16th Road Safety on Four Continents Conference
Beijing, China 15-17 May 2013

IDENTIFYING BLACK SPOTS USING PROPERTY DAMAGE ONLY EQUIVALENCY (PDOE) FACTORS

Simon Washington
Civil Engineering and Built Environment, Science and Engineering Faculty and
Centre for Accident Research and Road Safety (CARRS-Q), Faculty of Health
Queensland University of Technology
2 George St GPO Box 2434 Brisbane QLD 4001 Australia
Tel: +61 7 3138 9990
Email: simon.washington@qut.edu.au

Md. Mazharul Haque
Civil Engineering and Built Environment, Science and Engineering Faculty and
Centre for Accident Research and Road Safety (CARRS-Q), Faculty of Health
Queensland University of Technology
130 Victoria Park Rd Kelvin Grove QLD 4059 Australia
Tel: +61 7 3138 4511
Email: m1.haque@qut.edu.au

Jutaek Oh
The Korea Transport Institute
2311 Daehwa-dong, IIsanseo-gu, Goyang-si, Gyeonggi-do 411-701, Korea
Tel: 82-31-910-3174
Email: jutaek@koti.re.kr

Dongmin Lee
The Korea Transport Institute
2311 Daehwa-dong, IIsanseo-gu, Goyang-si, Gyeonggi-do 411-701, Korea
Tel: 82-31-910-3011
Email: dmlee@koti.re.kr

ABSTRACT
Methodologies for identifying 'sites with promise' have received considerable attention in the transportation safety literature. Robust hot spot identification (HSID) methodologies are vital as errors result in either identifying a safe site as high risk (false positive) or a high risk site as safe (false negative) misuse of public funds and lead to poor investment decisions and inefficient risk management. The historical pursuit of HSID methods development has lead to important insights; however, there remain at least several critical impediments to further progress: 1) a significant proportion of property damage only (PDO) and minor injury crashes are underreported (approximately 40%), affecting the reliability of count based models; 2) most methods ignore crash severity and costs; and 3) expected safety performance functions are heavily skewed by a preponderance of zeroes. This paper argues that it is possible and indeed desirable to incorporate crash costs into HSID. Moreover, a straightforward method is

proposed whereby crashes are intelligently weighted using property damage only equivalent (PDOE) crashes. The use of PDOEs enables identification of a set of high-risk sites that reflect the true safety costs to society and simultaneously reduces the influence of under-reported PDO crashes, thereby addressing impediments 1 and 2. Non-parametric Quantile regression is used to overcome the preponderance of zeroes problem (impediment 3). The proposed procedure is illustrated using rural road segment data from Korea.

1 INTRODUCTION

1.1 Hot Spot Identification Techniques

Identifying crash hot spots, also known as black spots, sites with promise, accident-prone locations, and priority investigation locations, that seeks to screen potentially hazardous locations in a roadway network for further improvement is an imperative—though challenging—step to improve safety performance of a road network. Challenges mainly arise due to heterogeneities in crash frequencies caused by heterogeneous driving population, traffic conditions, roadside features and design considerations. Common hot spot identification (HSID) techniques include ranking according to crash frequencies, ranking according to crash rates, the confidence interval technique, the crash reduction potential method and the empirical Bayes (EB) method. Previous research (e.g., Persaud, 1986, Persaud, 1988, Persaud and Hauer, 1984, Hauer, 1997) has reported that HSID methods relying on a simple ranking of crash counts or crash rates, due to random fluctuation of crashes from year to year, can produce large number of false positives (safe sites falsely identified as unsafe) and false negatives (truly hazardous sites escape identification). These errors result in inefficient use of federal and/or state aid and local government resources applied for safety improvements.

Hauer and Persaud (1984) drew an analogy between the first stage of identification of black-spots and a sieve, and discussed how to measure the performances of various methods of identifying hot spot sites. Based on this study, Higle and Hecht (1989) conducted a simulation experiment to evaluate and compare techniques for the identification of hazardous locations in terms of crash rates. Subsequent work by Hauer (1997) and others (Bauer and Harwood, 2000, Hadayeghi et al., 2003, Miaou and Lord, 2003) has shown that safety performance functions may be curvilinear with respect to vehicles mile travel (VMT), and therefore should not in general be used to rate the risk of various sites.

