



Relating crash frequency and severity: Evaluating the effectiveness of shoulder rumble strips on reducing fatal and major injury crashes



Kun-Feng Wu^{a,*}, Eric T. Donnell^b, Jonathan Aguero-Valverde^c

^a Department of Transportation and Logistics Management, National Chiao Tung University, Taiwan

^b Department of Civil and Environmental Engineering, The Pennsylvania State University, United States

^c Programa de Investigación en Desarrollo Urbano Sostenible Universidad de Costa Rica, Costa Rica

ARTICLE INFO

Article history:

Received 8 November 2013

Received in revised form 8 February 2014

Accepted 8 February 2014

Available online 28 February 2014

Keywords:

Crash severity
Shoulder rumble strips
Fixed-effects model
Heavy trucks
Motorcycles
Pedestrians

ABSTRACT

To approach the goal of “Toward Zero Deaths,” there is a need to develop an analysis paradigm to better understand the effects of a countermeasure on reducing the number of severe crashes. One of the goals in traffic safety research is to search for an effective treatment to reduce fatal and major injury crashes, referred to as severe crashes. To achieve this goal, the selection of promising countermeasures is of utmost importance, and relies on the effectiveness of candidate countermeasures in reducing severe crashes. Although it is important to precisely evaluate the effectiveness of candidate countermeasures in reducing the number of severe crashes at a site, the current state-of-the-practice often leads to biased estimates. While there have been a few advanced statistical models developed to mitigate the problem in practice, these models are computationally difficult to estimate because severe crashes are dispersed spatially and temporally, and cannot be integrated into the *Highway Safety Manual* framework, which develops a series of safety performance functions and crash modification factors to predict the number of crashes. Crash severity outcomes are generally integrated into the *Highway Safety Manual* using deterministic distributions rather than statistical models. Accounting for the variability in crash severity as a function geometric design, traffic flow, and other roadway and roadside features is afforded by estimating statistical models. Therefore, there is a need to develop a new analysis paradigm to resolve the limitations in the current *Highway Safety Manual* methods. We propose an approach which decomposes the severe crash frequency into a function of the change in the total number of crashes and the probability of a crash becoming a severe crash before and after a countermeasure is implemented. We tested this approach by evaluating the effectiveness of shoulder rumble strips on reducing the number of severe crashes. A total of 310 segments that have had shoulder rumble strips installed during 2002–2009 are included in the analysis. It was found that shoulder rumble strips reduce the total number of crashes, but have no statistically significant effect on reducing the probability of a severe crash outcome.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

“Toward Zero Deaths” is a national strategy, developed by the Federal Highway Administration (FHWA), to develop a systematic approach to eliminate highway traffic fatalities. The primary effort focuses on developing countermeasures that directly impact highway safety through engineering, enforcement, education, and emergency medical services (4 E’s) (FHWA, 2012a). This goal can be translated into searching for effective treatments to reduce fatal and major injury crashes, referred to as severe crashes, through the

4E’s. To achieve this goal, the selection of promising countermeasures is of utmost importance, and relies on the effectiveness of candidate countermeasures in reducing severe crashes. Although it is important to precisely evaluate the effectiveness of candidate countermeasures in reducing the number of severe crashes at a site, the evaluation is challenging because severe crashes are dispersed spatially and temporally. The current state-of-the-practice, found in the American Association of State Highway and Transportation Officials’ (AASHTO) *Highway Safety Manual* (HSM) [AASHTO, 2010], does not offer a consistent approach to jointly consider crash frequency and severity in the safety prediction algorithms for two-lane, two-way roads, rural multi-lane highways, and urban and suburban arterials. For example, the two-lane, two-way roads safety prediction algorithm evaluates the reduction in the total number of crashes resulting from implementation of a

* Corresponding author. Tel.: +8863-571-2121.

E-mail addresses: kfwu@nctu.edu.tw (K.-F. Wu), edonnell@engr.psu.edu (E.T. Donnell), jonathan.aguero@ucr.ac.cr (J. Aguero-Valverde).

Table 1

The crash data for the hypothetical example in the before period.

	Before	After
Total number of crashes	100	50 (predicted)
Number of severe crashes	20	
Probability of a crash becoming a severe crash	20%	

countermeasure, and then multiplies the proportion of severe crashes among all crashes to approximate the reduction in severe crashes. This approach essentially assumes that the crash severity distribution remains constant before and after the implementation of a countermeasure, which may lead to biased estimates. A similar approach is used when considering various crash types in the rural multi-lane and urban and suburban arterials crash prediction algorithms. Advanced statistical methods, which combine crash frequency and severity (e.g., [Ma and Kockelman, 2006](#); [Aguero-Valverde and Jovanis, 2009](#); [Chiou and Fu, 2013](#)), have recently been developed to improve the precision of traffic safety countermeasure effectiveness by considering the association among different crash severity levels. However, obtaining robust safety estimates using these methods is challenging due to the low relative proportion of severe crashes among all crashes. This study proposes a simple approach to estimate the effectiveness of a countermeasure based on the number of severe crashes. This approach cannot only identify the sources associated with the change in severe crash outcomes resulting from countermeasure implementation, but also builds on the HSM framework ([AASHTO, 2010](#)). The proposed approach is demonstrated by evaluating the effectiveness of shoulder rumble strips on reducing the number of severe crashes using data from Pennsylvania.

1.1. The limitations of current practice and statistical approaches

In general, there are two approaches in the current state-of-the-practice (e.g., HSM method) to assess crash severity. The first approach applies a safety performance function (SPF) to predict severe crash frequency, and then multiplies the result by a crash modification factor (CMF) that represents the effect of a specific countermeasure in reducing the number of severe crashes. Since severe crashes are dispersed temporally and spatially, the CMFs and SPFs are usually accompanied by high standard errors (e.g. [Torbic et al., 2009](#)). As high standard errors reduce the reliability of the estimated reduction in severe crash frequency, an alternative approach is often applied instead.

The alternative approach is a simplification of the first approach. Although the results may seem to be more efficient (lower standard errors), they may be biased. Consider a hypothetical example for a site where a countermeasure is planned for implementation. The purpose of the countermeasure is to reduce the number of severe crashes, and the “before” period crash data are shown in [Table 1](#). The approach would initially predict the reduction in the total number of crashes using a SPF, before and after countermeasure implementation. For the purposes of this example, assume that the predicted total number of “after” period crashes is 50. Further, assume that based on historical, reported crash data that the number of severe crashes before the countermeasure was implemented is 20. Therefore, there is a 20 percent chance that a reported crash will result in a severe outcome during the before period. Because the proportion of severe crashes at this site is 20 percent based on historical reported crash data, the HSM would predict that the number of severe crashes after implementing a countermeasure to be $(100 - 50) \times 0.2 = 10$, because the change in the proportion of crash severity outcomes is not explicitly considered in the crash prediction algorithms. The HSM method would thus suggest that

Table 2

The crash data for the hypothetical example in the after period.

	Before	After
Total number of crashes	100	50 (observed)
Number of severe crashes	20	15 (observed)
Probability of a crash becoming a severe crash	20%	30% (observed)

both total and severe crashes are reduced by 50 percent (100 – 50 total crashes and 20 – 10 severe crashes).

This approach may over- or underestimate severe crash reductions resulting from countermeasure implementation. The current state-of-the-practice does not consider how severe crash probabilities may change due to the implementation of the countermeasure.

Continuing with this hypothetical example, suppose there are 15 severe crashes reported after the countermeasure has been implemented, and the severe crash proportion increases to 30 percent of total crashes. The severe crash reduction can be decomposed into $(100 - 50) \times 0.2 + 50 \times (0.2 - 0.3) = 10 - 5 = 5$, which shows that ten severe crashes decrease to five severe crashes as a result of countermeasure implementation ([Table 2](#)).

