Modeling the Location of Crashes within Work Zones: Methodology and a Case-Study Application

Xiaoyu Zhu
University of Florida
365 Weil Hall
PO Box 116580
Gainesville, FL 32611

Sivaramakrishnan Srinivasan (*)
University of Florida
365 Weil Hall
PO Box 116580
Gainesville, FL 32611
Phone: 352-392-9537 ext. 1456
Fax: 352-392-3394
E-mail: siva@ce.ufl.edu

Grady Carrick
Florida Highway Patrol
7322 Normandy Blvd.
Jacksonville, FL 32205

Kevin Heaslip
Utah State University
233 Engineering
4110 Old Main Hill
Logan, UT 84322

(*) Corresponding Author

ABSTRACT

This study modeled the location of crashes within work zones as a function of the lengths of the different work-zone segments, traffic volume, weather, and other exogenous factors. A multinomial-logit model was developed that can be used to assess the relative safeness of the different work-zone segments in terms of the crash probabilities per lane-mile. Data from crash reports from a work zone in Florida were augmented with spatial attributes using geographic information systems and used in model estimation. For the work-zone studied, the results indicate that the "Advance Warning" area is particularly unsafe during times of peak traffic flow and bad weather. In contrast, the "Exit" area has a larger probability of crash (per lane-mile) during uncongested times. The probability of a crash per lane-mile in the "Work Area" was generally smaller compared to the other work-zone segments. However, "Work Area" crashes were also found to more severe in terms of the injuries sustained by the vehicle occupants. Policy implications of these results are discussed. While the empirical results are based on data from a specific work zone in Florida, the methodology is generic and can be applied to any region if appropriate spatial data are available.

Keywords:

Work zone, crash location, multinomial-logit model

1. **BACKGROUND AND OBJECTIVES**

The increasing demand for roadway capacity and the need for maintaining the aging transportation infrastructure have necessitated thousands of construction projects throughout the United States. A "fact-sheet" released by the Federal Highway Administration (FHWA) estimates that "more than 20% of the National Highway System (NHS) is under construction during the peak construction season" leading to over 3000 work zones (FHWA, 2008). While such projects are necessary to alleviate traffic congestion and to improve transportation safety, the roadway segments undergoing construction/maintenance pose their own safety issues. For instance, in the state of Florida alone, there were 361 fatalities and 10,415 injuries in highway work-zones crashes, between the years 2003 and 2005. At the national level, 40,000 persons are injured in work-zone crashes each year (FDOT, 2007).

In recognition of the need to ensure the safety of work-zones, a large number of studies have focused on analyzing work-zone crashes. Many of these studies have compared the crash rates and injury severities in work-zone segments with corresponding values from the regular highway segments. However, research designed to understand the location of crashes *within* work zones has been minimal. Since work zones can be several miles in length, knowledge of the locations that are most susceptible to crashes is important for the detection, response, and clearance activities associated with those crashes. In addition to mitigating congestion, identification of the relative crash-susceptibilities of the different work-zone segments (such as the "advance warning" area and the "taper" area) may help prioritize countermeasures aimed at reducing the incidence of crashes.

The police traffic crash report is the primary means by which researchers undertake safety analysis. However, the crash reports of most of the states do not have data elements to capture adequate details about the work zone in which a crash may have occurred (Washburn and Carrick, 2006). One such missing data element is the specific location of the crash within the work zone. Arguably, this lack of data is a critical reason for the limited research on the factors impacting the location of crashes within work zones.

The objective of this study was to develop a multinomial-logit model to determine the location of crashes within work zones. This is a methodological extension of past research which has relied largely on comparing frequency distributions of crash counts. The proposed method enables the assessment of the marginal impacts of several explanatory factors (traffic volume, weather, time-of-day, presence of a previous crash, etc.) simultaneously and also allows the calculation of crash probabilities per lane-mile of the different work-zone segments. The model was estimated using crash data from a work zone in Florida. The lack of detailed spatial data in the crash reports necessitated significant manual processing using Geographic Information Systems (GIS) for the determination of the crash locations. Hence the empirical scope of this work was limited to the one study region. However, the methodology is generic and can be applied to any region subject to availability of spatial data.

The rest of this paper is organized as follows. Section 2 presents a synthesis of relevant literature and identifies the primary contributions of this study. The econometric-modeling methodology is detailed in Section 3. Section 4 describes the data used in this analysis. Section 5 presents and discusses the empirical results. The paper ends (Section 6) with an outline of the salient features of this research, a summary of the major conclusions, and directions for further research.

2. LITERATURE REVIEW

Few studies have examined the location of crashes within work zones. Table 1 presents (in reverse chronological order) a summary of these studies, to be described briefly in the paragraphs that follow. Although all the studies presented in the table did not use the same classification scheme to locate crashes within work zones, the following four major segments may be identified: (1) the advance warning area, (2) the taper or the transition area, (3) the work or activity area (including any longitudinal buffer), and (4) the exit or termination area.

