CURS Report No. 2003-04

# Understanding the Role of Truck-Driver, Occupational and High-Risk Roadway Factors in Truck-Involved Collisions

**Asad J. Khattak, Ph.D**
Carolina Transportation Program & Department of City and Regional Planning
CB 3140 New East Building, UNC
Chapel Hill, NC 27599
Tel (919) 962-4760
Email khattak@email.unc.edu

**Daniel Rodriguez, Ph.D.**
Carolina Transportation Program & Department of City and Regional Planning
CB 3140 New East Building, UNC
Chapel Hill, NC 27599
Tel (919) 962-4763
Email danrod@email.unc.edu

Felipe Targa, MRP
Carolina Transportation Program & Department of City and Regional Planning
CB 3140 New East Building, UNC
Chapel Hill, NC 27599

Marta Rocha, MRP
Carolina Transportation Program & Department of City and Regional Planning
CB 3140 New East Building, UNC
Chapel Hill, NC 27599

August 2003

---

**Final Report Submitted to:**

**Southeastern Transportation Center**
**University of Tennessee**
**Knoxville, Tennessee**

# UNDERSTANDING THE ROLE OF TRUCK-DRIVER, OCCUPATIONAL AND HIGH-RISK ROADWAY FACTORS IN TRUCK-INVOLVED COLLISIONS

**Asad J. Khattak, Daniel Rodriguez, Felipe Targa & Marta Rocha**
University of North Carolina at Chapel Hill, NC 27599

This report develops a comprehensive understanding of the truck-safety problem from the perspective of:

1) Roadway hazards, involving analysis of truck-involved crashes in work zones;
2) Socio-economic and driver occupational factors, involving examination of driver and work characteristics on truck crash occurrence and frequency.

More specifically, the study is split into two parallel but complementary efforts, one relying on Highway Safety Information System crash and road inventory data and another looking at a longitudinal dataset of crashes, truck-driver characteristics and occupational factors obtained from a major for-hire truck-carrier. Two manuscripts resulting from each parallel research effort are included in the body of the report.

The first effort analyzes how work zone attributes (such as type of work zone, presence of warning signs and cones, type of activity in the work zone, location of the crash in the work zone, construction impact of the wok zone on the roadway and the type of work being done in the work zone) influence crash injury risk and specifically examine the factors that increase large truck crash involvement in work zones. In 2000, North Carolina changed their accident forms to include several new work zone related variables. The federally-sponsored Highway Safety Information System (HSIS) contains this work zone, crash and inventory data used in the project. The analysis was done separately for truck-involved and non-truck-involved collisions at the crash and vehicle levels. We found that injurious work zone crashes were 1) those involving a truck when the roadway was closed, requiring a detour on the opposite side, 2) before the actual work area, where traffic moves out of its normal path, 3) on two-way undivided roads and 4) when a Stop/Yield/Warning flashing sign was present. The results suggest a clear link between work zone characteristics and injury risk.

The second effort examines driver, compensation, and occupational factors that can increase the risk of truck-involved collisions, while controlling for exposure. This part of the study analyzes the dynamics of truck drivers' collision involvement through a unique longitudinal data set. By focusing on the role of driver and occupational factors in crash risk and severity, the study generates knowledge that is useful for public decision makers as well as the trucking industry. The dataset is unique because it contains human resources, operations, and safety data for more than 11,000 unscheduled over-the-road tractor-trailer drivers of a major U.S. for-hire truckload company over a 26-month period. The analysis is done at the driver level, examining the risk of crash involvement. The results indicate that higher truck driver compensation is strongly associated with better safety records. Although the data for the study come from a single firm, the evidence provided is a first step in examining the structural causes of unsafe driving behavior, such as driver compensation. These results can motivate other firms in modifying operations and driver hiring practices. Furthermore, to the extent that generalizations about the truckload sector can be made from this study, our findings suggest that human capital characteristics are important predictors of driver safety, but that motivational and incentive factors also play an important role in determining the safety outcomes of truck drivers.

# INJURY SEVERITY AND TOTAL HARM IN TRUCK-INVOLVED WORK ZONE CRASHES

PART 1: CRASH LEVEL ANALYSIS

**Asad J. Khattak & Felipe Targa**

University of North Carolina at Chapel Hill

**Abstract—**Society pays a high cost for work zone crashes in terms of operational disruptions, property damage, injuries and loss of life. Given narrow lanes in work zones, large trucks are of particular concern. This paper empirically examines truck-involved collisions, comparing them to non truck-involved collisions in North Carolina work zones. The paper helps understand which work zone attributes are statistically associated with the most seriously injured occupant and total harm in the crash. Specifically, the role of several new work zone variables is explored. They include the type of work zone, presence of warning signs and cones, type of activity in the work zone, location of the crash in the work zone, construction impact of the wok zone on the roadway and the type of work being done. These variables were obtained from a unique dataset based on a revised North Carolina accident form that included new work zone variables and additional coding of variables from the crash report narrative. Using this new work zone data for collisions during the year 2000 (N=3,383), we estimated ordered probit models for the most seriously injured occupant in the crash and linear regression models for the "total harm" in the crash. The models for total harm in the crash combine both frequency and severity of injuries by assigning economic cost to injuries. Overall, work zone crashes in North Carolina were more injurious compared to non-work zone involved crashes, especially those involving large trucks. Modeling results suggest that the most injurious work zone crashes were 1) those involving a truck when the roadway was closed, requiring a detour on the opposite side—there was a 38.5% higher chance of injury compared to if the roadway was not closed, 2) before the actual work area, where traffic moves out of its normal path, compared with crashes in the advance warning area or adjacent to the actual activity/work area, 3) on two-way undivided roads—there was a 19.1% higher chance of injury compared to two-way divided and protected roadway and 4) when a Stop/Yield/Warning flashing sign was present. Other crash- and environment-related effects are discussed in the paper. The results provide valuable information for work zone strategies intended to improve vehicle occupants' safety by reducing their injury severity.

---

**Keywords—**safety, work zones, collisions, injury, large trucks

## INTRODUCTION

Highway work zones are increasing due to deterioration of the highway infrastructure, increased maintenance spending, and greater attention to the management of existing highway assets. With the completion of the interstate highway system, a significant portion of all federal-aid highway funds in the United States are now geared toward roadway maintenance and rehabilitation. Additionally, due to the increase in traffic volumes, closing long segments of the roadway for rehabilitation is becoming very challenging. Indeed, 13% of the national highway system is under construction or maintenance at any given time (*1*). The presence of workers, construction machinery, roadside construction barriers, and other paraphernalia associated with work zones create a high degree of conflict that leads to hazardous conditions for vehicle occupants. These trends result in a significant increase of drivers' exposure to hazardous roadway environment demanding increased efforts to improve safety in work zones.

Society already pays a high cost for work zone crashes in terms of operational disruptions, property damage, injuries and loss of life. In 2000, work zone crashes across the U.S. resulted in 829 fatalities (3% of nation's total) and 52,000 people injured (2% of nation's total). Large truck-involved collisions accounted for 30% of work zone crashes and 24% of work zone fatalities (*2*). For North Carolina, 3% of total fatalities occurred in work zone areas (30% from large truck-involved collisions), showing similar outcomes.

