison of Models for Fatal Accidents**

The numerical values Predicted by models 2a and 2b and Elvik (power=3.6) are dissimilar. Model 2a predicts larger dN than Elvik and is somewhat more sensitive to mean speed. Model 2b is very sensitive to mean speed and predicts very large dN at low speeds.. The model would predict non-sense for v>73 mph. The sensitivity of 2b to mean speed and the simplicity of model 2a, commend model 2a for use.

The AMFs for 2a is approximately

$$\frac{N+dN}{N}=1+\frac{dN}{N}=1+0.2666(1-0.0098v)dv \qquad \text{... 32}$$

Thus, e.g., if the mean speed 'before' is 68 mph and a measure is expected to reduce it by 1.2 mph then the AMF is 1+0.2666(1-0.0098×68)×(-1.2)=1-0.107=0.893. Alternatively, using the dN computed earlier 1-53.4/500=0.893.

The functional relationship between N and v is obtained by solving the separable differential equation $\frac{dN}{N}=\alpha(1+\beta v)dv$. The solution for model 2a is

$$\ln(N)=\alpha(v+\frac{\beta}{2}v^2)+C$$

$$\text{or} \qquad \text{... 33}$$

$$N=e^{\alpha(v+\frac{\beta}{2}v^2)+C}$$

If for some system N=N* when v=v* then

$$e^C=N^*e^{-\alpha(v^*+\frac{\beta}{2}v^{*2})}$$

$$\text{and} \qquad \text{... 34}$$

$$N=N^*e^{\alpha[(v-v^*)+\frac{\beta}{2}(v^2=v^{*2})]}\cong N^*e^{\alpha\Delta v(1+\beta v^*)}$$

The AMF for v and v* is then $e^{\alpha[(v-v^*)+\frac{\beta}{2}(v^2=v^{*2})]}\cong e^{\alpha\Delta v(1+\beta v^*)}$. Thus, e.g., if v*=68 mph and v=68-1.2=66.8 mph then exp{0.2666×[-1.2+(-0.0098/2)(66.8$^2$-68$^2$)]=0.893

*Modeling Injury Accident*

The sequence of models and their evolution is described below. All estimated parameters are in Table F-13.

<u>Model Estimation</u>

Models 1. $dN=\alpha N^{\beta_N}dv$

In model 1a $\beta_N$ is set to 1. Therefore, α is estimated not by WLS but by equating the sum of estimated and predicted values of dN. The corresponding sum of weighted squared residuals is 70.8. In model 1b $\beta_N$ is estimated to be 1.17 and the sum of weighted squared residuals is 63.2.

Models 2. $dN = \alpha \times N^{\beta_N} \times f(v) \times dv$

In model 2a I use $f(v)=1+\beta_v v$ and $\beta_N=1$. The estimate of $\beta_v$ is -0.0051 and the sum of weighted squared residuals is 69.8. This is a very small improvement from model 1a indicating that the inclusion of the mean speed added little to prediction. If it was true that $\beta_v<0$ it would mean that the larger the speed the smaller the dN due to a fixed change dv. When $\beta_N$ is not set to 1 but is estimated (1.19 in model 2b) $\beta_v$ is estimated to be +0.0032. That the sign is reversed is another indication of the uncertain dependence on v. The sum of weighed square residuals in this case is reduced to 63.2.

I tried $f(v)=1/(1+\beta_v \times v)$ is used instead of $f(v)=(1+\beta_v \times v)$ and $f(v) = e^{\beta_v \times v}$ but these functional forms had a larger sum of squared residuals than models 2a and 2b. Thus, the simple linear form $f(v)=1+\beta_v v$ is to be preferred. (Note that the functions f(v) in models 3 are all very similar when $\beta_v \times v$ is small compared to 1 because $e^{\beta_v v} = 1 + \frac{\beta_v v}{1!} + \frac{(\beta_v v)^2}{2!} + \ldots$ and $(1+\beta_v v)^{-1} = 1 - \beta_v v + (\beta_v v)^2 - \ldots$ .).

**Table F-13. Summary of Model Evolution and Parameters for Injury Accidents**

| | | $\alpha$ | $\beta_N$ | $\beta_{dv}$ | $\beta_v$ | Weighted ssq |
|---|---|---|---|---|---|---|
| 1a | $dN=\alpha \times N \times dv$ | 0.0670 | Set to 1 | Set to 1 | | 70.8 |
| 1b | $dN=\alpha \times N^{\beta N} \times dv$ | 0.0138 | 1.17 | Set to 1 | | 63.2 |
| 2a | $dN=\alpha \times N*(1+\beta_v \times v) \times dv$ | 0.0838 | Set to 1 | Set to 1 | -0.0051 | 69.8 |
| 2b | $dN=\alpha \times N^{\beta N} \times (1+\beta_v * v)*dv$ | 0.0103 | 1.19 | Set to 1 | 0.0032 | 63.2 |

Model Examination and Modification

A. The relationship between the values of dN estimated directly from the data and the values predicted by model 2a is in Figure F-33.