The Empirical Bayes' (EB) method, formally introduced by Hauer (1997), has been adopted as the state of the practice HSID. The application of EB for HSID has received a great deal of attention as it accounts for both crash history and expected crashes on similar sites—two essential clues to safety at a site (Persaud, 1999). It follows that the safety of a site is affected by not only some common measurable factors shared by a corresponding reference population (generally captured in the safety performance function) but also some unique characteristics associated with the site (reflected in its crash history). In EB method, the safety of a site is estimated by a weighted average of observed crash count of the subject site and expected crashes of similar sites, where the weight is determined by the variance in estimating expected crashes of the reference sites. Hauer et al. (1988) applied Empirical Bayes' (EB) method to estimate the safety at signalized intersections, Persaud (1991) evaluated crash potential of Ontario road sections and Higle and Witkowski (1988) presented a Bayesian technique making use of crash rates. In a carefully controlled Monte Carlo simulation study comparing crash rate ranking, frequency ranking, accident reduction potential, and EB methods, Cheng and Washington (2005) showed that under controlled experimental conditions the EB method is in general superior to all other methods available for identifying high risk sites—revealing the lowest percentage of false positive and false negative errors. In subsequent work Cheng and Washington (2008) developed new criteria under which HSID methods can be evaluated and again the EB method yielded superior performance.

In a few studies on HSID methods researchers attempted to tackle the complex issue of crash severity. For example, the Missouri Department of Transportation identified seven methods for identifying high crash locations, two of which acknowledged the importance of crash severity (MDT, 1999). They detailed a crash severity method that weighed injury and fatal crashes, dictated by 'local policy' that appears to be somewhat arbitrary, by a factor of (for example) 6 compared to property damage only (PDO) crashes to obtain an EPDO estimate. The severity-rate method they identified takes the EPDO estimate and divides by exposure across locations to obtain an EPDO based rate.

Tarko and Kanodia (2003) recommended the index of crash frequency and index of crash cost as the 'best' methods for conducting HSID after conducting a thorough review. The index of crash frequency method in simple terms estimates safety performance functions by location (rural multi-lane roads, rural interstates, etc.) and compares the observed to expected total crash frequencies (divided by the standard deviation of the difference estimate) to rank sites for potential improvement. This method does not account for severity nor does it account for possible regression to the mean effects. Their second recommended method is similar to the first except that it uses crash costs to incorporate severity. Count models are estimated separately for PDOs and injuries and fatalities (I/Fs) (Tarko et al., 2000). Then, the average costs for PDOs and I/Fs (and other ancillary statistics) are used to calculate a severity-weighted index. This method accounts for severity, but requires as many regression models as there are severity classes which becomes cumbersome and requires estimation on increasingly smaller samples sizes. Ma et al. (2008) proposed even a more complex multivariate Poisson-lognormal model to consider severity and frequency in a safety performance function simultaneously. While this approach is extremely capable of accounting both frequency and severity for HSID, it is cumbersome for practitioners and safety managers to apply due to its significant complexity and time commitment to estimate.

1.1 Objective

This paper proposes a property damage only (PDO) equivalency technique for identifying hot spots on a transportation network. The proposed technique overcomes some of the difficulties of past methods, namely: 1) a significant proportion of PDO and minor injury crashes are underreported (approximately 40%), affecting the reliability of count based models; 2) most methods ignore crash severity and costs; and 3) expected safety performance functions are heavily skewed by a preponderance of zeroes. The under-reporting problem is addressed in much greater detail in a companion paper by Oh et al. (2010), leaving the focus here on the impact of crash costs and preponderance of zeroes. Performances of the proposed PDO equivalency method and state of the practice Empirical Bayes method are compared, contrasted, discussed.