The economic cost of a motor vehicle crash involving a fatality in the U.S. is estimated at 3 million dollars (*3, 4, 5*). Based on this estimate, the annual cost of work zone fatalities alone amounts to more than 2.5 billion dollars. Furthermore, costs for wage and productivity losses, medical and administrative expenses, and property damage from non-fatal injury crashes result in additional billions of dollars. Accordingly, mitigating injuries resulting from crashes in work zones is a national priority.

Owing to their larger size, limited maneuverability and sluggish performance, tractor-trailer trucks may experience greater increases in work zone crashes compared with smaller vehicles and truck-involved crashes may be relatively more injurious in work zones. Also, large truck-involved crashes in work zones have a tendency to be high-profile and newsworthy events, especially if they involve fatalities. So this paper examines the possible increase in severity of truck-involved crashes. Importantly, we know very little about how the attributes of work zones affect work zone safety, especially when large trucks are involved. Understanding the role of work zone attributes in injury risk and harm can potentially reduce the associated economic costs. In this paper, we empirically investigate the effects of work zone characteristics on the most seriously injured occupant and the total harm in the crash for truck-involved and non truck-involved collisions. Our objective is to identify work zone related risk factors that can reduce injury severity, given a crash.

## LITERATURE REVIEW

The literature provides significant information about rising work zone crashes and fatalities in the United States (*2, 6, 7, 8*). A majority of the studies regarding work zone crashes have focused on analyzing crash rates for pre-construction and during work zone operation periods. All of these studies have found that crash rates during work zone operation were higher compared with the pre-work zone period (*9, 10, 11, 12, 13, 14, 15*). Some of them found that a combination of cones, flashing and static signs, and flaggers could lower work zone crash rates. Furthermore, rear-end and sideswipe collisions occurred more frequently in work zones than in non-work zones, as expected.

Despite many work zone studies, the effect of work zone attributes on injury severity remains largely unexplored. The limited research available has analyzed the distribution and characteristics of injury and non-injury work zone crashes. These studies show that work zone crashes are slightly less severe than non-work zone crashes (*12, 13, 16, 17, 18*). However, inconsistencies appear with other studies showing that work zone crashes are more severe (*19, 20*) or that there is no significant difference between work zone and non-work zone crash severity (*14, 15*).

In terms of work zone attributes, Khattak et al. (*21*) found that longer work zone duration significantly increases both injury and property-damage-only crash frequencies. The location of the crash in the work area (e.g., activity area) is also significantly more severe, in both injury and fatal crashes (*22*).

Multi-vehicle collisions were consistently overrepresented in work zones (*13, 14, 15, 20, 22, 23*). This overrepresentation was also significant for large truck-involved collisions in work zone crashes (*14, 16, 19*). Moreover, large truck-involved collisions in work zones were more severe compared with non truck-involved collisions (*19, 24*).

Although crash rates often increase in work zones, questions about the effect of work zone attributes on the severity of these crashes are unanswered. A major gap in the literature seems to be lack of knowledge regarding work zone related contributing factors and their relationship with injury severity. While some studies have found that crashes involving large trucks in work zones are important contributors to fatalities, it is still unclear what specific work zone characteristics influence injury severity for truck-involved and non truck-involved collisions.

## DATA AND METHODS

Using work zone accident data for truck-involved and non truck-involved collisions, this paper explores the effect of work zone characteristics on qualitative differences between different injury categories. The study requires information on work zone characteristics, their influence on roadway configuration, and other crash and environmental attributes. The study uses data from the Highway Safety Information system (HSIS) with additional variables coded from accident reports. Developed and maintained by the Federal Highway Administration (FHWA), HSIS contains crash, and roadway inventory data from selected states. We used year 2000 HSIS data for North Carolina because it contained new work zone-related variables (see Figure 1 for a typical work zone location). These variables are:

- If the work zone was a contributing circumstance to the crash;
- Crash location, i.e., before work area, in work area approach taper, and adjacent to actual work area;
- If the crash occurred in or near a construction work area, maintenance work area, utility work area, and intermittent/moving work (e.g., patching pothole);
- If there was on-going work activity at the time of the crash; and
- If the work area was marked with warning signs and cones.

Additional work zone-related data, not included in HSIS, were coded directly from police officer accident reports (Crash Report Forms DMV-349) available from the North Carolina Division of Motor Vehicles (DMV). We were able to code two additional variables based on the narrative and crash diagram reported by the police officer. The additional coded variables were:

- If the construction had an effect on the roadway, i.e., lane closed, shoulder/median closed, roadway closed, lane shift, and other; and
- The type of work being done, i.e., repaving/resurfacing, shoulder/median work, new roadway, and other.

Because the process of downloading and coding DMV crash report forms required significant time and effort, we limited the coding of new variables to truck-involved collisions only. Overall, we developed a unique dataset that contains several new work zone-related variables, allowing us to explore their effects of injury severity.

A majority of the HSIS accident-file data required recoding and variable manipulation in order to isolate relevant factors for the study. Most of the independent variables, which included work zone/roadway, accident, and environmental factors, were categorical and were recoded as binary variables. Only posted speed limit, number of vehicles and number of persons involved in the crash were the continuous variables.

Injury severity is measured on the KABCO scale, which is ordinal; it was coded as 4 (K-killed), 3 (A-injury; severe incapacitating injury), 2 (B-injury; moderate or non-incapacitating injury), 1 (C-injury; minor or possible injury), and 0 (O-no injury; property damage only). Two measures of crash severity were explored (see Figure 2). One measure was the most seriously injured occupant in the crash. The ordered probit model is appropriate for such data, i.e., it can capture the qualitative differences between different injury categories. For example, it can capture the effect of a particular variable such as work zone activity on the likelihood of a fatality, differently from its influence on the likelihood of a minor or incapacitating injury (*25, 26*). The models interpretation is based on coefficient estimates and marginal effects or discrete changes in predicted probabilities for injury severity. Computation of marginal effects and discrete change for dummy variables is particularly meaningful for ordered probit models because the effect of explanatory variables on the intermediate injury categories is ambiguous if only the parameter estimates are available. For ordered probit models we use a pseudo-$R^2$ measure proposed by McKelvey & Zavoina (*27, 28*). The advantage of this goodness of fit measure is that it mimics the OLS $R^2$.

Another measure explored in the study is that of total harm in the crash. To create this variable, we assigned an economic value to each injury level and summed the costs for each injury (or property damage). Therefore, the total harm variable combines both injury frequency and injury severity in a crash. The injury and property damage costs for North Carolina were obtained from The Children's Safety Network Economics and Insurance Resource Center (*29*). They include:

- Medical costs, i.e., hospital, physician, rehabilitation, prescription and related cost.
- Pain, suffering and quality of life that the family loses because of a death or injury.
- Emergency service costs, i.e., police, fire, and ambulance and helicopter services.
- Victim work loss costs including wages, fringe benefits and household work.
- Employer costs including value of time lost, and the extra work and distractions for supervisors and coworkers that injuries cause.
- Traffic delay costs including values of time lost in traffic jams caused by crashes.
- Property damage costs including costs to repair or replace damaged vehicles and property.