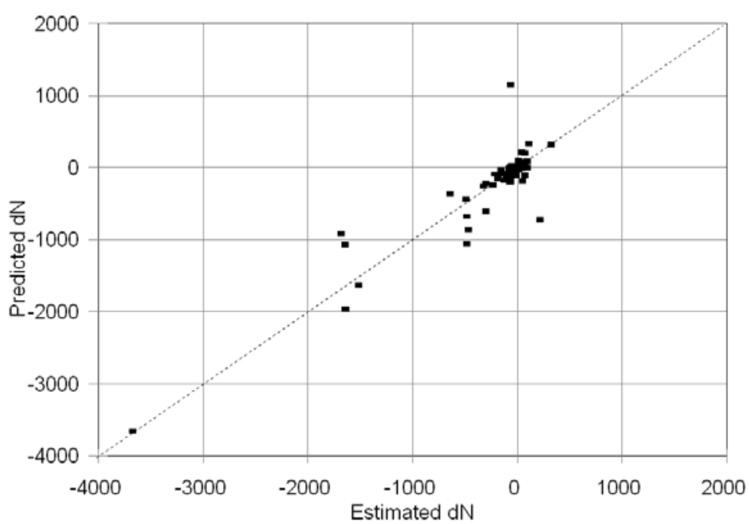


**Figure F-33. Comparison of all dN in the data with those predicted by model 2b.**

A. The correspondence between the dN from the data and that predicted by the model is reasonable but less pronounced than for fatal accidents. There is no indication that the model prediction quality depends on whether dv was positive (above 0,0) or negative (below 0,0)or on whether dN was large or small.

B. One question was whether a model based on experience in all countries applies to the USA. How model 2a predicts in the USA and in Other Countries is shown in Table F-14.

**Table F-14 Comparison of predicted and estimated dN for USA and Other Countries using model 2b**

| | | |
|---|---|---|
| Other Countries | Sum of Estimated dN<br>Sum of Predicted dN | -13904.7<br>-14453.0 |
| USA | Sum of Estimated dN<br>Sum of Predicted dN | -1126.1<br>-577.9 |
| Total Sum of Estimated | | -15030.8 |
| Total Sum of Predicted | | -15030.8 |

The sum of predictions for the USA is less than the sum of estimated dN. I attempted to add a multiplicative parameter to account for the difference between USA and Other Countries but it failed to reduce the discrepancy. Considering the large variance of the sum of dN the model is acceptable. An attempt to estimate a separate model using only USA Injury Accident data might be warranted.

C. The Cumulative Residuals plot against the Mean Speed Before for model 2a is in Figure F-34. It shows that the model fits well in all ranges of v. However, the vertical drops and rises indicate the presence of several outliers. Without reading the original reports it is difficult to make decisions about whether they should be excluded from analysis.

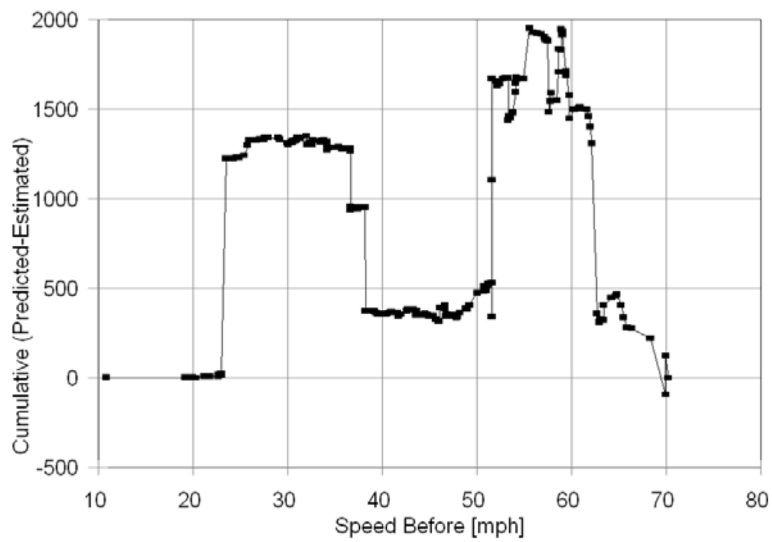


**Figure F-34. Cumulative Residuals Plot for v for model 2a**

D. The next question examined is whether the use of all data, including those studies in which the possibility of RTM bias was noted as present (YES) does not introduce bias into the results. How the sum of estimated dN compares to the sum of predicted dN is shown in Table F-15.

**Table F-15. Comparison of predicted and estimated dN for RTM using model 2a**

| | | |
|---|---|---|
| RTM not present | Sum of Estimated dN | -8475.0 |
| | Sum of Predicted dN | -8206.0 |
| RTM may be present | Sum of Estimated dN | -6555.8 |
| | Sum of Predicted dN | -6824.8 |
| Total Sum of Predicted dN | | -15030.8 |

The possibility of RTM bias is in this case not evident and the data indicating the possibility of RTM can be used for modeling.

E. The last question examined is whether model 2a performs well in all 'environments'. The comparison is in Table F-16.

**Table F-16. Comparison of predicted and estimated dN for Environment using model 2a.**

| | | |
|---|---|---|
| ALL | Sum of Estimated dN | -318.0 |
| | Sum of Predicted dN | -251.9 |
| FREEWAY | Sum of Estimated dN | -4978.7 |
| | Sum of Predicted dN | -4923.6 |
| RESIDENTIAL | Sum of Estimated dN | -55.0 |
| | Sum of Predicted dN | -42.4 |
| RURAL | Sum of Estimated dN | -7342.8 |
| | Sum of Predicted dN | -7625.2 |
| URBAN | Sum of Estimated dN | -2336.3 |
| | Sum of Predicted dN | -2187.7 |
| Sum of Predicted dN | | -15030.8 |

It appears that the model predicts well without adding an additional variable for Environment.