2 PROPOSED PDO EQUIVALENT METHOD FOR HSID

2.1 PDO Equivalent

Why do motor-vehicle crashes matter to society? According to Hauer (1997), “Road safety is manifest in the occurrence of accidents and their harm”. It is difficult to disagree with this simple assessment. Hence ‘the occurrence’ is measured by the count of crashes; but how is ‘their harm’ measured? Fortunately, this question has been addressed in considerable detail in a landmark study by Blincoe et al. (2002) which shows comprehensive national estimates of the harm of motor vehicle crashes on a per crash cost basis in US. The estimate shows that a fatal crash costs on average $3.4 million, while the average PDO crash costs $2,532. Estimates of the cost of road traffic crashes from Australia (BITRE, 2009) and Singapore (Chin et al., 2006) have also indicated a similar effect to the society.

How to accommodate crash cost in the HSID methodology is well illustrated by an example borrowed from the pavements profession. Experts in this field explicitly recognize that a load that vehicles impart on pavements is a non-linear function of the weight or load over the vehicle axle. As a result, the profession has developed axle load equivalency factors to aid in the assessment of pavement damage (Mannering et al., 2005). The pavement stresses and strains in fact are so non-linear that for a flexible pavement (with terminal serviceability index , TSI = 2.5 and Structural Number, SN = 3) a 20,000 lb single axle load does 1620 times the damage to pavement as does a 2000 lb single axle load. Moreover, 40,000 and 80,000 lb axle loads do 20,600 and 392,000 times the damage to flexible pavements as does a 2000 single axle load respectively. The main point is this: simply counting traffic over a section of road is insufficient for predicting or explaining pavement damage (and their associated costs), since all axle loads are not created equal. This phenomenon in fact is common, for example, it applies also to earthquakes, as the frequency alone is insufficient and the magnitude is necessary to assess the true impacts.

Clearly this same phenomenon exists in road safety. All crashes are not created equal, since fatal and severe injury crashes are far more costly to society than are property damage only (PDO) crashes. The average fatal crash in 2000 cost nearly $3,366,388, whereas the average PDO crash cost $2,532—a ratio of about 1330 to 1 (Blincoe et al., 2002). Societal crash costs include medical, emergency services, market productivity, household productivity, insurance administration, workplace cost, legal costs, travel delays, and property damage—the cumulative costs that society bears when a person is injured or killed in a motor vehicle crash.

In terms of societal safety cost, or total safety impact, the average fatal crash in 2002 was 1330 times more costly than the average PDO crash (on a per-person injured basis). Alternatively, a PDO-equivalency factor can be introduced so that a fatal crash is worth 1330 PDO crashes. PDO equivalency factors calculated for major and minor injury crashes (adapted from Blincoe et al., 2002) are 949 and 11 respectively. As a result a modified and straightforward HSID method is proposed that incorporates a PDO equivalency factor in estimating safety performance function. The PDO equivalency factors are given as:

$$y_i = PDOE_i = PDO_i + MinINJ_i(11) + MajINJ_i(949) + Fatal_i(1330) \qquad (1)$$

where the number of crashes at location i, y_i, is equal to the property damage equivalent (PDOE) crashes at this site calculated as the weighted sum of PDO, minor injury, major injury, and fatal crashes. The modified estimate of crashes across transportation system locations is now used in HSID.

2.2 Quantile regression models of PDOE

While crash frequencies lend themselves to generalized negative binomial models, the introduction of property damage only equivalents (PDOEs) complicates the analysis. First and foremost, the data are no longer approximately Poisson or negative binomial distributed. In fact the PDOE data are not 'nicely behaved' at all—and are not distributed like any well known statistical distribution. An examination shows that the average PDOE per unit of length at sites with varying posted speeds is near to zero for all speeds, suggesting that a regression through the means will yield a fruitless pursuit. Moreover, the expected safety performance function—which establishes relationships between predictors and expected or mean safety, is heavily skewed by the preponderance of zeroes in the data. This "zeroes" problem is not unique to modeling PDOEs and causes analysis issues with crash frequency models as well. For example, some experimentation with zero inflated models has been aimed at tackling this thorny issue (Nam and Lee, 2006, Shankar et al., 1997), but criticized for solving analysis problems but leaving interpretation and theoretical inconsistencies (Lord et al., 2007, Lord et al., 2005).