The estimated costs of motor-vehicle crashes including quality of life values is $2,925,100 for fatal K-injury, $144,796 for A-injury, $37,486 for B-injury, $17,916 for C-injury and $3,904 for property damage only (*29*). The ordinary least squares regression is the appropriate technique for modeling these cost data. Taking the log-transform of cost avoids negative predictions. Also the log-transform collapses the range of the original variable, pulling in the right tail, i.e., higher cost values in the transformed distribution are relatively closer to the lower ones. This transformation implies that every unit change in the independent variable is expected to multiply the original dependent variable by $e^b$. $b$ is the estimated coefficient; note that the percentage increase in cost is given by ($e^b$-1). In general, modeling allows us to explore the impacts of work zone attributes while controlling for roadway, crash and environmental factors as well as exposure factors, such as the number of vehicles involved in the crash, and number of people involved in the crash.

Conceptually, we expect that due to narrower lanes, more objects to strike, greater potential for distractions unexpected turns/lane drops, work zones will increase injury risk, given a crash. However, this effect might be mitigated if drivers slow down and drive more carefully in work zones. Work zones are likely to exacerbate the potential for truck-car collisions, given their large mass differentials, their close proximity in work zones as well as limited maneuverability and sluggish performance of large trucks. Furthermore, work zones that have more complex layouts might experience more injurious crashes. Typically, lower functional class roads are also more dangerous. Accordingly, work zones on two-lane undivided roadways might be particularly harmful. Certain locations within the work zone might be more dangerous than others due to the presence of more objects and/or more activity.

## RESULTS

### Descriptive Statistics

Table 1 shows that the proportion of fatalities (among fatal and non fatal injuries) in North Carolina work zone crashes was slightly higher (1.38%) compared with the proportion of fatalities in the total NC crashes (1.03%). Also, the proportion of fatalities in NC truck-involved work zone crashes was higher than the proportion of fatalities in work zone crashes. Thus, work zone crashes are more injurious than non-work zone crashes, consistent with other studies (*19 & 20*), and truck-involved crashes are more injurious in work zones (*19, 24*).

During 2000, 3,383 crashes occurred in North Carolina work zones; 675 (20.0%) were single-vehicle collisions and 2,708 (80.0%) involved more than one vehicle (Figure 2). Out of the 3,383 collisions, 660 (19.5%) involved at least one truck (88 single-vehicle and 572 multi-vehicle), and 2,669 (80.5%) did not involve any truck but involved a car, a Sports Utility Vehicle (SUV), or a van (553 single-vehicle and 2,116 multi-vehicle collisions). Though most crashes resulted in no injuries (PDO crashes ranged from 63% in non truck-involved crashes to 56% in truck-involved single-vehicle crashes), multi-vehicle truck-involved crashes appear to be the most injurious (Figure 3). Specifically, 1.05% crashes turned out to be fatal and 4.38% severe (injury type A) in truck-involved multi-vehicle collisions. Moreover, if we account for all injuries in the crash by assigning a dollar value as shown above, truck-involved multi-vehicle collisions are the most costly, on average, costing $81,300 per crash (Figure 3).

A brief description of the coding scheme for the variables and descriptive statistics for truck-involved and non truck-involved collisions is presented in Table 2. Compared with non truck-involved collisions, truck-involved collisions occurred more frequently when work zone activity was on-going and there was no traffic control device present. Moreover, truck-involved collisions occurred more frequently on two-way divided and protected roadways, as opposed to two-way undivided configurations. These relationships were investigated more rigorously by estimating regression models.

Although truck-involved collisions were 675, truck-involved single-vehicle collisions were only 88, precluding any modeling of single-vehicle collisions. Cross tabulations and their Chi-square tests for independence indicate that there is a significant negative association (at the 95% level of confidence) between the *Work Zone* variable (roadway contributing circumstances) and injury severity. This association suggests that if work zones have some contributing effect, in the opinion of the police officer, then the crash is less severe compared with other contributing circumstances (including “none”). Similarly, the *Rollover* variable (most harmful event) is significant and positively associated with injury severity, as expected. The rest of the variables for single-vehicle truck-involved collisions were not statistically significant. Given that single-vehicle truck collisions are the least costly ($16,800 per crash; see Figure 3), they seem the least problematic among work zone crashes.

## Modeling Results

Table 3 reports estimates from six models for multi-vehicle collisions: Three ordered probit models for injury severity and three OLS log-transformed models for total harm in the crash. Each three-set of models is estimated for the pooled data, truck-involved collisions and non truck-involved collisions. Models for non truck-involved collisions are the comparison group. Model interpretation for truck-involved collisions is also based on marginal effects (Table 4) and predicted probabilities for the ordered probit model. We focus on discussing model results for statistically significant coefficients at the customary 99%, 95% and 90% confidence levels.

All models are statistically significant as indicated by F-tests for the OLS models and Likelihood-Ratio tests for the probit models significant at the 99% confidence level. Adjusted $R^2$s range from 0.293 (all collisions, pooled data) to 0.343 (truck-involved collisions) for the OLS log-transformed models; and McKelvey & Zavoina pseudo-$R^2$s range from 0.146 (non truck-involved collisions) to 0.242 (truck-involved collisions) for the ordered probit models. These indicate reasonable goodness-of-fit.

Among the work zone/roadway-related variables coded from the DMV crash report forms (only for truck-involved collisions), *Roadway closed* (construction effect on the roadway) is significantly associated with higher injury and total harm. Roadway closed increases the cost of a crash by a factor of 143% ($e^{0.889}$-1) compared with other crashes. (Note that the intercept of the log-regression is difficult to interpret because it gives an estimate of cost when all independent variables are zero. However, it is positive ($e^{7.379}$ = $1602) and more reasonable compared with the negative constant in the equivalent (unreported) OLS model. The marginal effects for the injury severity model (Table 4) suggest that if a crash occurs in a work zone where the roadway is closed, requiring detour on the opposite side, then the chances of injuries increase by 38.5%, i.e., the chances of minor injuries are 9.5% higher, moderate injuries are 13.4% higher, severe injuries are 12.5% higher, and a small increase of 3.2% in the chances of fatalities.

For truck-involved collisions, the crash location in the work zone is a significant factor in injury severity and total harm. Truck-involved crashes that occurred before the work area or transition area, where traffic moves out of its normal path, or adjacent to the work area are associated with higher total harm than crashes in the approaching taper or advance warning area (see Figure 1). The harm model suggests that crashes before the work area cost 33% ($e^{0.282}$-1) more while those in the activity/work area cost 23% ($e^{0.210}$-1) more than crashes occurring in the advance warning area. Table 4 suggests that the chances of injuries are higher by 11.9%, (i.e., the chances of minor injuries are 6.2% higher, moderate injuries are 3.5% higher, severe injuries by about 2% higher and fatalities by 0.2%) if the if the truck-involved crash occurs adjacent to the actual activity/work area compared with a crash in the advance warning area. This warrants greater attention to the activity/work area as a cause of truck-car collisions.

Work zones located on certain roadways are associated with higher injury severity and total harm. Specifically, work zones located on two-way not divided roadways are associated with 37% ($e^{0.313}$-1) higher costs compared with work zones on two-way divided and protected roadways. This association is stronger for truck-involved than non truck-involved collisions, as indicated by the parameter estimates for the injury severity model. The chances of injuries are 19.1% higher, if a truck-involved crash occurs in a work zone located on a two-way not divided roadway. For truck-involved crashes only, a two-way divided but unprotected configuration was also more injurious than a two-way divided and protected configuration, increasing the chances of injuries by 13.3%. Thus to reduce injury severity and harm in truck-involved collisions, work zones located on two-way undivided and two-way divided but unprotected roadways need particular attention.