The HSM crash prediction algorithm employs SPFs and crash modification factors to estimate the expected number of severe crashes on rural two-lane, rural multi-lane, and urban/suburban arterials. Default proportions are then applied to estimate crash severity outcomes (see HSM, volume 2, Table 10-5). The HSM encourages users to adjust these default proportions based on crash data available at the study sites, but the HSM does not consider that the severity proportions may change before or after implementation of a countermeasure.

Although more advanced methods, which combine crash frequency and severity models, have recently been developed to utilize the associations among crash severity levels to more precisely estimate countermeasure effectiveness, it is difficult to obtain robust estimates from these advanced models due to the low proportion of severe crashes in crash data. Recent studies have found that there is a lack of independence in crash types or crash severities that constitute the total number of crashes ([Ma and Kockelman, 2006](#); [Park and Lord, 2007](#); [Ma et al., 2008](#); [Yannis et al., 2008](#); [Aguero-Valverde and Jovanis, 2009](#); [El-Basyouny and Sayed, 2009](#); [Ye et al., 2009](#)), which would lead to excess variation around fitted values that cannot be captured ([Berk and MacDonald, 2007](#)). Therefore, a multivariate Poisson log-normal model (MVPLN) has been proposed as a promising alternative for simultaneously modeling crash frequency in terms of different crash severity outcomes. [Chiou and Fu \(2013\)](#) advanced the MVPLN model, which considers the crash severity distribution, and takes both crash frequencies and severities into account. This is referred to as multinomial generalized Poisson (MGP) model. Although these two model types can be used to evaluate the effects that a countermeasure has on the number of severe crashes, these models are computationally difficult to estimate due to the low proportion of severe crashes in the total crash distribution. Neither the MVPLN nor MGP models can take event attributes into account (e.g., daytime/nighttime conditions), and they cannot be integrated into the HSM framework because these models are more generalized than the current HSM framework and require the correlation structure of different severe crash outcomes. Therefore, there is a need to develop a new analysis paradigm to resolve these challenges.

1.2. Evaluate the effectiveness of a countermeasure in reducing severe crashes

Consider [Fig. 1](#) as an illustration. A crash can be classified as a severe crash, or a less severe crash. In this illustration, a severe crash is defined as a fatal or major injury crash, while a less severe crash is

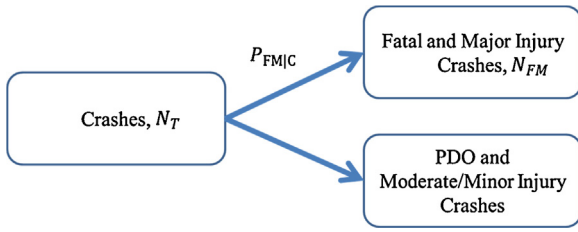


Fig. 1. Crash frequency and severity.

defined as a moderate injury or property damage only (PDO) crash. Therefore, on a roadway segment, the number of severe crashes, N_{FM} , can be decomposed into the number of crashes, N_T , times a proportion, $P_{FM|C}$, the probability of a crash being a severe crash. The latter variable ($P_{FM|C}$) is referred to as severe crash probability in the remainder of this article and shown in Eq. (1). In Eq. (1), application of N_{FM} as a safety measure consists of two components: crash frequency, N_T , and crash severity probability, $P_{FM|C}$.

$$N_{FM} = N_T \times P_{FM|C} \quad (1)$$

More insight can be obtained by decomposing Eq. (1) into Eq. (2), where $R=1$ indicates the implementation of a countermeasure and $R=0$ is the absence of a countermeasure. Eq. (2) applies the finite difference method to Eq. (1).

$$\begin{aligned} \Delta N_{FM} &= N_{FM|R=1} - N_{FM|R=0} = N_{T|R=1} \times P_{FM|C,R=1} - N_{T|R=0} \\ &\times P_{FM|C,R=0} = N_{T|R=1} \times P_{FM|C,R=1} - N_{T|R=1} \times P_{FM|C,R=0} \\ &+ N_{T|R=1} \times P_{FM|C,R=0} - N_{T|R=1} \times P_{FM|C,R=0} \\ &= N_{T|R=1}(P_{FM|C,R=1} - P_{FM|C,R=0}) \\ &+ P_{FM|C,R=0}(N_{T|R=1} - N_{T|R=0}) \end{aligned} \quad (2)$$

Therefore, Eq. (2) becomes:

$$\Delta N_{FM} = N_{T|R=1} \times \Delta N_{T|R=1} + P_{FM|C,R=0} \times \Delta N_T \quad (3)$$

As the previous derivation indicates, the effectiveness of the countermeasure on reducing severe crashes is influenced by $\Delta P_{FM|C}$ and ΔN_T . Since the magnitudes of $N_{T|R=1}$ and $P_{FM|C,R=0}$ are always non-negative, the effectiveness of the countermeasure on reducing severe crashes would be beneficial when both $\Delta P_{FM|C}$ and ΔN_T are negative. On the other hand, it is also possible to observe a severe crash reduction when ΔN_T is negative and $\Delta P_{FM|C}$ is positive, or when $P_{FM|C,R=0} \times \Delta N_T$ exceeds $\Delta N_{T|R=1} \times \Delta P_{FM|C}$, or vice versa; and when either $\Delta P_{FM|C}$ is zero and ΔN_T is negative, or vice versa. By contrast, the current version of the HSM essentially assumes $\Delta P_{FM|C}$ is zero, indicating that $P_{FM|C}$ does not change as a result of countermeasure implementation. Thus, ΔN_{FM} is solely determined by ΔN_T , as described in the earlier hypothetical example. Implicit in the HSM crash prediction algorithm is that the reduction in the total number of severe crashes is equivalent to the reduction in the product of $P_{FM|C,R=0}$ and ΔN_T , where $P_{FM|C,R=0}$ is a constant and does not change as a result of countermeasure implementation.

1.3. A case study: the effects of rumble strips on reducing severe crashes

Shoulder rumble strips have been shown to be effective in improving traffic safety (e.g., Griffith, 1999; Carrasco et al., 2004; Patel et al., 2007; Torbic et al., 2009). Shoulder rumble strips are generally considered a proven, effective safety countermeasure because this treatment provides drivers with auditory or tactile vibrations that facilitate recovery of roadway departure events.

A significant amount of research has been conducted concerning the effects of shoulder rumble strips in reducing the total number of crashes, the number of injury crashes, and the number of single vehicle run-off-road (SVROR) crashes (e.g. Torbic et al., 2009). Although these estimates are not entirely comparable to each other due to different crash types considered or the type of analysis employed (e.g. cross-sectional, naive before–after, or before–after Empirical Bayes (EB) analysis), it is difficult to refute the safety benefits attributed to shoulder rumble strips. Nevertheless, an important subsequent step is to determine if shoulder rumble strips are effective in reducing the frequency of severe crashes, which would then enable determination of how this treatment contributes to the FHWA Toward Zero Deaths strategy.

Table 3 is an excerpt from Torbic et al. (2009). Although the percent reductions for fatal and injury, and single vehicle run-off-road (SVROR) fatal and injury crashes are consistently negative, suggesting beneficial effects on reducing fatal and all levels of injury crashes, there are actually few estimates that are statistically significant (i.e., the confidence interval [plus and minus two standard errors (SE)] includes zero). This example illustrates the challenges of evaluating the effectiveness of shoulder rumble strips with regards to severe crashes. These challenges underscore the need to develop an advanced approach to more precisely estimate the effectiveness, and further explore the potential for shoulder rumble strips in reducing the frequency of severe crashes from $\Delta P_{FM|C}$ or ΔN_T .

1.4. Study objectives

The effects of a countermeasure on reducing severe crashes can be approached by examining the effects of the countermeasure on N_T and $P_{FM|C}$. This study proposes an approach that can be used to simultaneously consider the frequency and severity outcomes of crashes, and study the sources of severe crash reductions. More importantly, this approach can be built into the HSM framework. As discussed in Section 1.2, to supplement the existing method, the current CMFs are converted from crash frequency and severity models into the forms of $\Delta P_{FM|C}$ and ΔN_T . The next section describes the methodology used, a hybrid method that has been recently proposed to obtain more robust estimates compared to traditional approaches (e.g., Allison, 2009; Neuhaus and Kalbfleisch, 1998). Section 3 describes the data used in this study and relevant issues related to the data structure. A discussion of results follows with conclusions and suggestions for future research.