To solve simultaneously the poorly behaved distributional performance and the preponderance of zeroes, a non-parametric quantile regression technique (Koenker and Hallock, 2001) to model the PDOE data was employed. In contrast to linear regression models, which focus on the conditional mean (the mean conditioned on values of the model covariates), a quantile regression conditions on a particular quantile, such as the 90th percentile, 80th percentile, etc. The motivations for applying the quantile regression are interest in quantiles other than the mean and non-linearities of the regression function at values other than the mean. Quantile regression functions are optimization problems where asymmetrically weighted sum of absolute deviations of residuals are minimized for a particular quantile, in contrast to least squares estimation where the sum of squared residuals is minimized to find the mean function (Koenker and Hallock, 2001). In models that follow, bootstrap sampling was used to obtain estimates of the correct quantiles and parameter estimates.

2.3 HSID criteria

Application of the property damage only equivalent (PDOE) based model requires an estimation of excess PDOEs of a site than its expected PDOEs. Hot spots are identified by ranking sites by the observed PDOE minus their predicted PDOE, with the most high risk site identified as having the highest number of observed PDOEs compared to what is predicted for this site (in a multivariate model).

$$\text{Excess PDOEs} = \text{PDOE}_\text{i} - \widehat{\text{PDOE}}_\text{i} \tag{2}$$

Equation 2 will produce an estimate of the 'excess' number of PDO equivalent crashes occurring at site i. Since PDOEs instead of crash frequencies are used, fatal and major injury crashes will have an appropriate and disproportionately large influence on the statistic. Moreover, PDOEs are estimated at thresholds such as 90^{th} and 95^{th} percentiles, and thus represent the expected behavior of the worst performing sites against which the performance of individual sites is compared. The proposed HSID technique is compared against the state of the practice Empirical Bayes (EB) approach. The EB method is not described here, but the interested readers may consult a vast amount of available literature (e.g., Hauer et al., 1988, Persaud, 1999, Hauer, 1997).

3 DATASET FOR ANALYSIS

For the empirical investigation of the proposed hot spot identification (HSID) technique using property damage only equivalents (PDOEs), rural road segment data from South Korea are examined. Data for this study were collected under the explicit recognition that crashes at rural road segments are affected by both local and regional characteristics. Road segments in rural locations adjacent to metropolitan areas, for example, may consist of different driving populations and hence safety compared to intersections in rural locations. Roadside conditions are different depending on the location of rural segments within South Korea. Considering this expected heterogeneity across sites, road segment crash and geometric data were obtained from two sites in South Korea. The first location is outside of Seoul, the capital of South Korea and the largest city. The population of Seoul was about 10 million in 2008. The second location represents segment data collected from a more rural region, with no major cities nearby.

The data are based on a total of 2,916 highway road segments in rural areas and were obtained from two sources. First, detailed crash records from 2005 to 2007 were obtained from the national police agency. Roadway inventory data were obtained from field surveys that were conducted from 09/01/2008 through 11/30/2008. Based on underlying theories of crash causation and with the intent to establish defensible statistical models to enable the examination of possible relationships among crash frequencies, geometric, and traffic characteristics of road segments, a total of 45 possible variables were considered in the analysis. Explanatory variables include a wide range of roadway and traffic characteristics such as average traffic volume (ADT), heavy truck volume (HVADT), posted speed limit, roadway segment length, number of lanes, horizontal and vertical curvature, terrain condition, shoulder width, median width, shoulder type, lighting condition and land use characteristics.