The most recent of the studies presented in Table 1 is also the most extensive one. Based on a total sample of 1484 crashes, the largest fraction of the work-zone crashes were found to occur in the work area and the smallest fraction in the exit area, irrespective of the type of highway facility or the time-of-the-day (Garber and Zhao, 2002). However, the percentages of crashes that occur in the different segments of the work zone varied based on facility type and the time-of-the day. For example, a greater percentage of crashes occurred in the advance warning area in the day time compared to the night time. Similarly, a greater percentage of crashes occurred in advance warning areas on rural interstates than on urban interstates.

Based on an analysis of crashes in Kentucky, about 80% of all work zone crashes occurred in the work area (Pigman and Agent, 1990), a result that is consistent with the previous findings (Garber and Zhao, 2002). Additional studies were conducted using data from early 1980s or before and had much smaller sample sizes. Nonetheless, the results still indicate that the crashes in work zones were not equally distributed across the different segments and that most of the crashes happened in the work area (Nemeth and Rathi, 1983; Hargroves, 1981; Nemeth and Migletz, 1978).

Overall, the most striking observation that can be made from the Table 1 is that a substantial fraction of all work-zone crashes happen in the work area. This is logical, given that the work area may also be the longest segment of the typical work zone. None of the studies provide estimates of the relative lengths of the different segments. Hence, it is unclear from the current literature whether the work areas are indeed the most unsafe part of the work zone. It is certainly possible that large numbers of crashes are observed in work-areas because of the greater lengths of this segment. A comparison of crashes on a "per lane-mile" basis would enable a more realistic assessment of the relative safety of the different segments. This research presents such an approach. In addition, this research presents multivariate-analysis to examine the simultaneous impact of several exogenous factors (such as weather, day-of-the-week, and traffic volumes) on the crash location. This is a methodological extension of past research which has relied largely on frequency distributions and examined the impacts of the explanatory factors one-at-a-time.

3. METHODOLOGY

A Multinomial-Logit (MNL) model that explicitly accounts for the differences in the "size" (i.e., length or lane-miles) of the alternatives (i.e., the different work-zone segments) is proposed to relate the location of crashes within work zones to the exogenous factors. This is well established in the literature on using discrete-choice methods for travel behavior modeling. For example, Ben-Akiva and Abou-Zeid (2007) use this approach to model the choice of the time-of-day of trip making when the 24-hour day is discretized into periods (such as the "peak" and Off-peak" periods) of unequal durations.

This section presents a broad overview of this methodology. For simplicity, the model structure is described assuming three exogenous factors (size of the different segments, traffic volume, and weather). The procedure can be extended in a straightforward manner to accommodate more explanatory variables.

Consistent with the Model Minimum Uniform Crash Criteria (MMUCC) guidelines (NHTSA, 2003), the work zone is divided into five segments. Segment 1 is the area before the advance-warning sign (represented in short as "B"), Segment 2 is the advance-warning area ("A"), Segment 3 is the transition area or the taper ("T"), Segment 4 is the work area ("W"), and Segment 5 is the exit area ("E").

Let $i = \{B, A, T, W, E\}$ be the index for these work-zone segments. For a given crash, crash-propensity functions (U_i) are defined for each location as

$$U_i = \alpha_i + \ln(S_i) + \beta_i W + \gamma_i T + \varepsilon_i \qquad (1)$$

Where,

S_i = Size of the segment i (in lane-miles)
$W = 1$ if weather is clear and 0 otherwise (at the time of the crash)
T = traffic volume prevailing at the time of crash (veh/hour)
α_i, β_i, and γ_i = parameters of the model to be estimated from data
ε_i is the error term capturing the effect of unobserved factors
(i.e. other than size, weather, and traffic in this specification) on the propensity

For any location, the higher the value of the crash propensity (i.e., U_i), the greater is the probability that the crash will happen there. In fact, the probability that a crash will happen in a specific location is dependent on the relative propensity values of the different segments. Based on suitable assumptions about the error term, the mathematical expression for the probability that a crash will happen in any segment j $\left(\text{i.e., } \Pr(j)\right)$ can be stated as:

$$\Pr(j) = \frac{\exp\left(\alpha_j + \ln(S_j) + \beta_j W + \gamma_j T\right)}{\sum_i \exp\left(\alpha_i + \ln(S_i) + \beta_i W + \gamma_i T\right)} \qquad (2)$$

Therefore, the parameters of the model $\left(\alpha_i, \beta_i, \text{and } \gamma_i \forall i\right)$ capture the effect of the exogenous factors on the probability that a crash will happen in any segment. The coefficient on the “size” variable is fixed to 1 – this is an innocuous assumption required for model identification. In Equation 2, the βs capture the effect of weather, the γs capture the effect of traffic volume, and the αs are “alternative-specific constants” capturing the mean effect of all other (i.e., excluded) factors. These parameters can be estimated from data using the maximum-likelihood technique. In this analysis, the estimations were conducted using a procedure coded in the GAUSS programming language.