2. Methodology

Panel fixed-effects (FE) models have been recognized as robust statistical modeling methods for policy evaluation, since they provide consistent estimates of countermeasure effectiveness while controlling for unobserved effects (e.g., Hausman et al., 1984; Neuhaus and Kalbfleisch, 1998; Cameron and Trevidi, 2005; Angrist and Pischke, 2008; Allison, 2009; Hilbe, 2010; Wooldridge, 2010). Nevertheless, few researchers have adopted this specification to estimate countermeasure effectiveness in traffic safety (e.g., Houston, 1999; Greenstone, 2002; Wu et al., 2012, 2013). Law et al. (2009) used a panel negative binomial fixed effects model developed by Hausman et al. (1984) to analyze the relationship between motorcycle deaths and economic growth. Law et al. (2010) estimated the effect of per capita income and corruption on motor vehicle fatalities using a panel negative binomial fixed effects model.

The purpose of FE models is to condition out unobserved site-specific effects. For example, emergency medical services (EMS)

Table 3

The effectiveness of shoulder rumble strips on improving traffic safety, cross-referenced from Torbic et al. (2009).

	State	Number of sites	Fatal and injury		SVROR fatal and injury	
			Percent reduction	S.E.	Percent reduction	S.E.
Urban freeway	PA	53	-16.0*	7.2	-7.4	9.9
Rural freeway	Combined	47	-6.9	5.9	-17.1	7.3
	MO	29	-5.8	6.4	-15.6	8.2
	PA	18	-12.6	14.6	-23.2	15.7
Rural multilane divided highway (non-freeway)	Combined	25	-10.2	10.2	-2.6	13.5
	MN	6	-22.2	19.6	-10.3	28.6
	MO	14	-5.2	12.3	0.2	15.8
	PA	5	-40.1	42.5	-19.9	56.9
Rural two-lane road	Combined	53	-8.0	8.0	-36.4*	9.7
	MN	28	5.1	12.7	-32.4	17.6
	MO	5	-19.2	21.8	-44.6	23.2
	PA	20	-18.0	11.6	-36.7*	13.3

* Significant at 5 percent level of significance

response time to a crash, or roadside safety features present at a site are site-specific effects that are often not explicitly considered in crash frequency or severity models due to limited data availability. Consider two exactly identical injury crashes, one occurs in an area where it takes an hour for EMS to arrive at the scene; whereas, the EMS response time is 10 min in the other case. Although the authors do not have scientific evidence to prove the former is more likely to result in a more severe crash outcome than the other, EMS is undoubtedly a source of a site-specific effect. The same reasoning applies to the presence of roadside safety features, and how these features affect severe crashes. Although the FE model is rarely used due to data limitations (time invariant predictors cannot be included), other models such as the random-effects (RE) and mixed-effects models (ME), are widely used to consider site-specific effects (e.g. Gelman and Hill, 2007). This study seeks to apply a hybrid method, which incorporates the advantages of FE models while overcoming its limitation. This section first considers linear FE models (e.g. Houston, 1999; Greenstone, 2002; Wu et al., 2012) to help readers better understand the formulation. Since P_{FMIC} and N_T are categorical variables, Sections 2.2 and 2.3 discuss non-linear FE models. Section 2.4 introduces the specification of the hybrid method.

2.1. Panel fixed-effects (FE) models

Consider a linear FE model. For a safety measure on segment i at time t , y_{it} , the model formulation is as follows:

$$y_{it} = \alpha_i \mathbf{X}_{it} \beta + \varepsilon_{it} \quad (4)$$

for segment $i = 1, \dots, N$ and for each year $t, t = 1, \dots, T$. α_i represents site-specific effects, which measure unobserved heterogeneity that is possibly correlated with the regressors $\mathbf{X}_{it} = (x_{i1}, x_{i2}, \dots, x_{iN_i})$. The linear FE model is obtained by subtracting the time-average variables from the original model and by using the ordinary least squares estimator on the following:

$$(y_{it} - \bar{y}_i) = (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) \beta + (\varepsilon_{it} - \bar{\varepsilon}_i) \quad (5)$$

so the fixed-effect α_i is eliminated. In short panels, where T is less than 10 or 20, the estimated $\hat{\alpha}_i$ are inconsistent, but $\hat{\beta}$ is nonetheless consistent (e.g., Cameron and Trevisi, 2005).

The FE approach has two major limitations: (1) time-invariant covariates cannot be identified (i.e. $\mathbf{X}_{i2} - \mathbf{X}_{i1}$ will all be equal to one), and (2) segments without any crash records will be excluded from the model, resulting in larger standard errors (fewer observations results in lower efficiency). This specification leads to a loss of observations when y_{ij} is 0 for all J or y_{ij} is 1 for all J .

A RE model is another commonly used application in panel data analysis. The idea is to model all the site-specific effects, α_i , or the distribution that describes them, with an assumption that the site-specific effects are independent of other predictors. The former is often not desirable when the number of clusters is large due to the so-called incidental problem (Cameron and Trevisi, 2005), meaning that the variation of a model would be high when there are too many parameters to be estimated in a model, and would lead to inconsistent estimates. For the latter, RE models are sensitive to model misspecification such as making assumptions about the distribution of α_i (Cameron and Trevisi, 2005; Rabe-Hesketh and Skrondal, 2008). Moreover, RE models are only valid if a Hausman test does not reject the difference between RE and FE model with the same predictors (Hausman, 1978), which essentially tests whether the correlation among α_i and β would result in significantly different estimates between FE and RE models.

ME models are a “blend” of RE and FE models, and have recently been receiving more attention in the published literature (e.g., Cameron and Trevisi, 2005; Milton et al., 2008; Gkritza and Mannering, 2008; Hilbe, 2010). RE models, which are often referred to as random intercept models, do not allow correlation between α_i and β . On the other hand, ME models, which are often referred to as random-slope models, relax the assumption since this formulation essentially estimates site-specific slopes for each β . ME models allow the correlation between α_i and β by imposing a multivariate variance-covariance structure, and are therefore sensitive to the assumptions of the distributions (Train, 2009). In other words, ME models allow β to be modeled as random variables, but it is computationally intensive as there is no analytical closed-form solution for the log-likelihood function. Train (2009) provides more detailed mathematical formulations for these models. In summary, ME models balance the advantages and disadvantages of both RE and FE models, but the FE specification is more robust than the ME models. As will be discussed in Section 2.4, a hybrid approach has been proposed recently to take both the spirit of the FE models, and the flexibility of the ME models, into account (Allison, 2009; Neuhaus and Kalbfleisch, 1998).