Higher posted speed limit in the work zone is associated with higher injury severity and total harm, as expected. In terms of harm, every 10 mph increase in the authorized speed limit increases the total cost of the crash by 15%. For the same 10 miles per hour increase in the authorized speed limit, the chances of injuries in a crash increase by 8.0%. Figure 4 illustrates the effect of the increase in the work zone posted speed limit on the predicted probabilities of being injured (fatal and non-fatal injuries). It shows that the probability of injury increases substantially with higher posted speed limits, given a crash. This increase is somewhat uniform across truck-involved and non truck-involved crashes. Therefore, careful consideration needs to be given to setting up the speed limits in work zones.

The presence of stop/yield/warning flashing signs (traffic control devices) is associated with higher injury severity and harm, particularly for truck-involved collisions. However, this effect may not be causal, i.e., these devices typically prevent even more severe crashes from occurring. The presence of these devices is often associated with larger work zones and higher levels of hazardous activity at work zones. Nevertheless, the coefficient estimates suggest that crashes occurring in the presence of traffic control devices are more injurious and harmful than no control device present, so work zones with these control devices need to be further investigated.

In addition to work zone attributes, several other variables were controlled for. Crash-related factors such as the most harmful event in the crash was associated with higher injury severity and total harm. Also collisions with pedestrians, bicyclists, or animals are more injurious than collisions with other motor vehicles, mainly because the most seriously injured person is either the pedestrian or the bicyclist. Head-on and angle collisions are also more severe than other type of "more than two-vehicle" collisions. Crash exposure measures, such as number of vehicles and number of people involved in the crash are significantly associated with higher injury severity and total harm, as expected. Finally, environmental factors, such as weather and ambient light are significant but only for non truck-involved collisions.

## LIMITATIONS

If non-reporting of crashes varies systematically for work zone collisions, or any other type of collision such as truck-involved or non truck-involved, then it can bias the results. One could hypothesize both higher reporting in work zones (due to potentially higher police presence and less room to move damaged vehicles to the shoulders), and lower reporting (if vehicles in non-injury crashes are encouraged to leave the scene in order to decrease queues). Despite this potential bias, the accident data comes from a relatively clean, federally maintained Highway Safety Information System, there was very little missing data and the descriptive statistics and model results are quite reasonable.

## CONCLUSIONS

The purpose of this study was to investigate large truck-involved crashes in work zones. Using a unique dataset that contains new work zone-related variables, not available before in police crash reports and coding additional variables from reports, we present empirical evidence regarding injury and harm risks associated with work zone attributes. Given work zone crashes, we found that multi-vehicle truck-involved collisions were the most injurious and harmful. Harm is a measure that combines the effect of both frequency and severity of injuries, and helps estimate the economic costs.

Rigorously modeling injury and harm provided new insights into the role of work zones. Truck-involved multi-vehicle crashes were most injurious and harmful when the roadway was closed, requiring a detour on the opposite side. Furthermore, work zones located on a two-way not divided roadway substantially increased the chances of injuries compared with other roadway configurations. Within the work zone, truck-involved crashes occurring immediately before the actual work area or adjacent to the activity/work area were more injurious and harmful, compared to those occurring in the advance warning area. Higher speed limits was associated with higher injury and harm levels as was the presence of traffic control devices.

We found that North Carolina work zone crashes tend to be more injurious than non-work zone crashes, and that truck-involved crashes are more injurious in work zones than non-truck involved crashes. So work zones represent a high-risk injury-inducing environment to motorists and strategies that can reduce work zone durations need to be explored.

Overall, these results provide valuable information for developing work zone strategies intended to improve vehicle occupants' safety. Clearly reducing these high-risk factors will be beneficial. Specifically, we need a careful examination and countermeasure development for 1) optimizing physical layout in work zones with special attention given to lane closings/detour opposite side, 2) work zones that are located on two lane undivided roadways, 3) posting speed limits in work zones and managing speed, 4) transition areas within work zones and 5) the use of traffic control devices.

## ACKNOWLEDGEMENT

Financial support was provided in part by the Southeastern Transportation Center. The study was conducted by the Carolina Transportation Program, University of North Carolina at Chapel Hill, with help from the UNC Highway Safety Research Center. We thank Dr. Herman Huang for his help with the data.

## REFERENCES

(1) Kite, S. SMART Work Zones: Technology for work zone management. North Carolina Department of Transportation (NCDOT), 2003.

(2) Federal Motor Carrier Safety Administration. 2000 Work Zone Traffic Crash Facts. USDOT, Washington D.C., March, 2002.

(3) Elvik, R. A meta-analysis of value of life estimates for occupational and transport safety, mimeo. 1995.

(4) Blincoe, L. J. The Economic Costs of Motor Vehicle Crashes, 1994. National Highway Traffic Safety Administration, Washington, D.C. 1996.

(5) de Blaeij, A., R. J.G.M. Florax, P. Rietveld, and E. Verhoef. The value of statistical life in road safety: a meta-analysis. *Accident Analysis and Prevention*, Vol. 916, 2003, pp. 1–14.

(6) Hargroves, B., and M. Martin. Vehicle Accidents in Highway Work Zone. Report FHWA-RD-80-063. US Department of Transportation, 1980.
(7) Stammer, R. Highway safety: at the crossroads. *In Proceedings of ASCE Conference*. San Antonio, TX, 1988.
(8) Wang, J., W. E. Hughes, F. M. Council, and J. F. Paniati. Investigation of Highway Work Zone Crashes: What We Know and What We Don't Know. *Transportation Research Record*, No. 1529, TRB, National Research Council, Washington, D.C., 1996.
(9) Juergens, W. R. Construction Zone, Detour and Temporary Connection Accidents. Business and Transportation Agency, California Division of Highways, 1972.
(10) Liste, F., R. Bernard, and B. Melvin. Evaluation of Timber Barricades and Precast Concrete Traffic Barriers for Use in Highway Construction Areas. Virginia Highway and Transportation Research Council, 1976.
(11) Graham, J., R. Paulsen, and J. Glennon. Accident and Speed Studies in Construction Zones. Report FHWA-RD-77-80. US Department of Transportation, 1977.
(12) Nemeth, Z. A., and D. J. Migletz. Accident Characteristics Before, During, and After Safety Upgrading Projects on Ohio's Rural Interstate System. *Transportation Research Record*, No. 672, TRB, National Research Council, Washington, D.C., 1978.
(13) Rouphail, N., M. Zhao, S. Yang, and J. Fazio. Comparative Study of Short-and Long-Term Urban Freeway Work Zones. *Transportation Research Record*, No. 1163, TRB, National Research Council, Washington, D.C., 1988.
(14) Hall, J. W., and V. M. Lorenz. Characteristics of Construction-Zone Accidents. *Transportation Research Record*, No. 1230, TRB, National Research Council, Washington, D.C., pp. 20-27, 1989.
(15) Garber, N. J., and T. H. Woo. Accident Characteristics at Construction and Maintenance Zones in Urban Areas. VTRC 90-R12. Virginia Transportation Research Council, Charlottesville, 1990.
(16) Hargroves, B. T. Vehicle Accidents in Highway Work Zones. *Journal of Transportation Engineering*, Vol. 107, No. TE5, 1981, pp. 525-539.
(17) Nemeth Z. A., and A. Rathi. Freeway Work Zone Accident Characteristics. *Transportation Quarterly*, Vol. 37, No. 1, 1983, pp. 145-159.
(18) Ha, T., and Z. Nemeth. Detailed Study of Accident Experience in Construction and Maintenance Zones. *Transportation Research Record*, No. 1509, TRB, National Research Council, Washington, D.C., 1995.
(19) Pigman, J. G., and K. R. Agent. Highway Accidents in Construction and Maintenance Work Zone. *Transportation Research Record*, No. 1270, TRB, National Research Council, Washington, D.C., 1990, pp. 12-21.
(20) Standing Committee on Highway Traffic Safety, American Association of State Highway and Transportation Officials. Summary Report on Work Zone Accidents. Washington, D.C., 1987.
(21) Khattak A., A. J. Khattak, and F. M. Council. Effects of work zone presence on injury and non-injury crashes. *Accident Analysis and Prevention*, Vol. 34, 2002, pp. 19-29.
(22) Garber, NJ, and M. Zhao. Distribution and characteristics of crashes at different work zone locations in Virginia. *Transportation Research Record*, No. 1794, TRB, National Research Council, Washington, D.C., 2002, pp. 19-25.
(23) Daniel, J. K. Dixon, and D. Jared. Analysis of fatal crashes in Georgia work zones. *Transportation Research Record*, No. 1715, TRB, National Research Council, Washington, D.C., 2000, pp. 18-23.
(24) Chambless, J., A. M. Ghadiali, J. K. Lindly, and J. McFadden. Multistate work-zone crash characteristics. *Institute of Transportation Engineers Journal*, Vol. 72, no. 5, 2002, pp.46-50.
(25) O'Donnell C., and D. Connor. Predicting the Severity of Motor Vehicle Accident Injuries Using Models of Ordered Multiple Choice. *Accident Analysis and Prevention*, Vol. 28, No. 6, 1996, pp. 739-753.