Once the model parameters are estimated, the relative safety of the different work-zone segments under different conditions can be assessed by calculating the crash probabilities on a per-lane-mile basis as follows:

$$\Pr(j \mid lane-mile) = \frac{\exp\left(\alpha_j + \beta_j W + \gamma_j T\right)}{\sum_i \exp\left(\alpha_i + \beta_i W + \gamma_i T\right)} \qquad (3)$$

It should be noted that Equation 3 was derived from Equation 2 by simply setting all the size variables $\left(\text{i.e.}, S_i \ \ \forall i\right)$ to 1.

4. DATA

This study focused on the I-95 Trout River Bridge reconstruction area. Further, only crashes in the northbound direction were included in the analysis. Therefore, the study region includes the mainline I-95 between State Street and Dunn Avenue (based on observed traffic-flow and queuing patterns, it was decided that the chosen segment would be adequate for analyzing the work-zone). A schematic sketch of the study region is presented in Figure 1 identifying the different segments of the work zone (The details of the roadway geometry such as the curves are not presented in this figure). Segment 1, the area before the advance warning, is one-mile long with three lanes; Segment 2, the advance-warning area is approximately 1 mile long and has three lanes; Segment 3, the transition area or the taper is half-mile long tapers from three to two lanes, Segment 4, the work area is approximately three miles and has two lanes, and Segment 5, the exiting area is one mile long and has two lanes.

Two major types of data were obtained from the study region. These are the crash data and the traffic data. Each of these is discussed further in the following subsections.

4.1 Crash Data

This study focused on crashes that occurred for the one-year period between July 1, 2006 and June 30, 2007. Crash reports were obtained from two primary sources: the Florida Highway Patrol (FHP) and the Jacksonville Sheriff's Office (JSO). This is because JSO typically assists with crash investigation on the interstate even though FHP has primary jurisdiction for patrol on

I-95 in the study area. Overall, 254 crash reports were compiled, of which 190 were from FHP and 64 from JSO. 50% of the 254 reports were "short form" and the remaining were "long form" crash reports. An electronic database of crash records was created by manually entering the required fields from paper copies of the crash reports. Approximately 74% of the crashes in the sample were from weekdays; 74% of the crashes occurred during daylight hours; and 39% of the crashes occurred during cloudy/rainy conditions.

The major task in the assembly of crash data was to create a detailed spatial depiction of crash locations. Information from the crash reports (such as the distance to the nearest intersection, the narrative, and the crash diagram in the long-form records), imagery from Google Earth, and the researchers' knowledge of the study area were used to determine the crash location as accurately as possible. The "placemark" functionality in the Google Earth application was used to determine the latitudes and longitudes of all the crashes. This "placemark" file was converted to a shapefile and subsequently imported into ArcGIS. Roadway features were added from the Florida roadway inventory and point features were then created to depict the locations of the advanced warning signs, the taper, and the beginning and end of the work area. Finally, each crash was designated to belong to one of the five work zone segments (analysis was done using the ArcGIS software). As already defined, these five segments are Before Advance Warning (B), Advance Warning (A), Taper (T), Work Area (W), and Exit Area (E).

It is useful to acknowledge that there might be some subjectivity in the overall procedure used to determine the crash locations. The different elements of spatial data from the crash report may not have been internally consistent and the researcher's (all crash mapping was done by the same person) knowledge of the study segment could have influenced his interpretation of the location described in the narrative. Therefore, there is the possibility of misclassifying the location of crashes that happen near the boundaries of adjacent segments. Detailed spatial data collected using GPS devices would help eliminate the subjectivity associated with the determination of crash locations. Until such methods are available, it would be necessary to resort to manual methods as adopted in this study.

4.2 Traffic Data

One of the objectives of this research is to assess the impact of traffic characteristics on the location of work-zone crashes. Unfortunately, such data were not available for the analysis period (i.e., for the period from July 1, 2006 to June 30, 2007). A proxy timeframe immediately after the crash-data collection period (June 28, 2007 through December 31, 2007) was deemed suitable as the work-zone set up in the study area was identical in both periods. A detector located at the beginning of the taper was used to measure the traffic conditions. Specifically, the 15-minute traffic volumes at different times of the day were averaged over several incident-free days to determine a representative daily profile for the traffic volumes. The profiles were generated separately for weekdays and weekend days. Figure 2 presents these average traffic-volume profiles for incident-free weekdays and weekend days. The two profiles represent averages over 35 weekdays and 15 weekend days respectively. The graphs indicate an evening-peak in the traffic volumes during weekdays and a mid-day peak during weekend days. Traffic volumes that the time of each crash were then estimated by looking up the value (from either the weekday or weekend graphs in Figure 2) for the 15-minute period immediately prior to the time of the crash. It is useful to recognize that there are day-to-day and seasonal variations in the traffic-flow patterns, and hence, the “average” traffic determined as described above many not accurately reflect the traffic conditions on any specific day. However, given the data limitations, the average values are used as the best estimate of the traffic conditions during the time of the crash.