2.2. Investigating the effects of shoulder rumble strips on P_{FMIC} using FE models

For crash severity modeling, let the severe crash probability for crash j on segment i be modeled as:

$$\begin{aligned} \text{logit } \Pr(y_{ij} = 1 | \alpha_i, \mathbf{X}_{ij}) &= \log \left(\frac{\Pr(y_{ij} = 1 | \alpha_i, \mathbf{X}_{ij})}{1 - \Pr(y_{ij} = 1 | \alpha_i, \mathbf{X}_{ij})} \right) \\ &= \alpha_i + \beta \mathbf{X}_{ij} \end{aligned} \quad (6)$$

where $Pr(y_{ij} = 1|\alpha_i, \mathbf{X}_{ij})$ is the probability that the dependent variable, y_{ij} , a binary variable, is equal to one, which is the probability that a crash i on segment j is a severe crash. $Pr(y_{ij} = 1|\alpha_i, \mathbf{X}_{ij})/(1 - Pr(y_{ij} = 1|\alpha_i, \mathbf{X}_{ij}))$ is referred to as severe crash odds. The transformation from probability to odds is a monotonic transformation, meaning the odds increase as the probability increases, or vice versa. The unobserved site-specific effect, α_i , changes only across segments. \mathbf{X}_{ij} represents observed covariates for both event attributes and segment characteristics. Consider the simplest case-control study, where one observation is a case (severe crash), and the other is a control (moderate/minor injury crash). Condition on $y_{i1} + y_{i2} = 1$, so that $y_{ij} = 1$ in exactly one of the two crashes. Then,

$$Pr(y_{i1} = 0, y_{i2} = 1|y_{i1} + y_{i2} = 1) = \frac{Pr(y_{i1} = 0, y_{i2} = 1)}{Pr(y_{i1} = 0, y_{i2} = 1) + Pr(y_{i1} = 1, y_{i2} = 0)} \quad (7)$$

Now let $Pr(y_{i1} = 0, y_{i2} = 1) = Pr(y_{i1} = 0) \times Pr(y_{i2} = 1)$, assuming that y_{i1} and y_{i2} are independent given α_i and \mathbf{X}_{ij} . Based on the logistic formula, the following is obtained:

$$Pr(y_{i1} = 0, y_{i2} = 1) = \frac{1}{1 + \exp(\alpha_i + \beta\mathbf{X}_{ij})} \times \frac{\exp(\alpha_i + \beta\mathbf{X}_{ij})}{1 + \exp(\alpha_i + \beta\mathbf{X}_{ij})} \quad (8)$$

$$Pr(y_{i1} = 1, y_{i2} = 0) = \frac{\exp(\alpha_i + \beta\mathbf{X}_{ij})}{1 + \exp(\alpha_i + \beta\mathbf{X}_{ij})} \times \frac{1}{1 + \exp(\alpha_i + \beta\mathbf{X}_{ij})} \quad (9)$$

Substitute Eqs. (8) and (9) into Eq. (7), the denominators cancel, and the following expression is obtained:

$$Pr(y_{i1} = 0, y_{i2} = 1|y_{i1} + y_{i2} = 1) = \frac{\exp(\beta(\mathbf{X}_{i2} - \mathbf{X}_{i1}))}{1 + \exp(\beta(\mathbf{X}_{i2} - \mathbf{X}_{i1}))} \quad (10)$$

Therefore, the within segment conditioning eliminates the unobserved individual effect α_i by conditioning on $\sum_{j=1}^J y_{ij} = 1$. More generally, with up to J crashes, we can eliminate α_i by conditioning on $\sum_{j=1}^J y_{ij} = 2, \dots, \sum_{j=1}^J y_{ij} = J - 1$. The resulting conditional model is a logit model with the regressor $\mathbf{X}_{i2} - \mathbf{X}_{i1}$.

2.3. Investigating the effects of shoulder rumble strips on N_T using FE models

To model the total number of crashes, and the number of severe crashes, the negative binomial (NB) formulation is known to be useful for handling dispersion in count data, which is usually overdispersed. Therefore, it is intuitive to apply a fixed-effects negative binomial (FENB) model to model the number of crashes of interest. The model begins with the number of severe crashes on segment i in year t , y_{it} , which obeys a negative binomial distribution.

$$\ln(y_{it}) = \alpha_i + \beta\mathbf{X}_{ij} \quad (11)$$

For the FENB model, the joint probability of the counts for each segment is conditioned on the sum of the counts for the segment in the study period (i.e. the observed $\sum_{t=1}^{n_i} y_{it}$) (Hausman et al., 1984), and therefore, this FENB model is also referred to as a conditional FENB. Once conditioning on the count total $\sum_t y_{it}$ for each segment, a sufficient statistic of α_i , this yields a conditional likelihood that is proportional to Eq. (12), which also leads to the same estimator of β as the FE negative multinomial model.

$$\prod_i \prod_t \left(\frac{\exp(\beta\mathbf{X}_{it})}{\sum_S \exp(\beta\mathbf{X}_{is})} \right) y_{it} \quad (12)$$

2.4. Hybrid methods

A hybrid method has been proposed to incorporate the advantages of both the spirit of the FE formulation and the flexibility of the ME models (e.g., Allison, 2009; Neuhaus and Kalbfleisch, 1998). The hybrid method incorporates the advantages of both FE and ME model specifications, and has been proven to be able to obtain the same estimates as those obtained from the FE models through simulation studies (e.g., Allison, 2009; Neuhaus and Kalbfleisch, 1998). In other words, this approach incorporates not only the robustness of FE models, but also the efficiency of RE models and the flexibility of the ME models. This method was implemented by decomposing each time-varying independent variable into a within-group and between-group comparison, and then fitting a ME or RE model with both components. The between-group component is the group-specific mean of the variables, and the within-group component is the derivation from that group-specific mean.

Formally, the covariate \mathbf{X}_{ij} are decomposed into a between component, $\bar{\mathbf{X}}_i = n_i^{-1} \sum_{j=1}^{n_i} \mathbf{X}_{ij}$, where n_i is the number of observations in a cluster, and within-cluster components, $(\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)$. Eqs. (6) and (11) now become Eqs. (13) and (14), respectively. After transforming the independent variable of interest into deviations from their group-specific means, a ME model is then applied to estimate the relevant parameters.

$$\text{logit}Pr(y_{ij} = 1|\alpha_i, \mathbf{X}_{ij}) = \alpha_i + \beta_B \bar{\mathbf{X}}_i + \beta_W (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i) \quad (13)$$

$$\ln(y_{it}) = \alpha_i + \beta_B \bar{\mathbf{X}}_i + \beta_W (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i) \quad (14)$$

Conventional logistic and negative binomial regressions assume that $\beta_B = \beta_W$ in Eqs. (13) and (14), and the FE models focus on modeling β_W by using data from clusters with discordant outcomes; on the other hand, the hybrid models do not require $\beta_B = \beta_W$ and also take $\bar{\mathbf{X}}_i$ into modeling consideration, and hence provide a more unified approach. For those conventional models, such as standard logistic regression, RE models, and ME models, the estimated coefficients are weighted averages of the between and within coefficients, β_B and β_W . The hybrid model can also be used to validate the use of a conventional RE model, which implicitly assumes that $\beta_B = \beta_W$ in Eqs. (13) and (14). A joint test that all deviation coefficients are equal to the corresponding mean coefficients, a joint test for all the pairs of β_B and β_W , can be used to verify the critical assumption of the RE model.

Although the hybrid model includes $\bar{\mathbf{X}}_i$ as part of modeling process, the β_B in non-linear ME models are difficult to interpret since the coefficients are associated with cluster-level covariates. For example, consider logistic regression. Because Eq. (13) measures covariate effects conditional on the random effect, α_i , β_B actually measures differences on a logit scale among segments that share the same random effect. Hence, the probability, by inserting aggregate values of the explanatory variables, is not equal to the average probability due to non-linearity (Neuhaus and Kalbfleisch, 1998; Neuhaus et al., 1991; Train, 2009). On the other hand, the coefficients of β_W suggest that for a given segment, the crash odds will differ by β_W units between two crashes that differ by one unit on \mathbf{X}_{ij} . Allison (2009) and Neuhaus and Kalbfleisch (1998) provide a full derivation and discussion of this model.

2.5. Marginal effects of a countermeasure (evaluating ΔP_{FMIC} and ΔN_T)

Provided that the models discussed above are nonlinear and that R is a binary variable, evaluating ΔP_{FMIC} and ΔN_T is the same as computing the marginal effects for both logistic and count regressions. It should be noted that even though the within-cluster component has many different values (because of subtracting a segment-specific mean), the coefficient is interpreted as if the

Table 4
Descriptive statistics.