(26) Duncan C., A. J. Khattak, and F. M. Council. Using the ordered probit model to analyze factors that influence injury severity in truck-passenger car rear-end collisions on divided highways. *Transportation Research Record*, No. 1635, TRB, National Research Council, Washington, D.C., 1998, pp. 63-71.

(27) Hagle, T., and G. Mitchell. Goodness of Fit Measures for Probit and Logit. *American Journal of Political Science*, Vol. 36, 1992, pp. 762-784.

(28) Windmeijer, F. A. G. Goodness-of-fit measures in binary choice models. *Econometric Reviews*, Vol. 14, No. 1, 1995 pp. 101-116.

(29) Miller, T., and R. Spicer. Children's Safety Network Economics and Insurance Resource Center, PIRE, Calverton, MD, 2001.

(30) Kennedy, P. E. Estimation With Correctly Interpreted Dummy Variables in Semilogarithmic Equations. *The American Economic Review*, Vol. 71, No. 4, 1981, pp. 801.

## LIST OF TABLES AND FIGURES

**TABLE 1:** Fatalities and Non-Fatal Injuries from Motor Vehicle Collisions in Year 2000

| | **United States** | | | **North Carolina** | | |
|---|---|---|---|---|---|---|
| | Fatalities | Non-fatal injuries | Total | Fatalities | Non-fatal injuries | Total |
| **Total crashes** | 41,821 | 3,189,000 | 3,230,821 | 1,472 | 141,209 | 142,681 |
| **Row %** | 1.29% | 98.71% | 100% | 1.03% | 98.97% | 100% |
| **Work zones crashes** | 829 | 52,000 | 52,829 | 33 | 2,354 | 2,387 |
| **Row %** | 1.57% | 98.43% | 100% | 1.38% | 98.62% | 100% |
| **Work zone truck-involved crashes** | 264 | 15,600 | 15,864 | 10 | 394 | 404 |
| **Row %** | 1.66% | 98.34% | 100% | 2.48% | 97.52% | 100% |

*Source: FARS, 2002 and HSIS, 2000.*

**TABLE 2:** Descriptive Statistics for Year 2000 NC Work Zone Crashes Summarized at Accident Level

| | All collisions | | Truck-involved collisions | | Non truck-involved collisions | |
|---|---|---|---|---|---|---|
| **Variable** | Valid N | Mean or Proportion | Valid N | Mean or Proportion | Valid N | Mean or Proportion |
| Roadway contributing circumstances | | | | | | |
| Work zone | 3,366 | 21.3% | 656 | 24.8% | 2,656 | 20.5% |
| Other | 3,366 | 6.5% | 656 | 6.1% | 2,656 | 6.4% |
| None | 3,366 | 72.2% | 656 | 69.1% | 2,656 | 73.1% |
| Location of the crash | | | | | | |
| Before work area | 3,383 | 21.2% | 660 | 20.0% | 2,669 | 21.7% |
| Adjacent to actual work area | 3,383 | 44.9% | 660 | 45.6% | 2,669 | 44.6% |
| In work area approach taper | 3,383 | 33.9% | 660 | 34.4% | 2,669 | 33.7% |
| Work zone type | | | | | | |
| Construction | 3,383 | 81.6% | 660 | 82.0% | 2,669 | 81.5% |
| Maintenance | 3,383 | 9.2% | 660 | 8.6% | 2,669 | 9.3% |
| Utility | 3,383 | 5.5% | 660 | 3.9% | 2,669 | 5.9% |
| Intermittent/moving | 3,383 | 3.7% | 660 | 5.5% | 2,669 | 3.3% |
| Work zone activity (1 = On-going) | 3,383 | 51.1% | 660 | 61.8% | 2,669 | 48.6% |
| Work zone marked with sign/cones (1 = Yes) | 3,383 | 95.1% | 660 | 95.6% | 2,669 | 95.0% |
| Type of traffic control device | | | | | | |
| Stop/Yield/Warning sign | 3,383 | 23.3% | 660 | 18.3% | 2,669 | 24.6% |
| Stop/Yield/Warning flashing sign | 3,383 | 1.2% | 660 | 1.1% | 2,669 | 1.3% |
| Human control | 3,383 | 6.3% | 660 | 5.9% | 2,669 | 6.3% |
| Other | 3,383 | 11.6% | 660 | 11.4% | 2,669 | 11.6% |
| No control present | 3,383 | 57.7% | 660 | 63.3% | 2,669 | 56.2% |
| Dummy for missing cases | 3,383 | 6.7% | 660 | 7.3% | 2,669 | 6.5% |
| Roadway configuration | | | | | | |
| One-way, not divided | 3,363 | 3.7% | 658 | 3.3% | 2,653 | 3.8% |
| Two-way, not divided | 3,363 | 45.7% | 658 | 31.9% | 2,653 | 48.9% |
| Two-way, divided, unprotected | 3,363 | 17.0% | 658 | 16.7% | 2,653 | 17.3% |
| Two-way, divided, median barrier | 3,363 | 33.5% | 658 | 48.0% | 2,653 | 30.0% |
| Construction effect on the roadway[T] | | | | | | |
| Lane closed | | | 660 | 15.5% | | |
| Shoulder/Median closed | | | 660 | 17.0% | | |
| Roadway closed, detour opposing side | | | 660 | 2.6% | | |
| Lanes shift/become narrow | | | 660 | 1.8% | | |
| Other/Unknown | | | 660 | 21.5% | | |
| None | | | 660 | 41.7% | | |
| Type of work being done[T] | | | | | | |
| Repaving/Resurfacing | | | 660 | 8.8% | | |
| Shoulder/Median work | | | 660 | 14.5% | | |
| New roadway | | | 660 | 2.1% | | |
| Other | | | 660 | 6.1% | | |
| Unknown | | | 660 | 68.5% | | |
| Most harmful event (crash) | | | | | | |
| Ran off road | 3,367 | 2.3% | 656 | 2.1% | 2,657 | 2.1% |
| Rollover | 3,367 | 4.1% | 656 | 5.0% | 2,657 | 3.5% |
| Collision with pedestrian/bicyclist/animal | 3,367 | 2.5% | 656 | 1.5% | 2,657 | 2.5% |
| Collision with movable object | 3,367 | 2.6% | 656 | 3.2% | 2,657 | 2.3% |
| Collision with fixed object | 3,367 | 13.6% | 656 | 9.1% | 2,657 | 14.5% |
| Collision with 2+ motor vehicles (head on) | 3,367 | 1.1% | 656 | 0.9% | 2,657 | 1.1% |
| Collision with 2+ motor vehicles (angle) | 3,367 | 6.9% | 656 | 6.9% | 2,657 | 7.0% |
| Collision with 2+ motor vehicles (other) | 3,367 | 67.0% | 656 | 71.2% | 2,657 | 66.9% |
| Authorized speed limit (mph) | 3,351 | 50.08 | 653 | 52.51 | 2,646 | 49.56 |
| Number of vehicles involved in the crash | 3,383 | 1.97 | 660 | 2.07 | 2,669 | 1.96 |
| Number of persons involved in the crash | 3,383 | 3.01 | 660 | 2.78 | 2,669 | 3.07 |
| Weather (1 = Not Clear) | 3,383 | 28.5% | 660 | 26.5% | 2,669 | 29.0% |
| Ambient light (1 = Non Daylight) | 3,382 | 22.6% | 660 | 20.0% | 2,668 | 22.7% |