5. ANALYSIS

This section of the paper comprises three major components. In the first component, descriptive statistics on the crash locations are presented. The second component presents the core contribution of this study; the results from an MNL model for crash location. Both these components examine the location of *all* crashes irrespective of crash type or injury severity. Although the MNL methodology can be extended to examine the location of crashes by crash type and injury severity, the limited data available to this study did not permit such an analysis. However, in the third component of this section, descriptive statistics on the crash characteristics are presented for the different segments of the work zone.

5.1 Descriptive Statistics on Crash Location

Table 2 presents the distribution of crashes across these work-zone segments. As noted before, simply looking at the fraction of all crashes in each segment could be misleading as the segments are typically of unequal lengths. Therefore, Table 2 also presents the number of crashes per lane-mile. Approximately one-third of all crashes occurred in the advance warning area and this translates into 10.5 crashes per lane mile. Thus, the advance warning region has the highest number of crashes in the overall and per lane-mile. The work area or activity area had the second highest number of crashes (30%); however, on a per lane-mile basis, this segment has the second lowest crash rate. The taper region represented 23 crashes or just over 9% of all crashes, but this region has the second highest crash rate per lane-mile (7.3).   The exit area and the "before advanced warning” area and each made up about 1 in 8 crashes, and represented 6.7 and 4.3 crashes per lane-mile respectively.

5.2 MNL Model for Crash Location

The empirical results for the MNL model for crash location are presented in Table 3. In the final specification, weather conditions and traffic characteristics turned out to be statistically significant and intuitively reasonable predictors of the location of crashes within work zones.

The negative coefficient on the variable “Clear Weather” corresponding to the propensity function for the “Advance Warning” area indicates that the probability of a crash in this area is lesser during times of clear weather compared to times of either rainy or cloudy weather (the small sample size did not permit further distinguishing the effects between rainy and cloudy conditions).

The effect of traffic characteristics is captured via two variables. The first variable, “Current Traffic Volume” is the average volume of traffic in the 15-minute period immediately prior to the crash (determined from the graphs in Figure 2.) The coefficient on this variable is positive for the “Advance Warning” area and negative for the “Exit” area. This indicates that the probability of a crash is higher in the former location and lower in the latter location during congested times. The second traffic-related variable is the “% Change in Traffic Volume”. This variable is calculated as the percentage difference between the “Current Traffic Volume” and the volume in the 15-minute period immediately preceding it. A large positive value of this measure indicates growing congestion; a large negative value indicates congestion dissipation; and small values (close to zero) indicate “off-peak” periods in which there is little fluctuation in the traffic volumes with time. The coefficient on this variable is positive corresponding to the propensity function for the “Work Area”. This indicates that the probability of a crash in the work area in higher in times of growing congestion.

This research also attempted to determine whether the locations of “secondary” crashes (i.e., those crashes that could have been caused, in part, due to a previous crash in the study area) were different from those of primary crashes. A crash was defined as “secondary” if it occurred within one hour and one-mile upstream of another (i.e., the “primary”) crash. Of the 18 such crashes identified 50% were in the Advance Warning area, 16% were in the Work area, and the rest were equally distributed across the three other segments. Unfortunately, this small sample size did not permit the estimation of statistically-significant parameters on this variable.

In addition, the effects of the time-of-day and day-of-the-week were also examined. However, these turned out to be statistically insignificant. It is possible that the impacts of these variables are indirectly captured by the traffic variables because of strong correlations among the traffic, time-of-day, and day-of-week variables. Nonetheless, it should be emphasized that these effects can be captured within the adopted modeling framework if adequate data are available.

In order to assess the relative safety of the different work-zone segments, *crash probabilities per lane-mile* were calculated using equation (3) for eight different scenarios (Table 4). The scenarios were constructed by varying weather conditions, day of the week, and traffic characteristics. The model parameters presented in Table 3 were used to calculate the probabilities. For each scenario, the crash probability values sum to 100% across the five work-zone segments. Table 4 indicates that there is considerable variation in crash probabilities (especially in the “Advance Warning” and the “Exit” areas) across the different scenarios. Bad weather and high traffic volumes make the “Advance Warning” area relatively unsafe and the combined effect of the two conditions is particularly critical (see rows 4 and 8). The “Exit” area is the most unsafe segment during off-peak conditions (see rows 1, 2, 5, and 6). The probability of a crash per lane-mile in the “Work Area” varies from 11% to 18% across the eight scenarios. The “Before Advance Warning” area seems to be the safest of all segments as indicted by the low probability values relative to the other segments.