<u>Results for Injury Accidents</u>

I will discuss results for model 2a:

$$dN = 0.0838N(1 - 0.0051v)dv \qquad ... 35$$

Thus, e.g., if the mean speed 'before' is 68 mph, N 'before' is 500, and a measure is expected to reduce v by 1.2 mph then model 2a predicts dN=0.0838×500×(1-0.0051×68)×(-1.2)=-32.8.

Model 2a is compared with Elvik's model for Injury accidents in Figure F-35.

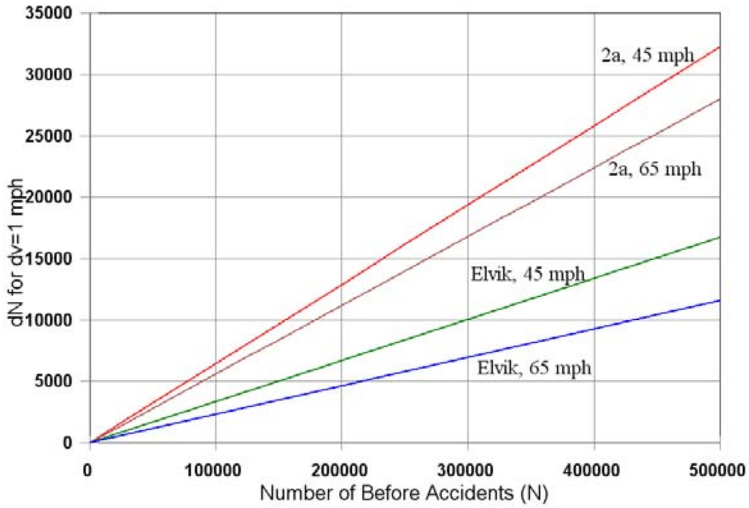


**Figure F-35. Comparison of Models for Injury Accidents**

The numerical values predicted by model 2a and 2b and those predicted by Elvik (power=1.5) are not similar. The simplicity of model 2a, commends it for use.

The AMFs for 2a is approximately

$$\frac{N+dN}{N} = 1+\frac{dN}{N} = 1+0.0838(1-0.0051v)dv \qquad \text{... 36}$$

Thus, e.g., if the mean speed 'before' is 68 mph and a measure is expected to reduce it by 1.2 mph then the AMF is 1+0.0838(1-0.0051×68)×(-1.2)=1-0.066=0.934. Alternatively, using the dN computed earlier 1-32.8/500=1-0.066=0.934.

The functional relationship between N and v is obtained by solving the separable differential equation $\frac{dN}{N} = \alpha(1+\beta v)dv$. The solution for model 2a is

$$\ln(N) = \alpha(v+\frac{\beta}{2}v^2)+C$$

or ... 37

$$N = e^{\alpha(v+\frac{\beta}{2}v^2)+C}$$

If for some system N=N* when v=v* then

$$e^{C} = N^{*}e^{-\alpha(v^{*}+\frac{\beta}{2}v^{*2})}$$

and ... 38

$$N = N^{*}e^{\alpha[(v-v^{*})+\frac{\beta}{2}(v^{2}=v^{*2})]} \cong N^{*}e^{\alpha\Delta v(1+\beta v^{*})}$$

The AMF for v and v* is then $e^{\alpha[(v-v^{*})+\frac{\beta}{2}(v^{2}=v^{*2})]} \cong e^{\alpha\Delta v(1+\beta v^{*})}$. Thus, e.g., if v*=68 mph and v=68-1.2=66.8 mph then $e^{0.0838\times[-1.2+\frac{-0.0051}{2}(66.8^{2}-68^{2})]} = 0.934$ or $e^{0.0838\times(-1.2)(1-0.0051\times 68)} = 0.936$

*Modeling PDO Accidents*

As for fatal and for injury accidents in sections 4.3.2 and 4.3.3, I attempted to develop and estimate models to predict dN as a function of N, f(v) and dv. However, modeling ran into the difficulty described below.

The simplest model used was always dN=α×N×dv. In this model α is estimated not by WLS but by equating the sum of estimated and predicted values of dN. It turned out that α=

-0.0077. That is that, which the sum of dN is positive (398) the sum of N×dv is negative (-51804) and the only way the make the two equal is to multiply by a negative number. But this would mean that when mean speed is decreased one predicts an increase in PDO accidents.

Instead of estimating α by ensuring that the sum of estimated dN equals the sum of predicted dN (see section 4.3.1.2) I tried to estimate it by minimizing the weighted sum of squares. Now the estimate of α was 0.0230. However, while the sum of estimated dN (398) is positive, the sum of predicted dN is now very different and negative (-1193). I am led to conclude that the data for PDO accidents does not lend itself to modeling.

This conclusion is not entirely unexpected. As noted earlier in Figure F-21, Figure F-22 and Figure F-23 the is no evident relationship between dN/dv and N. Logical reasoning leads one to expect that if on X miles of road a change of dv=Y yields a change in accidents dN=Z then on 2X of identical road the same dv=Y will change accidents by dN=2Z. That is, larger road systems are expected to be associated with larger accident changes per unit change in mean speed. This basic relationship is not observed in the data. The possible explanations are many. First, PDO accidents are not well reported and the change in what is reported may reflect the availability of police manpower before and after a speed change. Second, since a reduction in mean speed may reduce the chance of injury, some accidents reported before the change as injury or fatal may now fall into the PDO category. Third, perhaps Simpson's paradox is at work and aggregation over various variables (urban vs, rural, freeway vs. all etc.) may be obscuring some important regularities.