4 RESULTS AND DISCUSSION

Estimated safety performance functions for the quantile regression models of PDOEs that reflect 90%, 95%, and 97% percentiles are respectively shown in equation 3, 4, and 5. The safety performance function (SPF) of the traditional Negative Binomial model used for the EB method is also reported in equation 6.

$$Y_{(PDOE90)_i} = (length)\exp\begin{pmatrix} -0.08(level\_terrain) - 0.27(rolling\_terrain) \\ -2.66(concrete\_median) + 4.49(speed\_40kph) \\ +1.71(ADT * 10k) - 1.64 \end{pmatrix} \tag{3}$$

$$\gamma_{(PDOE95)_i} = (length)exp\begin{pmatrix} -4.25(paved\_shoulder) - 3.75(nonpaved\_shoulder) \\ -6.26(concrete\_median) - 13.48(speed\_50kph) \\ -12.63(speed\_60kph) - 10.43(speed\_80kph) + 4.51 \end{pmatrix} \quad (4)$$

$$\gamma_{(PDOE97)_i} = (length)exp\begin{pmatrix} -3.94(rolling\_terrain) - 8.34(concrete\_median) \\ -9.84(speed\_50kph) - 5.26(speed\_60kph) \\ +1.80(ADT * 10k) + 9.02 \end{pmatrix} \quad (5)$$

$$\gamma_{(Count)_i} = (length)exp\begin{pmatrix} 1.50(horizontal\_curvature) + .99(curve\_radius) \\ +1.32\ (ADT * 10k) + 3.21(HVADT * 1k) \\ +5.82(resident\_landuse) + 5.27(industrial\_landuse) \\ +0.69(level\_terrain) + 2.05(speed\_40kph) - 7.84 \end{pmatrix} \quad (6)$$

It is evident that both quantile regression models of PDOEs and negative binomial model of raw crash counts identify a substantially different set of statistically significant and logically defensible predictors. To remind the reader, the model on PDOEs is intended to reflect a more complete picture of segment safety by including both frequency and severity effects. The quantile regression models, which reflect the 90%, 95%, and 97% percentiles, reveal that significant predictors change quite considerable as increasingly smaller subsets of outliers are examined. Using the SPFs from equation 3 to 5, the hot spots are identified following the equation 2 and compared against the EB method that screens hot spots using the SPF in equation 6.

Table 1 shows the top 20 sites (out of 2916 rural road segments) identified using both methods. Note that 8 sites identified using crash frequencies were not identified using PDO equivalents, while 8 sites identified using PDO equivalents were not identified using crash frequencies. The only sites that appear in the top 10 of both methods are site ID 610 and 328, all others are unique to method. The table also shows how different percentiles of the quantile regression affect the ranking of top sites.

Table 1: High Risk Site ID Comparison: PDOEs vs. EB of Crash frequencies

High Risk Ranking	97% PDOE	95% PDOE	90% PDOE	EB Crash frequencies
1	**2810**	**2810**	**2810**	1073
2	**2043**	**2043**	**2043**	1068
3	**1384**	**1384**	**1384**	**328**
4	**1203**	**1203**	**38**	460
5	**1202**	**1202**	1203	**610**
6	**1318**	**1318**	1202	578
7	38	**1911**	1318	1029
8	**1911**	**610**	**610**	633
9	**2177**	**1478**	**2177**	323
10	**610**	**38**	**1911**	429
11	**2809**	**2177**	2809	1043
12	**1478**	**1488**	**1478**	618
13	**1381**	1995	**43**	64
14	**1488**	1265	352	7
15	**1995**	**20**	**20**	607
16	**328**	**2809**	**1381**	344
17	**1265**	**328**	**328**	1017
18	**43**	**2814**	1488	1018
19	**2814**	**1915**	**1995**	997
20	**1915**	1248	**1265**	458

Bold = Number is repeated in another list: Non-Bold = number unique to list

Table 2 lists the top ten most hazardous sites identified by each method, contrasting crash statistics of hot spots identified using the 97% quantile regression PDOE approach compared to state of the art EB approach. The contrast of sites identified using the two methods is stark. First, the top 3 high risk sites identified using crash frequencies are heavily skewed towards sites with many PDO and minor injury crashes—there were 6 major injury or fatal crashes that occurred at the top 3 sites. Note that this approach is commonly applied in practice. In contrast, the top 3 sites identified using PDO Equivalents are heavily skewed towards sites with major injury crashes and fatalities. There were 11 major injury and fatal crashes that occurred at the top 3 sites identified using the modified method. The reason behind this stark difference is due to the heavy emphasis placed on crash severity and its role in determining the safety of a site. When all crashes are treated as equal then frequencies will dominate—when crashes are properly weighted to reflect the true societal impact a different picture emerges.