[T] Coded only for truck-involved collisions from North Carolina police reports.

**TABLE 3:** Coefficients for Injury Severity and Total Harm Models (Multi-Vehicle Collisions)

| Variable | Injury (Ordered probit) | | | Harm (Semi-log) | | |
|---|---|---|---|---|---|---|
| | All | Truck | Non truck | All | Truck | Non truck |
| Roadway contributing circumstances | | | | | | |
| Work zone | -0.080 | -0.181 | -0.049 | -0.061 | -0.167* | -0.036 |
| Other | -0.003 | 0.047 | 0.004 | 0.006 | 0.079 | -0.010 |
| None(B) | | | | | | |
| Location of the crash | | | | | | |
| Before work area | -0.010 | 0.243 | -0.078 | 0.030 | 0.282** | -0.039 |
| Adjacent to actual work area | 0.010 | 0.321** | -0.082 | 0.012 | 0.210** | -0.051 |
| In work area approach taper(B) | | | | | | |
| Work zone type | | | | | | |
| Construction | -0.072 | -0.155 | -0.062 | -0.049 | -0.069 | -0.053 |
| Maintenance(B) | | | | | | |
| Utility(B) | | | | | | |
| Intermittent/moving(B) | | | | | | |
| Work zone activity (1 = On-going) | -0.040 | -0.045 | -0.036 | 0.044 | 0.062 | 0.025 |
| Work zone marked with sign/cones (1 = Yes) | -0.028 | 0.151 | -0.101 | -0.022 | -0.206 | 0.039 |
| Type of traffic control device | | | | | | |
| Stop/Yield/Warning sign | 0.071 | -0.120 | 0.119* | 0.059 | -0.048 | 0.090** |
| Stop/Yield/Warning flashing sign | 0.453** | 0.922** | 0.440** | 0.468*** | 1.207*** | 0.312** |
| Human control | 0.005 | -0.124 | 0.050 | -0.002 | -0.036 | -0.008 |
| Other | 0.140* | 0.159 | 0.180** | 0.113** | 0.108 | 0.122** |
| No control present(B) | | | | | | |
| Dummy for missing cases | 0.058 | 0.174 | 0.025 | 0.014 | 0.185 | -0.027 |
| Roadway configuration | | | | | | |
| One-way, not divided | 0.192 | 0.305 | 0.168 | 0.108 | 0.211 | 0.076 |
| Two-way, not divided | 0.428*** | 0.510*** | 0.398*** | 0.266*** | 0.313*** | 0.247*** |
| Two-way, divided, unprotected | 0.137* | 0.362** | 0.085 | 0.104** | 0.260** | 0.072 |
| Two-way, divided, median barrier(B) | | | | | | |
| Construction effect on the roadway(T) | | | | | | |
| Lane closed | | -0.166 | | | -0.096 | |
| Shoulder/Median closed | | 0.278 | | | 0.230 | |
| Roadway closed, detour opposing side | | 1.011*** | | | 0.889*** | |
| Lanes shift/become narrow | | -0.689 | | | -0.381 | |
| Other/Unknown | | 0.122 | | | 0.136 | |
| None(B) | | | | | | |
| Type of work being done (T) | | | | | | |
| Repaving/Resurfacing | | -0.059 | | | 0.033 | |
| Shoulder/Median work | | 0.046 | | | 0.069 | |
| New roadway | | -0.534 | | | -0.116 | |
| Other | | -0.048 | | | 0.085 | |
| Unknown(B) | | | | | | |
| Most harmful event (crash) | | | | | | |
| Ran off road | -0.124 | -0.184 | -0.093 | -0.065 | -0.020 | -0.069 |
| Rollover | 1.133*** | 0.748 | 1.481*** | 0.740*** | 0.516 | 1.037** |
| Collision with pedestrian/bicyclist/animal | 1.674*** | 1.582*** | 1.698*** | 1.662*** | 1.431*** | 1.880*** |
| Collision with movable object | -0.565** | -0.623 | -0.576* | -0.468*** | -0.320 | -0.387* |
| Collision with fixed object | 0.159 | -0.138 | 0.271 | 0.007 | -0.191 | 0.059 |
| Collision with 2+ motor vehicles (head on) | 1.445*** | 1.566*** | 1.456*** | 1.639*** | 1.742*** | 1.618*** |
| Collision with 2+ motor vehicles (angle) | 0.175** | 0.305 | 0.137 | 0.204*** | 0.307** | 0.163** |
| Collision with 2+ motor vehicles (other)(B) | | | | | | |
| Authorized speed limit (mph) | 0.016*** | 0.022*** | 0.015*** | 0.013*** | 0.014*** | 0.011*** |
| Number of vehicles involved in the crash | 0.366*** | 0.314*** | 0.360*** | 0.364*** | 0.181** | 0.323*** |
| Number of persons involved in the crash | 0.029*** | 0.093** | 0.024** | 0.141*** | 0.291*** | 0.146*** |
| Weather (1 = Not Clear) | -0.104* | -0.050 | -0.149** | -0.083** | 0.012 | -0.111*** |
| Ambient light (1 = Non Daylight) | 0.190*** | 0.128 | 0.210*** | 0.149*** | 0.124 | 0.163*** |
| Constant | | | | 7.768*** | 7.379*** | 7.973*** |
| Threshold 1 | 2.172*** | 2.872*** | 2.007*** | | | |
| Threshold 2 | 3.275*** | 3.770*** | 3.181*** | | | |
| Threshold 3 | 4.016*** | 4.383*** | 4.020*** | | | |