**Summary of Modeling Results**

Modeling was done by weighted least squares (WLS) estimation. The weight of an estimated value of dN was the reciprocal value of its variance. To facilitate model comparisons and to limit the WLS to predictions that closely mimic the data, the estimates were constrained by the relation: Sum of predicted values of dN = Sum of estimated value of dN.

For <u>Fatal Accidents</u> I found in model 2a:

$$dN = 0.2666N(1-0.0098\,v)dv \qquad ...\ 39$$

Thus, e.g., if the mean speed 'before' (v) is 68 mph, N=500, and a measure is expected to reduce v by 1.2 mph then model 2a predicts dN=0.2666×500×(1-0.0098×68)×(-1.2)=-53.4 fatal accidents.

The numerical values of dN predicted by model 2a and by Elvik et al. are dissimilar.

The AMFs for model 2a is approximately

$$\frac{N+dN}{N} = 1+\frac{dN}{N} = 1+0.2666(1-0.0098v)dv \qquad ...\ 40$$

Thus, e.g., if the mean speed 'before' is 68 mph and a measure is expected to reduce it by 1.2 mph then the AMF is 1+0.2666(1-0.0098×68)×(-1.2)=1-0.107=0.893. Alternatively, using the dN computed earlier 1-53.4/500=0.893.

The functional relationship between N and v is obtained by solving the separable differential equation $\frac{dN}{N} = \alpha(1+\beta v)dv$. If for some system N=N* when v=v* then

$$N = N^{*} e^{\alpha[v-v^{*}+\frac{\beta}{2}(v^{2}-v^{*2})]} \qquad ...\ 41$$

In this α=0.2666 and β=-0.0098.
The AMF for v and v* is then

$$AMF = \frac{N}{N^{*}} = e^{\alpha[v-v^{*}+\frac{\beta}{2}(v^{2}-v^{*2})]} \qquad ...\ 42$$

For the numerical values above AMF=0897.

For <u>Injury Accidents</u> I found in model 2a:

$$dN = 0.0838N(1-0.0051v)dv \qquad ...\ 43$$

Thus, e.g., if the mean speed 'before' is 68 mph, N 'before' is 500, and a measure is expected to reduce v by 1.2 mph then model 2a predicts dN=0.0838×500×(1-0.0051×68)×(-1.2)=-32.8.

The numerical values predicted by model 2a and Elvik (power=1.5) are similar.

The AMFs for model 2a the AMF is approximately

$$\frac{N+dN}{N} = 1+\frac{dN}{N} = 1+0.0838(1-0.0051v)dv \qquad ...\ 44$$

Thus, e.g., if the mean speed 'before' is 68 mph and a measure is expected to reduce it by 1.2 mph then the AMF is 1+0.0838(1-0.0051×68)×(-1.2)=1-0.066=0.934. Alternatively, using the dN computed earlier 1-32.8/500=1-0.066=0.934.

The functional relationship between N and v is in equations 41 and 42 with α=0.0838 and β=-0.0051. For the numerical values above the AMF=0.934.

For PDO accidents I found that the data contain no indication that dN/dv is proportional to N. The possibility that the change in accidents due a change in mean speed is not related to the size of the system makes modeling of PDO accidents with this data questionable.

## COMPARING THE RESULTS OF MODELING BY APPROACHES 1 AND 2

Two different modeling approaches were applied to the very same data. The prediction models of both approaches for injury and for fatal accidents are summarized below.

Approach 1:

$$\text{Injury Accident AMF} = e^{1.368\times\ln\left(\frac{v_{with}}{v_{without}}\right)-(70.9-51.2\text{ if urban})\times(1/v_{with}-1/v_{without})} \quad \ldots 45$$

$$\text{Fatal Accident AMF} = e^{2.742\times\ln\left(\frac{v_{with}}{v_{without}}\right)-(70.9-51.2\text{ if urban})\times(1/v_{with}-1/v_{without})} \quad \ldots 46$$

Approach 2:

$$\text{Injury Accident AMF} = 1 + 0.0838\times(1-0.0051\times v_{without})\times\Delta v \quad \ldots 47$$

$$\text{Fatal Accident AMF} = 1 + 0.2666\times(1-0.0098\times v_{without})\times\Delta v \quad \ldots 48$$

The AMFs computed by approaches 1 and 2 for selected values of $v_{without}$ and Δv are in Table F-17. The AMFs for approach 1 are based on rural freeway and highways.