In summary, the results from the MNL model clearly indicate that all work-zone segments do not have the same likelihood of crashes. In fact, the model indicates that the “Exit” area is relatively unsafe. It is possible that the crashes in this segment are tied to driver perceptions of reduced danger leading to increased speeding. Therefore, speed enforcement at the “Exit” areas of work-zones might be beneficial for improving the safety of work zones. Further, the analysis also indicates that the crash probabilities at the different segments vary based on external factors such as weather and traffic volumes. This result has implications for the design and prioritization of counter measures aimed at reducing work-zone crashes. Specifically, the results indicate the value of “dynamic” counter-measures that are sensitive to the prevailing conditions. For instance, posted speed-limits that vary based on weather conditions may be considered. Alternately, ITS technology such as variable-message boards could be activated at appropriate conditions/locations that warn the drivers of adverse conditions (such as queuing or visibility).

5.3 Descriptive Statistics on Crash Characteristics by Location

The descriptive statistics presented in Table 2 and the MNL model results presented in Table 3 have both examined the location of work zone crashes irrespective of the crash type or injury severity. However, for a comprehensive assessment of the safety of work zones, it is important to study not only the number of crashes but also the characteristics of these crashes.

The MNL modeling procedure presented in this paper can be extended to study the crash locations by type and injury severity. However, the small sample size available to the authors did not permit such an extension of analysis.

As a first step in the direction of a more detailed analysis of work zone crashes, Table 5 presents the distribution of crash type and injury severity by work zone location. Crashes in the Advance Warning and Taper area were most likely to be “two vehicle rear-end” crashes. Work area crashes were most likely to be “two vehicle other” although the likelihood of one- and three-vehicle crashes is also high (46% total). Overall, single vehicle crashes made up nearly 15% of crashes, with two-vehicle rear-end type crashes comprising the majority at 35.83%. Crashes involving two vehicles that were not rear-end type made up approximately 28% and those involving three or more vehicles were 21.26%. On examining the distribution of injury severity, one observes that the crashes in the Advance Warning and the Taper areas were likely to be less severe (80% no-injury crashes). In contract, crashes in work areas were more likely to be severe. The reader will note, from the MNL model, that the probability of a crash (per lane mile) is relatively low in the work areas compared to the other segments of the work zone. However, Table 5 highlights that such work-area crashes could be more severe than crashes in the other segments. Thus, an assessment of work zone safety requires an understanding of both the number of crashes as well as their characteristics.

6. SUMMARY AND CONCLUSIONS

This study modeled the location of crashes within work zones as a function of the lengths of the different work-zone segments, traffic volume, weather, and other exogenous factors. Data from crash reports were augmented with spatial attributes by using GIS. The empirical results indicate that weather conditions and traffic characteristics are statistically-significant and intuitively-reasonable predictors of the location of crashes within work zones. The model was also applied to assess the relative safety of the different work-zone segments in terms of the probability of a crash per lane-mile of the segment. The results indicate that the "Advance Warning" area is particularly unsafe (relative to other segments) during times of peak traffic flow and bad weather. Further, the "Exit" area was also found to be relatively unsafe, especially during off-peak periods. Therefore, speed enforcement at the "Exit" areas of work-zones might be beneficial for improving the safety of work zones. The probability of a crash per lane-mile in the "Work Area" was generally smaller compared to the other work-zone segments but varied from 11% to 18% across eight scenarios evaluated. At the same time, the 'Work Area" crashes were found to more severe in terms of the injury sustained by the vehicle occupants. Overall, the results from this study have implications for the design and prioritization of counter measures aimed at reducing work-zone crashes. In particular, it is found that the critical segments (in terms of safety) of the work-zone vary based on exogenous factors such as weather and traffic. Therefore, countermeasures should be focused on different segments depending on the time-of-the-day and season of the year.

While the methodology employed in this study is generic, the data were from a single Florida work zone. Thus, the empirical results may not be true for all work-zones. Therefore, an empirical extension of this research that includes a wide variety of work zones is recommended. However, an important deterrent to pursuing such a study is the lack of detailed data on crash locations in the police reports. In this study, the latitudes and longitudes of the crashes were determined using manual methods. This was feasible given the small size of the sample and the researchers' knowledge of the study region. In a large-scale study covering several work-zones, such an approach may not be efficient or even feasible. Therefore, improving the spatial accuracy of the crashes locations in police reports using the GPS technology would be significantly beneficial for furthering the analysis of work zone safety. This would also remove any subjectivity in the determination of the precise location of the crashes. Subsequently, the methodology adopted in this research can be used to develop rigorous models for work-zone safety, which, in turn, would inform policy decisions.

ACKNOWLEDGEMENTS

The Southeastern Transportation Center (STC) provided funding for this research (Grant #: DTRS99-G-0004). The authors would also like to acknowledge feed back provided by Dr. Scott Washburn on an earlier draft of this manuscript.