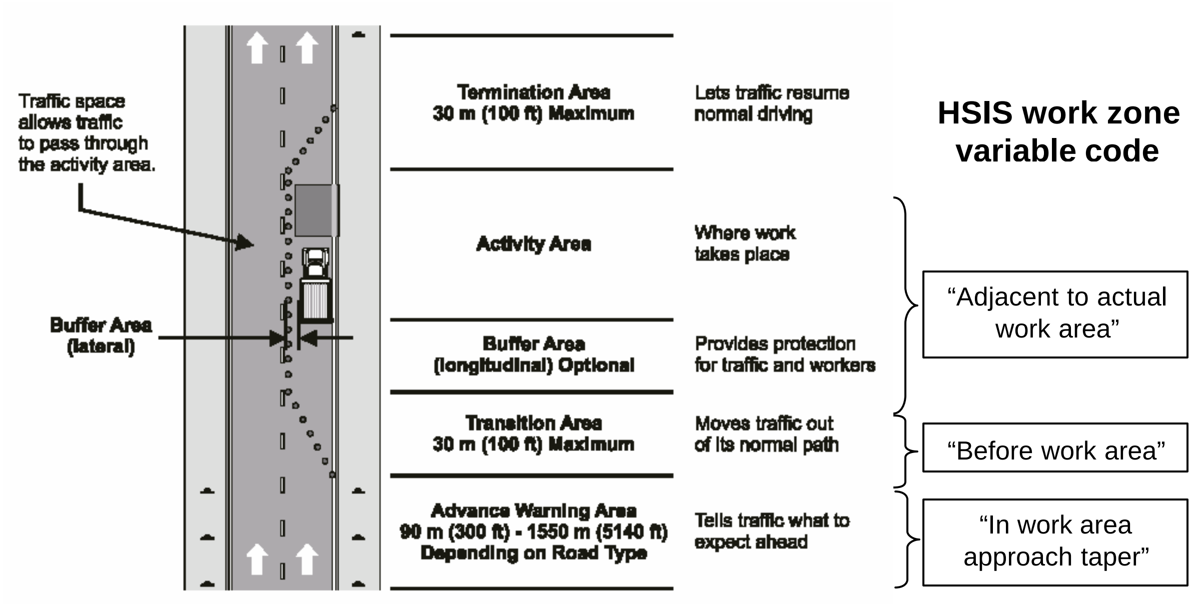
| | | | | | | |
|---|---|---|---|---|---|---|
| Threshold 4 | 4.690*** | 5.312*** | 4.581*** | | | |
| N | 2,649 | 557 | 2,074 | 2,649 | 557 | 2,074 |
| Log likelihood intercept only | -2,723.88 | -572.26 | -2,102.64 | | | |
| Log likelihood at convergence | -2,548.70 | -512.51 | -1,971.07 | | | |
| P> $\chi^2$: LR test ($\chi^2$) / P> $\chi^2$: F test ($\chi^2$) | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| Adjusted-$R^2$ | | | | 0.293 | 0.343 | 0.300 |
| McKelvey & Zavoina pseudo-$R^2$ | 0.150 | 0.242 | 0.146 | | | |

***, **, and * denote coefficient significantly different from zero at the 1%, 5%, and 10% level of significance (two-tail test), respectively.
(B) Base category for coefficient comparison.
(T) Coded only for truck-involved collisions (police officer accident reports).
Coefficients for dummy categorical variables in the Harm models are adjusted for interpretation according to Kennedy (*30*) for a semi-log specification.