In contrast to Table 2, which shows detailed statistics on the top 10 sites for the two methods (PDOE at 97th percentile), Table 3 shows the aggregate differences between the two methods (for 90th, 95th, and 97th percentiles). It shows crash statistics for the top 2.5%, 5%, and 10% of sites identified as high risk using both crash frequencies and PDO equivalents.

It can be seen, for example, that when the top 2.5% of sites are identified using the EB approach using crash frequencies, there are a total of 225 crashes that occurred at those sites, including 8 fatals, 96 major injuries, and 72 minor injury crashes. In contrast the PDO equivalents method (97% percentile) identifies a total of 209 crashes at the top 73 sites, with 26 of them fatal, 110 as major injuries, 33 as minor injuries, and 40 PDOs. Hence the PDOE method has identified sites with 18 more fatal crashes and 14 major injury crashes than did the EB method using crash frequencies.

As the number of sites is increased the two methods start to converge. By the time 291 high risk sites are identified—10% of the sample—fatal crahses are within 1 crash by each method, major injuries are within 13 crashes, while differences still exist among minor injuries and PDOs between the two models. In the middle—where 5% of high risk sites are identified—there are still significant discrepancies especially with regard to fatal and major injury crashes.

Table 2: Crash statistics of Identified High Risk Sites: 97.5% PDOEs vs. EB

Site ID	Crash Types			
	Fatal	**Major Injury**	**Minor Injury**	**Property Damage Only**
Top 10 High Risk Sites Identified using 97% PDOEs				
2810	0	3	0	1
2043	0	3	1	1
1384	1	4	2	0
1203	0	2	1	1
1202	0	2	0	1
1318	2	1	0	1
38	1	0	0	0
1911	1	2	1	1
2177	0	4	0	2
610	0	3	1	0
Top 10 High Risk Sites Identified using EB of Crash frequencies				
1073	0	2	8	3
1068	0	2	5	5
328	0	2	1	0
460	0	2	2	0
610	0	3	1	0
578	0	3	1	0
1029	0	2	3	4
633	0	1	2	0
323	0	1	1	0
429	0	2	1	0

In practice the number of high risk sites identified will be limited because of restricted financial resources, limited human resources for auditing sites, and competing safety priorities. Thus, when the top proportion of sites identified is relatively small (5% or less) the two methods produce quite different results. The PDO equivalent method identifies sites where crashes have been quite severe, whereas the conventional method identifies sites where minor injury and PDO crashes are abundant.

Table 3: High Risk Site Comparison Statistics: PDOEs vs. Crash frequencies

High Risk Percentage (sample size)	HSID Method	Fatal crashes	Major injuries	Minor Injuries	PDOs	Total Crashes
2.5% n = 73	*97% PDO Equivalents*	*26*	*110*	*33*	*40*	*209*
	EB Crash Counts	8	96	72	49	225
5% n = 146	*95% PDO Equivalents*	*28*	*201*	*72*	*69*	*370*
	EB Crash Counts	20	167	121	86	394
10% n = 291	*90% PDO Equivalents*	*38*	*221*	*85*	*86*	*430*
	EB Crash Counts	37	234	176	127	574
		Fatal crashes/ km	Major injuries/ km	Minor Injuries/ km	PDOs/ km	Total Crashes/ km
2.5% n = 73	*97% PDO Equivalents*	*2.7*	*11.5*	*3.4*	*4.2*	*21.8*
	EB Crash Counts	0.7	8.6	6.4	4.4	20.1
5% n = 146	*95% PDO Equivalents*	*2.7*	*11.4*	*3.2*	*3.5*	*20.8*
	EB Crash Counts	0.8	6.9	5.0	3.6	16.4
10% n = 291	*90% PDO Equivalents*	*0.7*	*3.8*	*1.5*	*1.5*	*7.4*
	EB Crash Counts	0.7	4.7	3.5	2.5	11.3