## Table F-17. AMFs by Both Approaches and their Ratios.

**Injury Approach 1**

| Δv [mph] | $v_{without}$ [mph] 30 | 40 | 50 | 60 | 70 | 80 |
|---|---|---|---|---|---|---|
| -5 | 0.49 | 0.65 | 0.74 | 0.80 | 0.84 | 0.86 |
| -4 | 0.57 | 0.71 | 0.79 | 0.84 | 0.87 | 0.89 |
| -3 | 0.67 | 0.78 | 0.84 | 0.88 | 0.90 | 0.92 |
| -2 | 0.77 | 0.85 | 0.89 | 0.92 | 0.93 | 0.94 |
| -1 | 0.88 | 0.92 | 0.94 | 0.96 | 0.97 | 0.97 |
| 0 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 1 | 1.13 | 1.08 | 1.06 | 1.04 | 1.03 | 1.03 |
| 2 | 1.27 | 1.16 | 1.11 | 1.09 | 1.07 | 1.06 |
| 3 | 1.41 | 1.25 | 1.17 | 1.13 | 1.10 | 1.09 |
| 4 | 1.57 | 1.34 | 1.23 | 1.18 | 1.14 | 1.12 |
| 5 | 1.73 | 1.43 | 1.30 | 1.22 | 1.18 | 1.14 |

**Fatal Approach 1**

| Δv [mph] | $v_{without}$ [mph] 30 | 40 | 50 | 60 | 70 | 80 |
|---|---|---|---|---|---|---|
| -5 | 0.38 | 0.54 | 0.64 | 0.71 | 0.75 | 0.79 |
| -4 | 0.47 | 0.62 | 0.70 | 0.76 | 0.80 | 0.83 |
| -3 | 0.58 | 0.70 | 0.77 | 0.82 | 0.85 | 0.87 |
| -2 | 0.70 | 0.79 | 0.84 | 0.87 | 0.90 | 0.91 |
| -1 | 0.84 | 0.89 | 0.92 | 0.94 | 0.95 | 0.96 |
| 0 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 1 | 1.18 | 1.12 | 1.09 | 1.07 | 1.05 | 1.05 |
| 2 | 1.38 | 1.24 | 1.18 | 1.14 | 1.11 | 1.09 |
| 3 | 1.61 | 1.38 | 1.27 | 1.21 | 1.17 | 1.14 |
| 4 | 1.86 | 1.53 | 1.37 | 1.29 | 1.23 | 1.19 |
| 5 | 2.14 | 1.68 | 1.48 | 1.36 | 1.29 | 1.24 |

**Injury Approach 2**

| Δv [mph] | $v_{without}$ [mph] 30 | 40 | 50 | 60 | 70 | 80 |
|---|---|---|---|---|---|---|
| -5 | 0.65 | 0.67 | 0.69 | 0.71 | 0.73 | 0.75 |
| -4 | 0.72 | 0.73 | 0.75 | 0.77 | 0.78 | 0.80 |
| -3 | 0.79 | 0.80 | 0.81 | 0.83 | 0.84 | 0.85 |
| -2 | 0.86 | 0.87 | 0.88 | 0.88 | 0.89 | 0.90 |
| -1 | 0.93 | 0.93 | 0.94 | 0.94 | 0.95 | 0.95 |
| 0 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 1 | 1.07 | 1.07 | 1.06 | 1.06 | 1.05 | 1.05 |
| 2 | 1.14 | 1.13 | 1.12 | 1.12 | 1.11 | 1.10 |
| 3 | 1.21 | 1.20 | 1.19 | 1.17 | 1.16 | 1.15 |
| 4 | 1.28 | 1.27 | 1.25 | 1.23 | 1.22 | 1.20 |
| 5 | 1.35 | 1.33 | 1.31 | 1.29 | 1.27 | 1.25 |

**Fatal Approach 2**

| Δv [mph] | $v_{without}$ [mph] 30 | 40 | 50 | 60 | 70 | 80 |
|---|---|---|---|---|---|---|
| -5 | 0.06 | 0.19 | 0.32 | 0.45 | 0.58 | 0.71 |
| -4 | 0.25 | 0.35 | 0.46 | 0.56 | 0.67 | 0.77 |
| -3 | 0.44 | 0.51 | 0.59 | 0.67 | 0.75 | 0.83 |
| -2 | 0.62 | 0.68 | 0.73 | 0.78 | 0.83 | 0.88 |
| -1 | 0.81 | 0.84 | 0.86 | 0.89 | 0.92 | 0.94 |
| 0 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 1 | 1.19 | 1.16 | 1.14 | 1.11 | 1.08 | 1.06 |
| 2 | 1.38 | 1.32 | 1.27 | 1.22 | 1.17 | 1.12 |
| 3 | 1.56 | 1.49 | 1.41 | 1.33 | 1.25 | 1.17 |
| 4 | 1.75 | 1.65 | 1.54 | 1.44 | 1.33 | 1.23 |
| 5 | 1.94 | 1.81 | 1.68 | 1.55 | 1.42 | 1.29 |

**AMF Ratio App.1/App.2** (Injury)

| Δv [mph] | $v_{without}$ [mph] 30 | 40 | 50 | 60 | 70 | 80 |
|---|---|---|---|---|---|---|
| -5 | 0.75 | 0.97 | 1.08 | 1.12 | 1.14 | 1.15 |
| -4 | 0.80 | 0.97 | 1.05 | 1.09 | 1.11 | 1.11 |
| -3 | 0.85 | 0.97 | 1.03 | 1.06 | 1.07 | 1.08 |
| -2 | 0.90 | 0.98 | 1.02 | 1.04 | 1.05 | 1.05 |
| -1 | 0.95 | 0.99 | 1.01 | 1.02 | 1.02 | 1.02 |
| 0 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 1 | 1.05 | 1.01 | 0.99 | 0.99 | 0.98 | 0.98 |
| 2 | 1.11 | 1.03 | 0.99 | 0.97 | 0.96 | 0.96 |
| 3 | 1.16 | 1.04 | 0.99 | 0.96 | 0.95 | 0.95 |
| 4 | 1.22 | 1.06 | 0.99 | 0.95 | 0.94 | 0.93 |
| 5 | 1.28 | 1.07 | 0.99 | 0.95 | 0.93 | 0.92 |