Variable	Unit	Mean	S.D.	Min	Max
<i>Predictors in severity analysis (analysis unit : crash; observation = 4465)</i>					
Fatal and major injury crashes	Proportion (1 = yes)	0.05	0.22	0	1
Shoulder rumble strips	Proportion (1 = yes)	0.68	0.47	0	1
Heavy truck involvement	Proportion (1 = yes)	0.07	0.26	0	1
Motorcycle involvement	Proportion (1 = yes)	0.03	0.17	0	1
Bicyclist/pedestrian involvement	Proportion (1 = yes)	0.01	0.12	0	1
Daylight condition	Proportion (1 = yes)	0.68	0.47	0	1
Head-on crash	Proportion (1 = yes)	0.04	0.20	0	1
Number of occupants unbelted	Persons	0.19	0.54	0	6
Speed limit	MPH	42.44	7.93	25	65
Driver over 65 years of age	Proportion (1 = yes)	0.16	0.37	0	1
<i>Predictors in frequency analysis (analysis unit: segment; observation = 2168)</i>					
Total number of crashes	Crashes	2.07	2.72	0	26
Shoulder rumble strips	Proportion (1 = yes)	0.68	0.47	0	1
AADT	Vehicles	9732	5438	582	22,607
Segment length	Feet	2365	685	322	4296
Rural area	Proportion (1 = yes)	0.62	0.48	0	1
Right roadside hazard rating greater than 4 and a horizontal curve	Proportion (1 = yes)	0.22	0.41	0	1
Divided median	Proportion (1 = yes)	0.99	0.11	0	1
Other expressway and principal arterials	Proportion (1 = yes)	0.31	0.46	0	1
Minor arterials	Proportion (1 = yes)	0.58	0.49	0	1
Collectors	Proportion (1 = yes)	0.11	0.32	0	1
Local roads	Proportion (1 = yes)	0.003	0.06	0	1

variable is still a dummy variable. That is, the effect of R is evaluated from zero to one (Allison, 2009). The marginal effect of R in a logistic regression is computed as the discrete change in the expected probability for a change in R :

$$\Delta P_{FM|C}(y = 1|\alpha_i, \mathbf{X}_{ij}, R) = P_{FM|C}(y = 1|\alpha_i, \mathbf{X}_{ij}, R = 1) - P_{FM|C}(y = 1|\alpha_i, \mathbf{X}_{ij}, R = 0) \quad (15)$$

The marginal effect of R in a count regression is computed as the discrete change in the expected count for a change in R :

$$\Delta E(y|\alpha_i, \mathbf{X}_{ij}, R) = E(y|\alpha_i, \mathbf{X}_{ij}, R = 1) - E(y|\alpha_i, \mathbf{X}_{ij}, R = 0) \quad (16)$$

The present study adopted the average marginal effect (AME) approach, since AME has been shown to be more appropriate for providing a realistic interpretation of estimation results than the marginal effects at the mean (MEM) approach (e.g. Train, 2007; Hilbe, 2010). The AME computes the average of discrete or partial changes over all observations; whereas, the MEM computes the marginal effects at fixed values of the independent variables, where the most often used values are sample means. To account for the uncertainty in the estimated coefficients in both Eqs. (13) and (14), both the expected probabilities and expected counts in the right-hand side of Eqs. (15) and (16) are evaluated using the delta method (Long, 1997), an approximation appropriate in large samples, as well as Eq. (3).

3. The data

This study includes 310 segments in Pennsylvania, covering the period 2002–2009 (inclusive). Shoulder rumble strips were installed during 2004 and 2006, including 269 in 2004 and 41 in 2006. Rumble strip installations in Pennsylvania are often done based on geographic region. A single contractor often installs the same rumble strip pattern throughout the region (several different road segments) during a construction season, which covers the spring, summer, and fall periods (e.g., April through October). As such, it is difficult to determine the precise date that rumble strips are installed; rather, it is often only possible to identify the year when the installation was completed. Therefore, the authors have eliminated the entire year of data from the analysis file at

locations where rumble strips were installed during the analysis period. For example, if rumble strips were installed during the 2005 construction season, the traffic volume, crash, and other analysis data were compiled for the years 2002–2004 (before period) and for the period 2006–2008 (after period).

There were 5629 reported crashes in total on the study segments during the analysis period, of which 4 percent were categorized as fatal and major injury, 44 percent moderate/minor injury, and 52 percent property damage only (PDO) crashes. The Pennsylvania Department of Transportation (PennDOT) defines crash severity levels as fatal, major injury, moderate injury, minor injury, and PDO, which are analogous to the KABCO scale. In this study, we refer to fatal and major injury crashes as severe crashes. For severity modeling, the analysis unit in this study is at the crash-level, and therefore, the crash severity is based on the most severe outcome reported in a crash.

Among the study segments, about 59 percent of the segments included curve sections with an equal number of left- and right-hand curves. The presence of sharp curves (based on advanced curve warning signs) and roadside hazard rating (RHR) information for the study segments were collected using PennDOT online video photo logs. It was found that 21 percent of the curves in the study segments were designated as sharp curves. Please refer to Zegeer et al. (1988) for more details concerning the RHR rating scale. Approximately 42 percent of the segments had a rating of 3 or higher (marginally recoverable with a side slope of 1 V:3H or 1 V:4H) for the right side of the roadway and 38 percent for the left side of the roadway. Less than 1 percent of the segments had RHR higher than 4 for both the left- and right-hand sides of the roadway.

The data shows that 42.5 percent of the total crashes occurred in rural areas, which includes 49 percent of the reported fatal and severe injury crashes. Approximately 3.2 percent of the reported crashes involved motorcycles and 7.1 percent involved heavy trucks. Less than 1 percent of crashes involved a pedestrian and 17 percent involved drivers 65 years or older. 35 fatal or severe injury crashes were reported to have occurred during adverse weather conditions such as rain, fog, sleet, etc. Another 35 percent of the crashes occurred during conditions other than daylight (dark with no street lights, dark, dawn, dusk, etc.). The most frequent collision types reported for crashes included head-on, angle or hit fixed object crashes. Occupants of the vehicle were unbelted in about

Table 5
 P_{FMIC} , N_T , and N_{FM} on study segments between 2002 and 2009.

Year	AADT	N_T	P_{FMIC} (%)	N_{FM}
2002	9513	633	4.11	26
2003	9553	696	5.46	38
2004	9840	640	4.22	27
2005	10,055	630	5.08	32
2006	10,028	665	5.11	34
2007	9821	650	5.08	33
2008	9561	606	4.46	27
2009	9553	620	4.84	30

14 percent of reported crashes. Approximately 10 percent of the crashes with unbelted occupants resulted in a fatal or major injury crash. More than 70 percent of the crashes occurred on segments with a posted speed limit of 45 mph or above. The average annual daily traffic (AADT) before installation of rumble strip was 9430.8 vehicles per day and after installation was 9863.4 vehicles per day. Please refer to Table 4 for descriptive statistics of the predictors included in this study.

Table 5 shows the year-wise P_{FMIC} , N_T , and N_{FM} during the analysis period. As shown, there is an increase in N_{FM} during 2007–2008. Note that these decompositions were computed without controlling for any confounding factors. The next section will evaluate the effects of installing shoulder rumble strips on these decompositions.

4. Data analysis

This section evaluates the effects of shoulder rumble strip installation on P_{FMIC} and N_T using the models described in Section 2 of this paper. Once the models described in Section 2.4 are fit to the data, the signs and magnitudes of ΔP_{FMIC} and ΔN_T in Eq. (3) can be obtained. To demonstrate without losing generality, although ME models are applied in this study, the β_B were not modeled as random variables. All the variable names beginning with “M” refer to segment-specific means, and all the variable names beginning with “D” refer to deviations from those means. The coefficients for the deviation variables are functionally equivalent to fixed-effects coefficients, because they are estimated using only within-group variation and therefore control for all stable predictors (Allison, 2009). The M-variables, the estimated mean coefficients, indicate variability across segments in the effects of the predictors, but they are difficult to interpret and are not the focus of this study. Hence, only the D-variables will be interpreted. Section 4.1 presents the model for P_{FMIC} , and Section 4.2 presents the model for N_T .