REFERENCES

Ben-Akiva, M and Aboud Zeid, M, 2007. Methodological Issues in Modeling Time-of-Travel Preferences. *Presented at the 11th World Conference on Transportation Research.*

FDOT, 2007. FDOT Continues Safety Push for National Work Zone Awareness Week, http://itseveryonesjob.com/media/2007NationalWorkZoneAwarenessWeek.pdf. Access Date: Nov 15, 2008

FHWA, 2008. Ninth Annual National Work Zone Awareness Week 2008, http://safety.fhwa.dot.gov/wz/nwzaw_events/factsheet08.htm. Access Date: Nov 15, 2008

Garber, N.J and Zhao, M, 2002. Distribution and Characteristics of the Crashes at Different Work Zone Locations in Virginia. *Transportation Research Record*, No. 1794, pp. 19-25

Hargroves, B. T., 1981. Vehicle Accidents in Highway Work Zones. *Journal of Transportation Engineering,* Vol. 107, No. TE5, pp. 525-539.

Nemeth, Z. A. and Migletz, D. J., 1978. Accident Characteristics Before, During, and After Safety Upgrading projects on Ohio's Rural Interstate System. *Transportation Research Record 672*, pp. 19-23.

Nemeth, Z.A. and Rathi, A., 1983. Freeway Work Zone Accident Characteristics. *Transportation Quarterly*, Vol. 37, No. 1, pp. 145-159.

NHTSA (National Highway Traffic Safety Administration), 2003. Guideline for Minimum Uniform Crash Criteria (MMUCC), 2nd Edition. Washington, D.C.: National Highway Traffic Safety Administration. .

Pigman, J. G. and Agent, K. R., 1990. Highway Accidents in Construction and Maintenance Work Zone. *Transportation Research Record 1270*, pp. 12-21.

Washburn, S. S. and Carrick, G. T., 2006. Data Collection Needs for Work Zone Incidents. *Final Report. Southeast Transportation Center, Univ. of Tennessee, Knoxville, TN*. pp. 104.

Table 1. Location of Crashes within Work Zones: Summary Findings from Literature

Study	Location	Sample Characteristics	Sample Size	Percentage of all Crashes in Location..			
				Advance Warning	Taper	Work Area	Exit Area
Garber and Zhao (2002a and 2002b)	Virginia	All Crashes	1484	10.00	13.00	75.00	2.00
		Day-Time Crashes	1085	12.00	14.00	72.00	2.00
		Night-Time Crashes	399	6.00	12.00	80.00	2.00
		Urban Interstate Crashes	544	7.00	17.00	75.00	1.00
		Rural Interstate Crashes	159	14.00	14.00	72.00	0.00
		Urban Primary Crashes	339	7.00	10.00	80.00	3.00
		Rural Primary Crashes	206	18.00	16.00	63.00	3.00
		NOVA Urban Secondary Crashes	94	10.00	5.00	80.00	5.00
		Other Secondary Crashes	140	14.00	10.00	75.00	1.00
Pigman and Agent (1990)	Kentucky	All Crashes	1361	8.30	11.70	80.00	NA
Nemeth and Rathi (1983)	Ohio Turnpike	All Crashes	185	6.40	9.20	23.20	NA
		Night-Time Crashes	80	1.25	7.50	16.25	NA
Hargroves (1981)	Virginia	All Crashes	566	12.70	85.90		1.40
Nemeth and Migletz (1978)	Ohio Rural Interstates	All Crashes	151	15.90	22.50	55.60	NA

NOTE (1) For the Nemeth and Rathi (1983) study, the percentages do not sum to 100 -- a large number of the crashes were classifed as happening in locations such as "crossover zones" and "bi-directional zones"; locational descriptions which are rather specific to this study

NOTE (2) For the Nemeth and Migletz (1978) study, the percentages do not sum to 100 -- about 6% of the crashes were in ramps and such "other" locations

Table 2. Sample Characteristics

Crash Location	Frequency	%	Crashes per Lane mile
Before Advance Warning (B)	33	12.99	4.3
Advance Warning (A)	88	34.65	10.5
Taper (T)	23	9.06	7.3
Work Area (W)	76	29.92	5.2
Exit Area (E)	34	13.39	6.7
Total	254	100	-

Table 3. Empirical Model Results

	Parameters of the model corresponding to the propensity fucntions for the different segments									
	Before Advance Warning (B)		Advance Warning (A)		Taper (T)		Work Area (W)		Exit Area (E)	
	Param.	t stat.	Param.	t stat.	Param.	t stat.	Param.	t stat.	Param.	t stat.
Constant (α)	-1.195	-2.262	-0.933	-1.507	-0.676	-1.250	-1.030	-2.009	0.000	fixed
Ln(Size)	1.000	fixed	1.000	fixed	1.000	fixed	1.000	fixed	1.000	fixed
Clear Weather (β)	-		-0.770	-2.791	-		-		0.000	fixed
Current Traffic Volume (γ)	0.000	fixed	0.149	2.421	-		-		-0.118	-1.561
% Change in Traffic Volume (δ)	-		-		-		0.036	1.648	0.000	fixed
Goodness-of-Fit Measures										
Log likelihood for final model	-364.3									
Log likelihood for constants-only model	-375.945									
Rho-squared w.r.t constants-only-model	0.0201									

NOTE(1) "fixed" parameters are those that are assumed to have specific values -- an innocuous assumption but necessary for model identification