5 CONCLUSIONS

Other engineering disciplines have recognized the importance of weighing the magnitude of an event and not just the frequency. Pavement damage and the occurrence of earthquakes are both examples of where damage is a function of both frequency and magnitude of events. Safety researchers have recognized the importance of crash severity and have looked for ways to include it in efficient, straightforward, and defensible ways. Unfortunately, most past efforts have resulted in fairly complex and hard to implement modeling frameworks or arbitrary weighting of injury and fatal crashes. A method is proposed and described in this paper using property damage only equivalents (PDOEs), weighting the crash count sum according to the direct cost impact that crashes have on society. It is argued that this weighting is justified based on the public ownership of the transportation system and the total societal costs of motor vehicle crashes.

Econometric issues arise in the PDO equivalency approach, namely that the standard distributional assumptions of crash being negative binomial distributed are questionable. Additional variation is added through the PDO calculation, while a large proportion of zeroes are retained. Essentially these features add to the unobserved heterogeneity already complicating crash count modeling efforts. A non-parametric technique—quantile regression—is applied to overcome both the distributional limitations as well as the disproportionate influence of the preponderance of zeroes in the data.

After describing the analytical process in detail, the use of PDO equivalents is demonstrated and compared to current state of the practice using Korean road segment data.

Individual sites are examined and differences shown as to which sites are identified as "high risk" based on the PDO equivalency approach compared to the state of the practice EB method. As expected, the method using PDO equivalents places more emphasis on costly and therefore more harmful crashes than does the conventional EB method. Comparison of aggregate statistics of 2.5%, 5%, and 10% hot spots identified suggests that the methods differ considerably in their performance, with the PDO equivalency method identifying far more harmful crash locations, on average.

The method proposed here may serve as a substitute for or complement to existing methods. It is intended to be straightforward, non-arbitrary in the weighting, and overcome the econometric problems. Certainly it is recommended that this method be included in a network screening process that has to date ignored crash severity. The appeal of the proposed approach is its practical, intuitive, defensible, and straightforward application.

ACKNOWLEDGEMENTS

This research was supported by a grant (06 Transportation Core C01) from the Transportation System Innovation Program (TSIP) of the Ministry of Land, Transportation and Maritime Affairs (MLTM) of Korea.

REFERENCES

BAUER, K. M. & HARWOOD, D. 2000. Statistical models of at-grade intersection accidents—Addendum. Washington, DC: Federal Highway Administration.

BITRE 2009. Cost of road crashes in Australia 2006. Canberra: Bureau of Infrastructure, Transport and Regional Economics.

BLINCOE, L., SEAY, A., ZALOSHNJA, E., MILLER, T., ROMANO, E., LUCHTER, S. & SPICER, R. 2002. The economic impact of motor vehicle crashes, 2000. Washington, D.C.: National Highway Traffic Safety Administration.

CHENG, W. & WASHINGTON, S. 2008. New criteria for evaluating methods of identifying hot spots. *Transportation Research Record: Journal of the Transportation Research Board,* 2083**,** 76-85.

CHENG, W. & WASHINGTON, S. P. 2005. Experimental evaluation of hotspot identification methods. *Accident Analysis and Prevention,* 37**,** 870-881.

CHIN, H. C., HAQUE, M. M. & JEAN, Y. H. 2006. An estimate of road accident costs in Singapore. *In Proceedings of International Conference on Road Safety in Developing Countries.* Dhaka, Bangladesh.

HADAYEGHI, A., SHALABY, A. S. & PERSAUD, B. N. 2003. Macrolevel accident prediction models for evaluating safety of urban transportation systems. *Transportation Research Record: Journal of the Transportation Research Board,* 1840**,** 87-95.