**AMF Ratio App.1/App.2** (Fatal)

| Δv [mph] | $v_{without}$ [mph] 30 | 40 | 50 | 60 | 70 | 80 |
|---|---|---|---|---|---|---|
| -5 | 6.42 | 2.84 | 2.00 | 1.57 | 1.30 | 1.11 |
| -4 | 1.90 | 1.75 | 1.54 | 1.36 | 1.20 | 1.08 |
| -3 | 1.32 | 1.36 | 1.30 | 1.22 | 1.13 | 1.05 |
| -2 | 1.12 | 1.17 | 1.16 | 1.12 | 1.08 | 1.03 |
| -1 | 1.03 | 1.06 | 1.06 | 1.05 | 1.03 | 1.01 |
| 0 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 1 | 0.99 | 0.96 | 0.96 | 0.96 | 0.97 | 0.99 |
| 2 | 1.01 | 0.94 | 0.92 | 0.93 | 0.95 | 0.98 |
| 3 | 1.03 | 0.93 | 0.90 | 0.91 | 0.93 | 0.97 |
| 4 | 1.06 | 0.93 | 0.89 | 0.89 | 0.92 | 0.97 |
| 5 | 1.10 | 0.93 | 0.88 | 0.88 | 0.91 | 0.97 |

In the figures below we compare the magnitude of dNs predicted by the models produced by the two approaches to the estimates of the dNs that served as data. For injury accidents the comparison for approach 1 is in Figure F-36 and for approach 2 in Figure F-37.

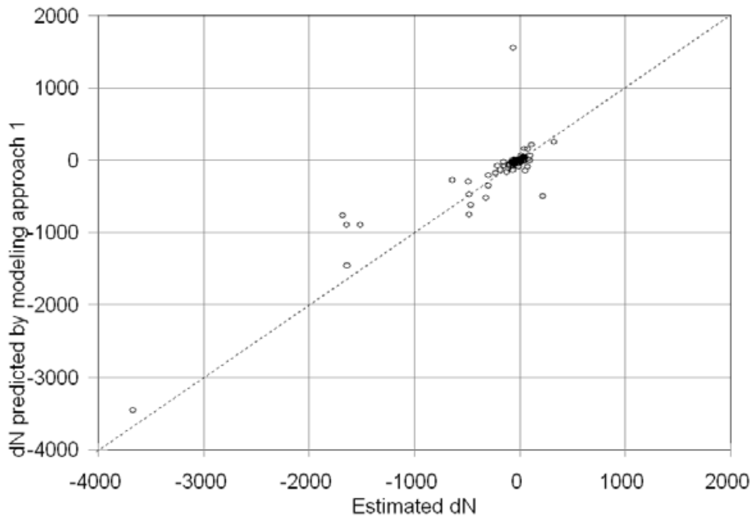


**Figure F-36. Injury Accidents, Modeling Approach 1**

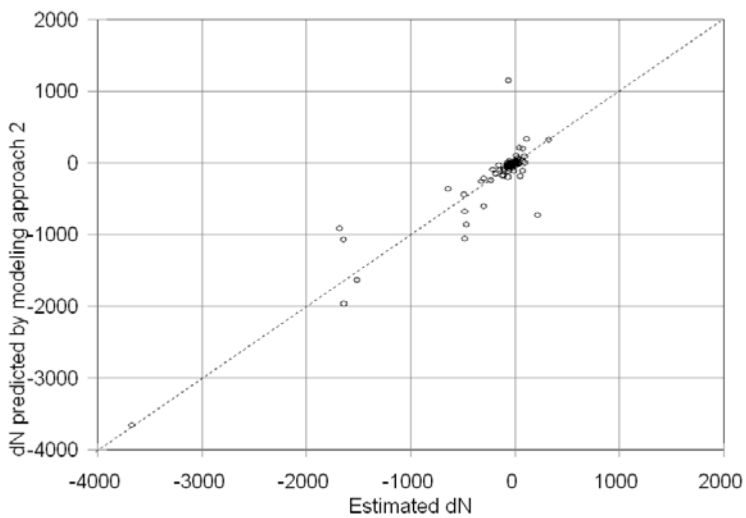


**Figure F-37. Injury Accidents, Modeling Approach 2**

For fatal accidents the comparison is in Figure F-38 and Figure F-39.