NOTE(2) "-" indicates that the corresponding parameter was found to be statistically insignificant

Table 4. Model Application: Crash Probabilities (%) per Lane-Mile of Work-Zone Segments

S. No.	Scenario Description	Predicted probability of a crash (per lane-mile) in segment…...				
		Before Advance Warning (B)	Advance Warning (A)	Taper (T)	Work Area (W)	Exit Area (E)
1	Weekday, Off-peak, Clear Weather	13.36	9.37	22.45	15.74	39.08
2	Weekday, Off-peak, Rainy/Cloudy Weather	12.05	18.25	20.25	14.20	35.25
3	Weekday, Peak, Clear Weather	13.12	36.34	22.05	15.49	13.01
4	Weekday, Peak, Rainy/Cloudy Weather	9.23	55.21	15.51	10.90	9.15
5	Weekend-day, Off-peak, Clear Weather	13.36	9.39	22.45	15.76	39.04
6	Weekend-day, Off-peak, Rainy/Cloudy Weather	12.05	18.28	20.25	14.21	35.21
7	Weekend-day, Peak, Clear Weather	14.29	23.39	24.00	16.89	21.43
8	Weekend-day, Peak, Rainy/Cloudy Weather	11.24	39.73	18.88	13.29	16.86

The peak period is taken as 4:45 PM for weekday and 1:45 PM for weekend. The off-peak period was taken as 3:15 AM for weekday and 3:35 AM for weekend. For these times, the traffic volumes and % changes in traffic were obtained from the graphs in figure 2

Table 5. Percentages of Crashes by Type, Injury Severity, and Crash Location

	Crash Type			
	Single Vehicle	Two Vehicle Rear-end	Two Vehicle Other	Three or More Vehicles
Before Advance Warning (B)	9.09	36.36	30.30	24.24
Advance Warning (A)	13.64	47.73	17.05	21.59
Taper (T)	21.74	39.13	30.43	8.70
Work Area (W)	21.05	19.74	34.21	25.00
Exit Area (E)	5.88	38.24	38.24	17.65
Total	14.96	35.83	27.95	21.26
	Injury Severity			
	None	Possible	Non-Incapacitating	Incapacitating
Before Advance Warning (B)	75.76	15.15	9.09	0.00
Advance Warning (A)	81.82	7.95	9.09	1.14
Taper (T)	82.61	13.04	4.35	0.00
Work Area (W)	75.00	14.47	6.58	3.95
Exit Area (E)	82.35	5.88	8.82	2.94
Total	79.13	11.02	7.87	1.97