4.1. The effects of shoulder rumble strips on P_{FMIC}

A Wald test to jointly test whether all seven deviation coefficients are equal to the corresponding mean coefficients is first performed to test the assumption of $\beta_B = \beta_W$. The joint test, as shown in the bottom row of Table 6, clearly indicates a need to reject the applicability of RE models (p -value = 0.0018), suggesting that the α_i are correlated with other predictors so that RE estimates are biased, and that a FE model approach is superior to a RE approach.

There is no evidence that the presence of shoulder rumble strips would affect P_{FMIC} (p -value = 0.417). The marginal effect of installing shoulder rumble strips is estimated as a mean of 0.0056 with standard deviation of 0.0066, and is not statistically significant with regards to P_{FMIC} . This result is not unexpected. Past research has found that the effects of shoulder rumble strips on crash severity are ambiguous. Several studies suggest that a driver may panic when running over shoulder rumble strips, and subsequently swerve to hit another vehicle, causing multiple-vehicle

crashes, and hence increase crash severity (Griffith, 1999; Smith and Ivan, 2005; Geedipally et al., 2014). Conversely, some studies show that shoulder rumble strips are helpful in reducing crash severity (e.g., Sayed et al., 2010). Geedipally et al. (2014) suggest that the inconsistent results are confounded because shoulder rumble strips are likely to be installed at locations where severe crash outcomes are high; therefore, the results from cross-sectional studies may report that crashes at locations with shoulder rumble strips, resulting in overestimating the reduction in severe crash outcomes attributed to this safety countermeasure. Another possibility is that the installation is a routine practice among state transportation agencies, thus the effects of shoulder rumble strips are aggregated so that the results are “mixed.” Because the hybrid model can mitigate the issues discussed above, this study concludes that there is no evidence showing that the installation of shoulder rumble strips increases the probability of a severe crash outcome.

In terms of statistical significance, other factors that are associated with P_{FMIC} include heavy truck, motorcyclist, and bicyclist/pedestrian involvements, crash types, daylight conditions, seatbelt usage, and posted speed limit. The involvement of a heavy truck in a crash increases the odds of a crash being severe by 2.4 ($\exp(0.874)$), or by 140 percent ($(2.4 - 1) * 100$). Similarly, the involvement of a bicyclist/pedestrian or motorcycle increases the severe crash odds by 16.5 and 24.2 times, respectively. Occupant seatbelt use and age indicated the vulnerability of vehicle occupants during a crash, and both signs are positive and statistically significant. Head-on crashes have been found to be the type of crash that often results in severe outcomes, and is consistent with the results in past research (e.g. Kockelman and Kweon, 2002; Lenguerand et al., 2006). Although one would expect that drivers have better visibility, thus resulting in improved safety performance, the effects of daylight on crash severity are somewhat ambiguous in the literature. For example, Kockelman and Kweon (2002) adopted an ordered probit formulation, and found that without distinguishing the number of vehicles involved in a crash, more severe single-vehicle crashes tend to occur in nighttime conditions when compared to daytime conditions (p -value = 0.2071). With differentiation, more severe single-vehicle crashes tend to occur in daytime conditions than at night (p -value = 0.1244); whereas, more severe two-vehicle crashes tend to occur at night relative to the daytime (p -value = 0.1525).

One of the advantages of the hybrid approach considered in the present study is the ability to keep time-constant predictors such as speed limits and roadside hazard ratings in the model. Segments with higher speed limits and roadside hazard ratings have higher P_{FMIC} . The combined effect of left RHR greater than 4 and the presence of a horizontal curve was not found to significantly affect P_{FMIC} . Although roadside hazard rating was found to be positively associated with crash severity, the posted speed-limit is the only roadway feature that was found to achieve statistical significance. Crashes occurring on segments with higher posted speed limits were found to increase the probability of a severe crash outcome. Every 5 mph increase in the posted speed-limit was estimated to increase the probability of a severe crash by 1.22 times.

4.2. The effects of shoulder rumble strips on N_T and N_{FM}

To estimate count regression models, all of the crash records need to be summarized by segment on a yearly basis. Note that event attributes such as the individual(s) involved in the crash, and daylight conditions are aggregated, which is a limitation of a count regression.

For the N_T model, the joint test, as shown in the bottom row of Table 7, indicates that the RE model can be rejected, suggesting that a FE model approach is superior to a RE approach. The installation of shoulder rumble strips was estimated to reduce the total

Table 6
The effects of shoulder rumble strips on P_{FMC} .

	Coeff.	S.E.	P-value
D.Shoulder rumble strips	0.149	0.183	0.417
D.Motorcycle involvement	3.185***	0.280	<0.001
D.Byclist/pedestrian involvement	2.795***	0.455	<0.001
D.Daylight condition	-0.424*	0.179	0.018
D.Head-on crash	2.429***	0.247	<0.001
D.Number of occupants unbelted	0.854***	0.098	<0.001
D.Driver over 65	1.094***	0.214	<0.001
D.Heavy truck involvement	0.874**	0.270	0.001
Speed limit	0.039**	0.012	0.001
Left roadside hazard rating greater than 4 and a horizontal curve	0.092	0.228	0.685
M.Shoulder rumble strips	0.207	0.578	0.720
M.Motorcycle involvement	1.112	1.532	0.468
M.Byclist/pedestrian involvement	-6.861	3.592	0.056
M.Daylight condition	-0.061	0.630	0.923
M.Head-on crash	2.127	1.605	0.185
M.Number of occupants unbelted	2.204***	0.486	<0.001
M.Driver over 65	-1.363	0.865	0.115
M.Heavy truck involvement	1.006	1.179	0.393
Constant term	-5.832***	0.877	<0.001
Number of crashes	4465		
Constant term only log-likelihood	-879.734		
Convergent log-likelihood	-679.508		
Combined test, Chi-square test (<i>p</i> -value)	24.66 (0.0018)		

* Significant at 10% level.

** Significant at 5% level.

*** Significant at 1% level.

number of crashes by seven percent ($1 - \exp(-0.072)$). This estimate is consistent with the 6.5 percent estimate reported by Torbic et al. (2009), which applied the empirical Bayes (EB) method with a smaller sample of sites in Pennsylvania. The marginal effect of the installation of shoulder rumble strips was estimated as a mean of -0.1479 with standard deviation of 0.0814. All the signs and magnitudes of the other variables are consistent with past research (e.g., FHWA Crash Modification Factors Clearinghouse, 2013). The study sites in rural areas have significantly fewer crashes than those in urban areas (13 percent), possibly due to fewer access points.

Minor arterials, collectors, and local roads have higher crash rates than expressways when controlling for other variables such as traffic volume, segment length, area type, and roadway features, as shown in Table 7. The results support national crash statistics (FHWA, 2012b) which indicate that fatal crash rates are higher on lower functional class roads (i.e., local roads and collectors) when

compared to fatal crash rates on higher functional class roads (i.e., arterials, expressways, and interstates). The sample of local roads with rumble strips in this study is comparatively small relative to other road classes, because local roads with rumble strips are not a significant part of the sample as these road types do not often contain rumble strips due to lower traffic volumes, low travel speeds, and short trips durations relative to higher functional class roads. Separate models without the local road data were estimated, but the regression coefficients did not change the interpretation of the model.

The N_{FM} model is reported in Table 7. The coefficient of installing shoulder rumble strips is estimated as a mean of -0.004 with standard deviation of 0.148. Not surprisingly, except for segment length (exposure variable), none of the predictors are statistically significant, including the presence of shoulder rumble strips. The N_{FM} model indicates that the installation of shoulder rumble

Table 7
The effects of shoulder rumble strips on N_T and N_{FM} .