HAUER, E. 1997. *Observational before-after studies in road safety,* Tarrytown, NY, Pergamon Elsevier Science, Inc.

HAUER, E., NG, J. C. N. & LOVELL, J. 1988. Estimation of safety at signalized intersections. *Transportation Research Record: Journal of the Transportation Research Board,* 1185**,** 48-61.

HAUER, E. & PERSAUD, B. N. 1984. Problem of identifying hazardous locations using accident data. *Transportation Research Record: Journal of the Transportation Research Board,* 975**,** 36-43.

HIGLE, J. L. & HECHT, M. B. 1989. A comparison of techniques for the identification of hazardous locations. *Transportation Research Record: Journal of the Transportation Research Board,* 1238**,** 10-19.

HIGLE, J. L. & WITKOWSKI, J. M. 1988. Bayesian identification of hazardous locations. *Transportation Research Record: Journal of the Transportation Research Board,* 1185**,** 24-36.

KOENKER, R. & HALLOCK, K. 2001. Quantile regression. *Journal of Economic Perspectives,* 15**,** 143-156.

LORD, D., WASHINGTON, S. & IVAN, J. N. 2007. Further notes on the application of zero-inflated models in highway safety. *Accident Analysis & Prevention,* 39**,** 53-57.

LORD, D., WASHINGTON, S. P. & IVAN, J. N. 2005. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accident Analysis & Prevention,* 37**,** 35-46.

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 & Prevention,* 40**,** 964-975.

MANNERING, F., KILARESKI, W. & WASHBURN, S. 2005. *Principles of highway engineering and traffic analysis*, John Wiley & Sons.

MDT 1999. Identification, analysis, and correction of high-crash locations (the HAL Manual), 3rd edition. Technology Transfer Assistance Program, Missouri Department of Transportation

MIAOU, S.-P. & LORD, D. 2003. Modeling traffic crash-flow relationships for intersections: Dispersion parameter, function form, and Bayes' Versus Empirical Bayes' methods. *Transportation Research Record: Journal of the Transportation Research Board,* 1840**,** 31-40.

NAM, D. & LEE, J. 2006. Accident frequency model using zero probability process. *Transportation Research Record: Journal of the Transportation Research Board,* 1973**,** 142-148.

OH, J., WASHINGTON, S. & LEE, D. 2010. Property damage crash equivalency factors to solve crash frequency-severity dilemma: Case study on South Korean rural roads.

PERSAUD, B. N. 1986. Safety migration, the influence of traffic volumes, and other issues in evaluating safety effectiveness. *Transportation Research Record: Journal of the Transportation Research Board,* 1086**,** 33-41.

PERSAUD, B. N. 1988. Do traffic signals affect safety? Some methodological issues. *Transportation Research Record: Journal of the Transportation Research Board,* 1185**,** 37-46.

PERSAUD, B. N. 1991. Estimating accident potential of Ontario road sections. *Transportation Research Record: Journal of the Transportation Research Board,* 1327**,** 47-53.

PERSAUD, B. N. 1999. Empirical Bayes' procedure for ranking sites for safety investigation by potential for safety improvement. *Transportation Research Record: Journal of the Transportation Research Board,* 1665**,** 7-12.

PERSAUD, B. N. & HAUER, E. 1984. Comparison of two methods for de-biasing before-and-after accident studies. *Transportation Research Record: Journal of the Transportation Research Board,* 975**,** 43-49.

SHANKAR, V., MILTON, J. & MANNERING, F. 1997. Modeling accident frequencies as zero-altered probability processes: an empirical inquiry. *Accident Analysis and Prevention,* 29**,** 829-837.

TARKO, A. P. & KANODIA, M. 2003. Hazard elimination program – manual on improving safety of Indiana road intersections and sections. US Department of Transportation.

TARKO, A. P., SINHA, K. C., ERANKY, S., BROWN, H., ROBERTS, E., SCINTEIE, R. & ISLAM, S. 2000. Crash reduction factors for improvement activities in Indiana. West Lafayette: Joint Transportation Research Program: Purdue University.