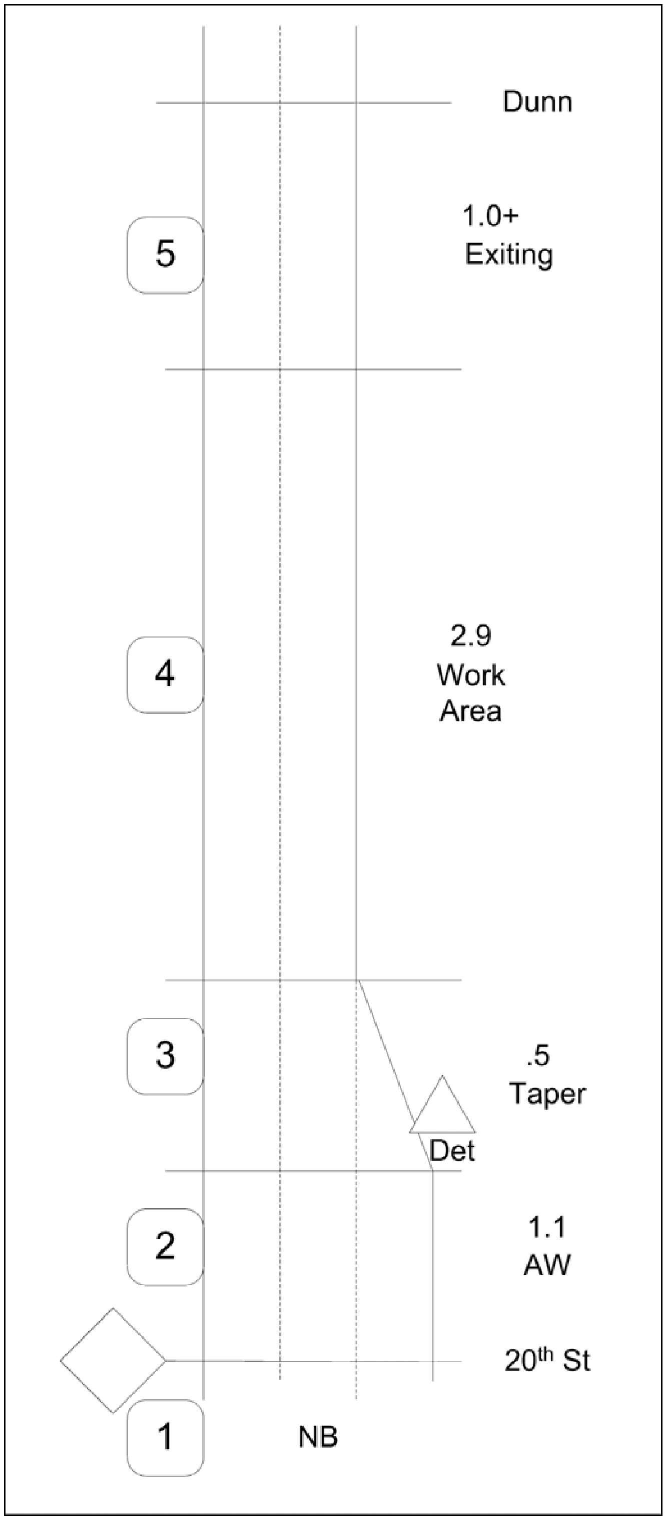

Figure 1. Schematic Representation of the Study Area

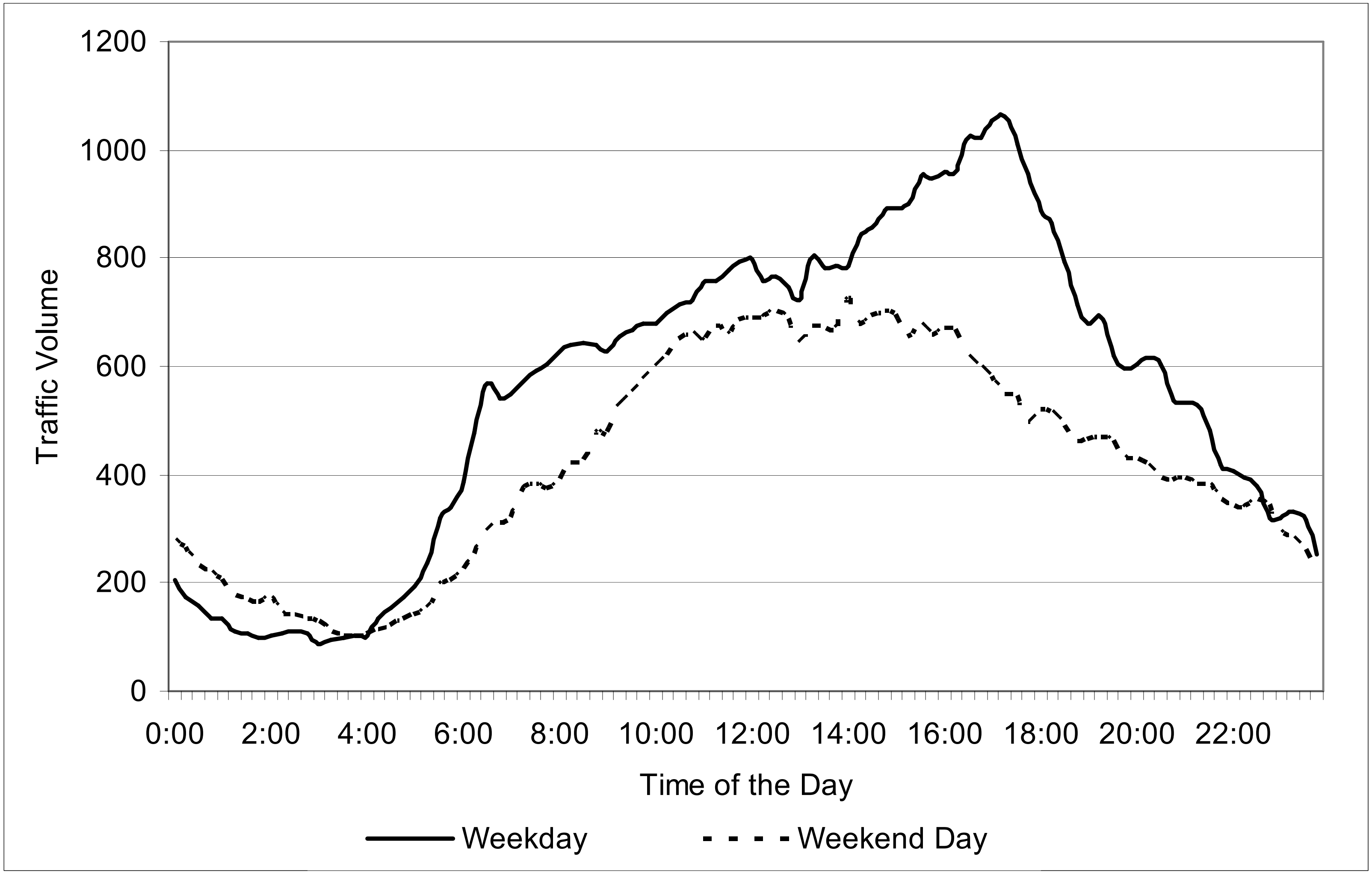

Figure 2. Profile of Traffic Volumes