	N_T			N_{FM}		
	Coef.	S.E.	P-value	Coef.	S.E.	P-value
D.Shoulder rumble strips	-0.072*	0.040	0.070	-0.004	0.148	0.976
D.In(AADT)	0.082	0.186	0.660	0.286	0.784	0.715
ln(segment length)	0.814***	0.126	<0.001	0.853***	0.272	0.002
Rural area	-0.138**	0.064	0.030	NA		
Right roadside hazard rating greater than 4 and a horizontal curve	0.106	0.102	0.297	0.097	0.192	0.615
Divided median	-0.864**	0.385	0.025	Baseline		
Minor arterials	0.132	0.096	0.170			
Collectors	0.337**	0.142	0.018			
Local roads	2.659***	0.698	<0.001			
Other expressway and principle arterials	Baseline	0.233	0.180	0.194		
M.Shoulder rumble strips	1.007**	0.433	0.020	0.008	0.823	0.992
M.In(AADT)	0.959***	0.069	<0.001	0.422***	0.134	0.002
Constant	-12.54***	1.230	<0.001	2.714	505.26	0.996
Sample size	2168	2168				
Constant term only log-likelihood	-3656	-718				
Convergent log-likelihood	-3546	-710				
Combined test, Chi-square test (<i>p</i> -value)	27.02 (<0.001)	0.03(0.98)				

* Significant at 10% level.

** Significant at 5% level.

*** Significant at 1% level.

Table 8
Summary of Eq. (3).

ΔN_{FM}	$N_{TR=1} \times \Delta P_{FMIC} + P_{FMIC,R=0} \times \Delta N_T$
	$N_{TR=1} = 2.07$
	$P_{FMIC,R=0} = 0.04898$
ΔN_{FM} (S.E.) = -00046 (0.015)	ΔP_{FMIC} (S.E.) = 0.0056 (0.0066)
	ΔN_T (S.E.) = -0.1479 (0.0814)
95% confidence interval = (-0.03, 0.029)	95% confidence interval = (-0.027, 0.036)

strips did not significantly reduce the number of severe crashes (p -value = 0.98). This finding is also consistent with Torbic et al. (2009), as shown in Table 3.

4.3. The relationship between P_{FMIC} , N_T , and N_{FM}

This section compares the results obtained from the N_{FM} model, the left-hand side of Eq. (3), to that from the right-hand side of Eq. (3). For the right-hand side, as shown in Table 8, except for $N_{TR=1} = 2.07$ and $P_{FMIC,R=0} = 1.04898$, the values of ΔP_{FMIC} and ΔN_T are estimated using the marginal effects of ΔP_{FMIC} and ΔN_T in terms of R . From the results in Sections 4.1 and 4.2, ΔP_{FMIC} is estimated as a normal distribution with a mean of 0.0056 and a standard error of 0.0066, and ΔN_T is estimated as a normal distribution with a mean of -0.1479 with a standard error of 0.0814. Altogether, the right-hand side of Eq. (3) was estimated as a mean of .0044 with a standard error of 0.014, as shown in Table 8.

With regard to the marginal effect of N_{FM} , ΔN_{FM} is estimated as a mean of -0.00046 and a standard error of 0.015. The 95 percent confidence intervals for the left- and right-hand side of Eq. (3) are (-0.03, 0.029) and (-0.027, 0.036), respectively, as shown in the last row of Table 8. The empirical results show that the confidence intervals overlap, consistent with the derivation of Eq. (3). Both of the results indicate that shoulder rumble strips do not significantly reduce the number of severe crashes.

5. Conclusion and discussion

To move toward the goal of Toward Zero Deaths, there is a need to develop an analysis paradigm to better understand the effects of a countermeasure on reducing the number of severe crashes. This study first showed that the reduction in the number severe crashes is not only associated with the reduction in the total number of crashes but also the reduction in the severe crash probability, as derived in Eq. (3). Eq. (3) also presents a way to connect the relationship between the total number of crashes, N_{FM} , the number of severe crashes, P_{FMIC} , and severe crash probability, N_{FM} , which simultaneously consider the frequency and severity outcomes of crashes. The crash frequency and severity models are constructed using the hybrid method. This method not only provides robust estimates as FE models, but also incorporates the advantages of the ME models, as discussed in Section 2 of this paper. The implementation of shoulder rumble strips in Pennsylvania was used as an example to demonstrate the proposed approach.

Eq. (3) shows how the sources of severe crash reduction can be decomposed and matched. Key findings of the hybrid models for both crash frequency and severity are:

- The factors that are associated with severe crash probability include heavy truck, motorcyclist, and bicyclist/pedestrian involvements, crash types, daylight conditions, seatbelt usage, and posted speed limit. There is no evidence, however, that the presence of shoulder rumble strips affect severe crash outcomes.
- The installation of shoulder rumble strips was estimated to reduce the total number of crashes by seven percent. This estimate is consistent with the 6.5 percent estimate reported by

Torbic et al. (2009), which applied the empirical Bayes (EB) method with a smaller sample of sites in Pennsylvania. The hybrid, cross-sectional modeling approach used in this study estimated a countermeasure safety effect consistent with the EB method. To further determine if the hybrid approach used in this paper consistently produces results that compare to the EB method, additional countermeasure safety treatments should be evaluated.

- Although shoulder rumble strips are beneficial in reducing the total number of crashes, there is no evidence in the present study to indicate that they effectively reduce severe crash outcomes after controlling for various factors. Nevertheless, shoulder rumble strips were found to be beneficial in reducing the total number of crashes, consistent with past research.

Although the derivation in Eq. (3) and the empirical results are consistent, additional studies are needed to confirm this relationship. Nevertheless, this study serves as a first step toward an approach to simultaneously consider crash frequency and severity in traffic safety research. As discussed above, it is clear that to precisely estimate the effects of a countermeasure on the number of severe crashes, information about the change in the total number of crashes, and severe crash probability are required. Most importantly, it is crucial to evaluate the number of severe crashes in terms of the total number of crashes and severe crash outcomes, so that the sources of severe crash reduction can be better understood.

The most important implication of this study to safety policy is that to effectively reduce the number of severe crashes a countermeasure that could both reduce the total number of crashes and severe crash outcomes would be the most desirable, since this condition guarantees a reduction in the number of severe crashes. A countermeasure that can reduce the total number of crashes may not necessarily be able to reduce crash severity. Similarly, a countermeasure that can reduce crash severity may not necessarily be able to reduce crash frequency. A countermeasure beneficial to either the total number of crashes or severe crash probability does not guarantee a reduction in the number of severe crashes, but a countermeasure beneficial to the total number of crashes and severe crash outcomes guarantees the reduction of the number of severe crashes.

Recently researchers and practitioners have called for the need to further develop crash severity CMFs in addition to current crash frequency CMFs. The approach developed in this study can be used to advance the current HSM. Specifically, as shown in the right-hand side of Eq. (3), the currently available frequency CMFs for the total number of crashes could be converted to the form of marginal effect (ΔN_T), and the total number of crashes after implementing a countermeasure, $N_{TR=1}$, could be projected using the current HSM framework. Since the severe crash probability before implementing a countermeasure, $P_{FMIC,R=0}$, could be readily obtained from crash records in the past, it is clear that once severity CMFs are available, the form of marginal effects (P_{FMIC}) would be available as well. Thus, the change in the number of severe crashes could be evaluated. Moreover, the sources leading to the changes, either from ΔN_T or P_{FMIC} , can then be discerned.

Future research should be directed at further testing the proposed analysis paradigm presented in this manuscript. In particular, the methodology presented should be compared to MVPLN and MGP models using both empirical and simulation studies. The present study considered safety countermeasure treatment effects using a binary variable (i.e., before vs. after), while future research should consider treatments that can be evaluated using a continuous variable. For example, the shoulder or median width, or horizontal curve radius, often vary significantly across a sample of road analysis segments that a total derivative technique could be

employed to develop a safety effect estimate for these variables when coded as a continuous variable.

Acknowledgement

The first author would like to thank the National Science Council, Taiwan, Republic of China, for financially supporting his contribution of this research (NSC 102-2218-E-009-022).