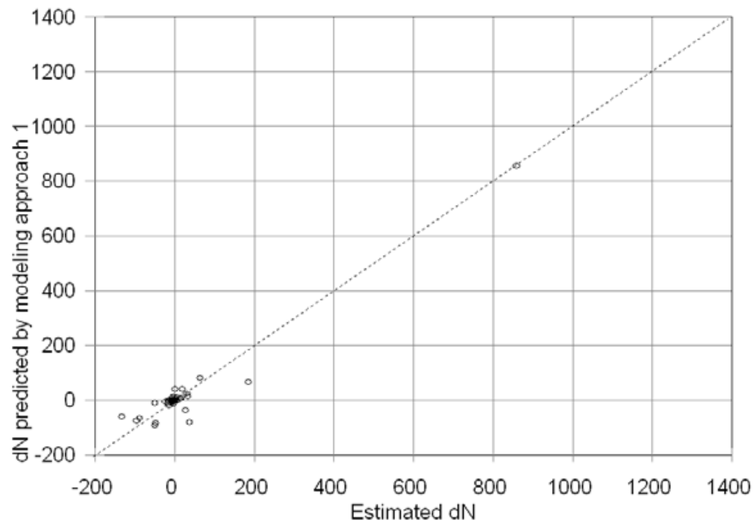


**Figure F-38. Fatal Accidents, Modeling Approach 1**

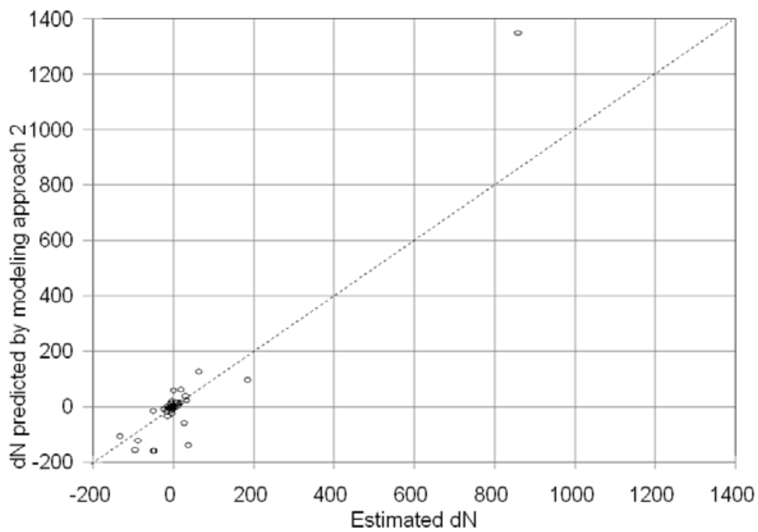


**Figure F-39. Fatal Accidents, Modeling Approach 2**

The results are very similar. For injury accidents the dNs predicted by modeling approach 1 (Figure F-36) are slightly biased (the equality line does not go through the centre of the points but is a bit too low). For fatal accidents the predictions by approach 2 seem somewhat better clustered around the equality line than those by approach 1. A direct comparison of the prediction quality by the two alternative approaches is not possible because approach 1 uses three estimated parameters while approach 2 uses two. The ratio of AMFs produced by the two approaches is usually[1] quite close to 1 and therefore both may be used.

## SUMMARY AND CONCLUSIONS

Nilsson (1984, 2005)[2] used the 'Power' model $N_1/N_0=(\bar{v}_1/\bar{v}_0)^{\alpha}$ to represent the relationship between accidents and mean speed. Elvik et al. (2004)[3] assembled a large dataset from 97 published studies containing 460 results about $\bar{v}_0$, $\bar{v}_1$, $N_0$ and $N_1$. To illustrate, the data about fatal accidents is shown in Figure F-40. In the large majority of cases when the mean speed increased, so did the number of fatal accidents and vice versa.

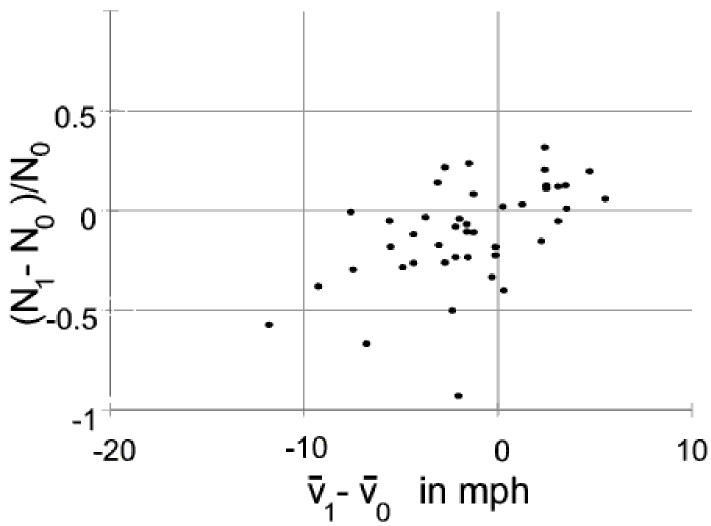


**Figure F-40 . Change in mean speed vs. relative change in fatal accidents**

Elvik et al. used Nilsson's power model. Their estimates of α are in Table F-18.

---

[1] The exception is the top left corner for fatal accidents in Table F-17.