References

- AASHTO, Washington, DC 2010. *Highway Safety Manual*.
- Aguero-Valverde, J., Jovanis, P.P., 2009. Bayesian multivariate Poisson–log normal models for crash severity modeling and site ranking. *Transportation Research Record: Journal of the Transportation Research Board* 2136, 82–91.
- Allison, P.D., 2009. *Fixed Effects Regression Models*. SAGE Publications, Inc., Thousand Oaks, CA.
- Angrist, J.D., Pischke, J.S., 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.
- Berk, R., MacDonald, J., 2007. Overdispersion and Poisson regression. *Journal of Quantitative Criminology* 24, 269–284.
- Cameron, A.C., Trivedi, P.K., 2005. *Microeconometrics: Methods and Applications*. Cambridge University Press, New York, NY.
- Carrasco, O., McFadden, J., Chandhok, P., Patel, R., 2004. Evaluation of the Effectiveness of Shoulder Rumble Strips on Rural Multilane Divided Highways in Minnesota. In: Presented at the 83rd Annual Meeting of the Transportation Research Board, Washington, DC.
- Chiou, Y.-C., Fu, C., 2013. Modeling crash frequency and severity using multinomial-generalized Poisson model with error components. *Accident Analysis and Prevention* 50, 73–82.
- El-Basyouny, K., Sayed, T., 2009. Collision prediction models using multivariate Poisson–lognormal regression. *Accident Analysis and Prevention* 41 (4), 820–828.
- FHWA, 2012a. HSIP Website. U.S. Department of Transportation, Washington, DC <http://safety.fhwa.dot.gov/hsip/> (accessed 06.11.13).
- FHWA, 2012b. *Highway Statistics*. U.S. Department of Transportation, Washington, DC.
- FHWA, 2013. *Crash Modification Factors Clearinghouse*. U.S. Department of Transportation, Washington, DC <http://www.cmfclearinghouse.org/> (accessed 06.11.13).
- Gelman, A., Hill, J., 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York, NY.
- Gkritza, M., 2008. Mixed logit analysis of safety belt use in single- and multioccupant vehicles. *Accident Analysis and Prevention* 40 (2), 443–451.
- Greenstone, M., 2002. A reexamination of resource allocation responses to the 65-mph speed limit. *Economic Inquiry* 40 (2), 271–278.
- Geedipally, S.R., Bonneson, J.A., Pratt, M.P., Lord, D., 2013. Severity distribution function for freeway segments. In: *Transportation Research Board Annual Meeting, Paper #13-2873*.
- Griffith, M.S., 1999. Safety evaluation of rolled-in continuous shoulder rumble strips installed on freeways. *Transportation Research Record: Journal of the Transportation Research Board*, 28–34, No. 1665.
- Hausman, J., 1978. Specification tests in econometrics. *Econometrica* 46 (6), 1251–1271.
- Hausman, J., Hall, B.H., Griliches, Z., 1984. Econometric models for count data with an application to the patents – R&D relationship. *Econometrica* 52, 909–938.
- Hilbe, J.M., 2010. *Negative Binomial Regression*, 2nd ed. Cambridge University Press, New York, NY.
- Houston, D.J., 1999. Implications of the 65 mph speed limits for traffic safety. *Evaluation Review* 23 (3), 304–315.
- Kockelman, K.M., Kweon, Y.-J., 2002. Driver injury severity: an application of ordered probit models. *Accident Analysis and Prevention* 34 (3), 313–321.
- Law, T.H., Noland, R.B., Evans, A.W., 2009. Factors associated with the relationship between motorcycle deaths and economic growth. *Accident Analysis and Prevention* 41 (2), 234–240.
- Law, T.H., Noland, R.B., Evans, A.W., 2010. The direct and indirect effects of corruption on motor vehicle crash deaths. *Accident Analysis and Prevention* 42 (6), 1934–1942.
- Lenguerrand, E., Martin, J.L., Laumon, B., 2006. Modeling the hierarchical structure of road crash data – application to severity analysis. *Accident Analysis and Prevention* 38 (1), 43–53.
- Long, J.S., 1997. *Regression Models for Categorical and Limited Dependent Variables*. SAGE Publications, Thousand Oaks, CA.
- Ma, J., Kockelman, K., 2006. Bayesian multivariate Poisson regression for models of injury count, by severity. *Transportation Research Record* 1950, 24–34.
- Ma, J., Kockelman, K.M., Damien, P., 2008. A multivariate Poisson–lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis and Prevention* 40, 964–975.
- Milton, J., Shankar, V., Mannering, F., 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. *Accident Analysis and Prevention* 40 (1), 260–266.
- Neuhaus, J.M., Kalbfleisch, J.D., Hauck, W.W., 1991. A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. *International Statistical Review* 59, 25–35.
- Neuhaus, J.M., Kalbfleisch, J.D., 1998. Between- and within-cluster covariate effects in the analysis of clustered data. *Biometrics* 54 (2), 638–645.
- Patel, R.B., Council, F.M., Griffith, M.S., 2007. Estimating the Safety Benefits of Shoulder Rumble Strips on Two Lane Rural Highways in Minnesota: An Empirical Bayes Observational Before–After Study. In: Presented at the 86th Annual Meeting of the Transportation Research Board, Washington, DC.
- Park, E.S., Lord, D., 2007. Multivariate Poisson–lognormal models for jointly modeling crash frequency by severity. *Transportation Research Record: Journal of the Transportation Research Board* 2019, 1–6.
- Rabe-Hesketh, S., Skrondal, A., 2008. *Multilevel and Longitudinal Modeling Using Stata*, 2nd ed. Stata Press, College Station, TX.
- Sayed, T., de Leur, P., Pump, J., 2010. Impact of rumble strips on collision reduction on highways in British Columbia, Canada. *Transportation Research Record: Journal of the Transportation Research Board* 2148, 9–15.
- Smith, E.B., Ivan, J.N., 2005. Evaluation of safety benefits and potential crash migration due to shoulder rumble strip installation on Connecticut freeways. *Transportation Research Record: Journal of the Transportation Research Board* 1908, 104–113.
- Torbic, D.J., Hutton, J.M., Bokenkroger, C.D., Harwood, D.W., Gilmore, D.K., Dunn, J.M., Ronchetto, J.J., Donnell, E.T., Sommer, H.J., Garvey, P., Persaud, B., Lyon, C., 2009. NCHRP Report 641: Guidance for the Design and Application of Shoulder and Centerline Rumble Strips. *Transportation Research Board of the National Academies*, Washington, DC.
- Train, K., 2009. *Discrete Choice Methods with Simulation*, 2nd ed. Cambridge University Press, New York, NY.
- Wooldridge, J.M., 2010. *Econometric Analysis of Cross Section and Panel Data*, 2nd ed. The MIT Press, London, England.
- Wu, K., Himes, S.C., Pietrucha, M.T., 2012. An evaluation of the effectiveness of the Federal Highway Safety Improvement Program. *Transportation Research Record: Journal of the Transportation Research Board* 2318, 23–34.
- Wu, K., Himes, S.C., Pietrucha, M.T., 2013. Examining fatal crash reductions by first harmful events since the introduction of the Federal Highway Safety Improvement Program. *Transportation Research Record: Journal of the Transportation Research Board* 2364, 51–61.
- Yannis, G., Papadimitriou, E., Antoniou, C., 2008. Impact of enforcement on traffic accidents and fatalities: a multivariate multilevel analysis. *Safety Science* 46 (5), 738.
- Ye, X., Pendyala, R.M., Washington, S.P., Konduri, K., Oh, J., 2009. A simultaneous model for crash frequency by collision type for rural intersections. *Safety Science* 47 (3), 443–452.
- Zegeer, C.V., Reinfurt, D.W., Hummer, J., Herf, L., Hunter, W., 1988. Accident effects of sideslope and other roadside features on two-lane roads. *Transportation Research Record: Journal of the Transportation Research Board* 1195, 20–32.