[2] Nilsson, G., Hastigheter, olycksrisker och personskadekonsekvenser I olika vägmiljöer. VTI Report 277, Swedish Road and Traffic Research Institute.. 1984.
Nilsson, G., Traffic safety dimensions and the power model to describe the effect of speed in safety. Bulleting 221, Lund Institute of Technology, Department of Technology and Society, Traffic Engineering, Jund, Sweden. 2004

[3] Elvik R., P. Christensen and A. Amundsen, Speed and road accidents. TØI Report 740. Institute of Transport Economics, Oslo, 2004

**Table F-18. Estimates of α by Elvik et al.**

| Severity | Estimate of α | 95% Confidence Interval |
|---|---|---|
| Fatalities | 4.5 | 4.1-4.9 |
| Seriously Injured Road Users | 2.4 | 1.6-3.2 |
| Slightly Injured Road Users | 1.5 | 1.0-2.0 |
| All Injured Road Users (Including Fatally) | 1.9 | 1.0-2.8 |
| Fatal Accidents | 3.6 | 2.4-4.8 |
| Serious Injury Accidents | 2.0 | 0.7-3.3 |
| Slight Injury Accidents | 1.1 | 0.0-2.4 |
| All Injury Accidents (Including Fatal) | 1.5 | 0.8-2.2 |
| PDO Accidents | 1.0 | 0.0-2.0 |

In this report we used the data prepared by Elvik et al. (2004) to fit alternative model forms. Two different approaches to model fitting were used yielding very similar results. The average of results obtained by the two approaches is in Table F-19.

**Table F-19. Accident Modification Factors**

| **Injury Accidents** | $\overline{v}_0$ [mph] | | | | | |
|---|---|---|---|---|---|---|
| $\overline{v}_1 - \overline{v}_0$ [mph] | 30 | 40 | 50 | 60 | 70 | 80 |
| -5 | 0.57 | 0.66 | 0.71 | 0.75 | 0.78 | 0.81 |
| -4 | 0.64 | 0.72 | 0.77 | 0.80 | 0.83 | 0.85 |
| -3 | 0.73 | 0.79 | 0.83 | 0.85 | 0.87 | 0.88 |
| -2 | 0.81 | 0.86 | 0.88 | 0.90 | 0.91 | 0.92 |
| -1 | 0.90 | 0.93 | 0.94 | 0.95 | 0.96 | 0.96 |
| 0 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 1 | 1.10 | 1.07 | 1.06 | 1.05 | 1.04 | 1.04 |
| 2 | 1.20 | 1.15 | 1.12 | 1.10 | 1.09 | 1.08 |
| 3 | 1.31 | 1.22 | 1.18 | 1.15 | 1.13 | 1.12 |
| 4 | 1.43 | 1.30 | 1.24 | 1.20 | 1.18 | 1.16 |
| 5 | 1.54 | 1.38 | 1.30 | 1.26 | 1.22 | 1.20 |

| **Fatal Accidents** | $\overline{v}_0$ [mph] | | | | | |
|---|---|---|---|---|---|---|
| $\overline{v}_1 - \overline{v}_0$ [mph] | 30 | 40 | 50 | 60 | 70 | 80 |
| -5 | 0.22 | 0.36 | 0.48 | 0.58 | 0.67 | 0.75 |
| -4 | 0.36 | 0.48 | 0.58 | 0.66 | 0.73 | 0.80 |
| -3 | 0.51 | 0.61 | 0.68 | 0.74 | 0.80 | 0.85 |
| -2 | 0.66 | 0.73 | 0.79 | 0.83 | 0.86 | 0.90 |
| -1 | 0.83 | 0.86 | 0.89 | 0.91 | 0.93 | 0.95 |
| 0 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 1 | 1.18 | 1.14 | 1.11 | 1.09 | 1.07 | 1.05 |
| 2 | 1.38 | 1.28 | 1.22 | 1.18 | 1.14 | 1.10 |
| 3 | 1.59 | 1.43 | 1.34 | 1.27 | 1.21 | 1.16 |
| 4 | 1.81 | 1.59 | 1.46 | 1.36 | 1.28 | 1.21 |
| 5 | 2.04 | 1.75 | 1.58 | 1.46 | 1.36 | 1.27 |

To illustrate the use of these AMFs consider a road on which the mean speed is 60.0 mph. If the some measure is expected to increase the mean speed by 2.0 mph, injury accidents are expected to increase by a factor of 1.10 and fatal accidents by a factor of 1.18. Thus, what may appear to be a small change mean speed has a large impact on accidents.

Even though the data are international, the results were found to apply also to North American data. As is obvious from Figure F-40, there is considerable noise in the data. This noise reflects, in part, the randomness of accident counts, in part the variety of circumstances under which the data were obtained, and in part the variety of causes of change in mean speed. The question of whether these results would apply irrespective of the cause of the change in mean speed cannot be well answered at this time. If the change in accident frequency reflects mainly the associated change in severity then the AMFs in Table F-19 apply generally.

## REFERENCES


1. Elvik R., P. Christensen and A. Amundsen, Speed and road accidents. TØI Report 740. Institute of Transport Economics, Oslo, 2004 and Elvik, R. "Speed and Road Safety." *Transportation Research Record 1908.* Transportation Research Board, National Research Council, Washington, D.C., 2005, pp. 59-69.

2. Kockelman, K. *NCHRP Web-Only Document 90: Safety Impacts and Other Implications of Raised Speed Limits on High-Speed Roads - Final Report.* National Cooperative Highway Research Board, National Research Council, Washington, D.C., 2006.

3. Hauer, E. *Observational Before-After Studies in Road Safety.* Pergamon Press, Oxford, England, 1997.

4. Nilsson, G., Hastigheter, olycksrisker och personskadekonsekvenser I olika vägmiljöer. VTI Report 277, Swedish Road and Traffic Research Institute. 1984; and Nilsson, G., Traffic safety dimensions and the power model to describe the effect of speed in safety. Bulleting 221, Lund Institute of Technology, Department of Technology and Society, Traffic Engineering, Lund, Sweden. 2004