**TABLE 4:** Marginal Effects for Truck-Involved Injury Model (Multi-Vehicle Collisions)

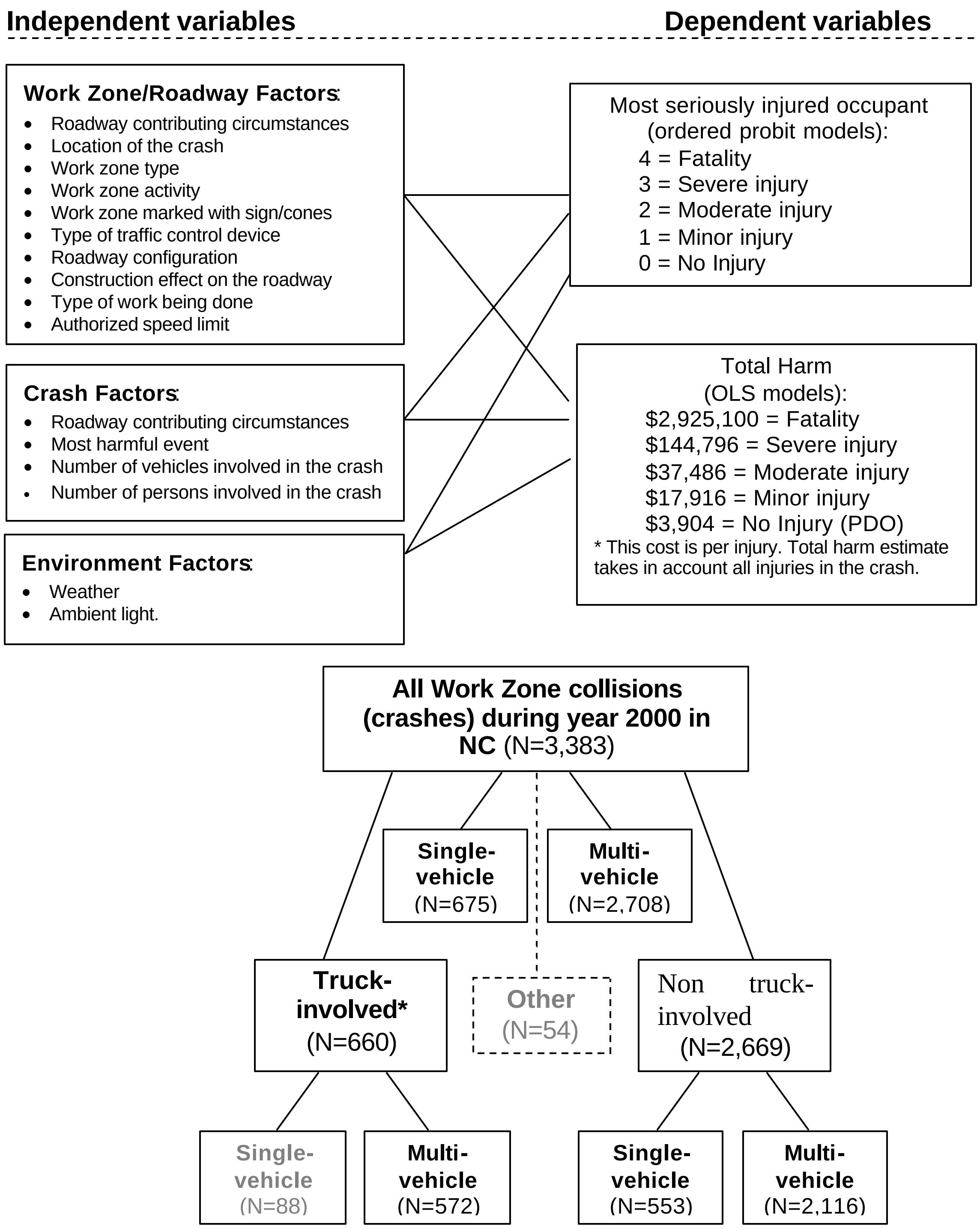
| Variable | Injury (Ordered probit) | | | | |
|---|---|---|---|---|---|
| | No injury | Minor | Moderate | Severe | Fatality |
| Roadway contributing circumstances | | | | | |
| Work zone | 0.067 | -0.035 | -0.020 | -0.011 | -0.001 |
| Other | -0.018 | 0.009 | 0.006 | 0.003 | 0.000 |
| None[B] | | | | | |
| Location of the crash | | | | | |
| Before work area | -0.088 | 0.048 | 0.026 | 0.013 | 0.002 |
| Adjacent to actual work area | **-0.119** | **0.062** | **0.035** | **0.019** | **0.002** |
| In work area approach taper[B] | | | | | |
| Work zone type | | | | | |
| Construction | 0.059 | -0.029 | -0.018 | -0.011 | -0.001 |
| Maintenance[B] | | | | | |
| Utility[B] | | | | | |
| Intermittent/moving[B] | | | | | |
| Work zone activity (1 = On going) | 0.017 | -0.009 | -0.005 | -0.003 | -0.000 |
| Work zone marked with sign/cones (1 = Yes) | -0.055 | 0.030 | 0.016 | 0.008 | 0.001 |
| Type of traffic control device | | | | | |
| Stop/Yield/Warning sign | **0.044** | **-0.024** | **-0.013** | **-0.007** | **-0.001** |
| Stop/Yield/Warning flashing sign | -0.353 | 0.088 | 0.124 | 0.114 | 0.028 |
| Human control | 0.045 | -0.025 | -0.013 | -0.007 | -0.001 |
| Other | -0.061 | 0.029 | 0.019 | 0.011 | 0.002 |
| No control present[B] | | | | | |
| Dummy for missing cases | -0.067 | 0.032 | 0.021 | 0.012 | 0.002 |
| Roadway configuration | | | | | |
| One-way, not divided | -0.111 | 0.060 | 0.032 | 0.017 | 0.002 |
| Two-way, not divided | **-0.191** | **0.094** | **0.058** | **0.034** | **0.005** |
| Two-way, divided, unprotected | **-0.133** | **0.070** | **0.039** | **0.021** | **0.003** |
| Two-way, divided, median barrier[B] | | | | | |
| Construction effect on the roadway[T] | | | | | |
| Lane closed | 0.059 | -0.033 | -0.016 | -0.008 | -0.001 |
| Shoulder/Median closed | -0.107 | 0.050 | 0.034 | 0.020 | 0.003 |
| Roadway closed, detour opposing side | **-0.385** | **0.095** | **0.134** | **0.125** | **0.032** |
| Lanes shift/become narrow | 0.207 | -0.133 | -0.050 | -0.021 | -0.002 |
| Other/Unknown | -0.046 | 0.023 | 0.014 | 0.008 | 0.001 |
| None[B] | | | | | |
| Type of work being done[T] | | | | | |
| Repaving/Resurfacing | 0.022 | -0.011 | -0.007 | -0.004 | 0.000 |
| Shoulder/Median work | -0.018 | 0.009 | 0.005 | 0.003 | 0.000 |
| New roadway | 0.177 | -0.106 | -0.047 | -0.022 | -0.002 |
| Other | 0.018 | -0.009 | -0.005 | -0.003 | 0.000 |
| Unknown[B] | | | | | |
| Most harmful event (crash) | | | | | |
| Ran off road | 0.065 | -0.037 | -0.018 | -0.009 | -0.001 |
| Rollover | -0.292 | 0.095 | 0.100 | 0.080 | 0.016 |
| Collision with pedestrian/bicyclist/animal | **-0.535** | **0.017** | **0.170** | **0.243** | **0.104** |
| Collision with movable object | 0.193 | -0.122 | -0.048 | -0.021 | -0.002 |
| Collision with fixed object | 0.050 | -0.028 | -0.014 | -0.007 | -0.001 |
| Collision with 2+ motor vehicles (head on) | **-0.532** | **0.020** | **0.170** | **0.240** | **0.101** |
| Collision with 2+ motor vehicles (angle) | -0.118 | 0.054 | 0.037 | 0.023 | 0.003 |
| Collision with 2+ motor vehicles (other)[B] | | | | | |
| Authorized speed limit (mph) | **-0.008** | **0.004** | **0.003** | **0.001** | **0.000** |
| Number of vehicles involved in the crash | **-0.118** | **0.060** | **0.035** | **0.020** | **0.003** |
| Number of persons involved in the crash | **-0.035** | **0.018** | **0.011** | **0.006** | **0.001** |
| Weather (1 = Not Clear) | 0.019 | -0.010 | -0.006 | -0.003 | 0.000 |
| Ambient light (1 = Non Daylight) | -0.049 | 0.024 | 0.015 | 0.009 | 0.001 |

Note: For dummy variables the marginal effect corresponds to the discrete change from 0 to 1.
(B) Base category for coefficient comparison. (T) Coded only for truck-involved collisions (police officer accident reports).



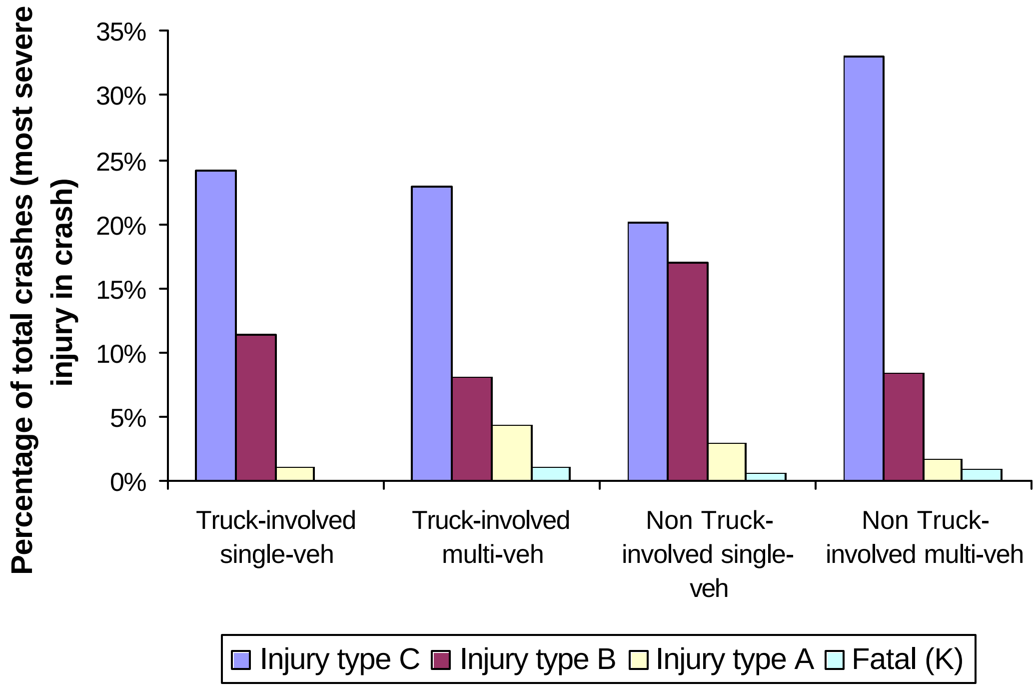

*Source: Manual on Uniform Traffic Control Devices* (*MUTCD*) *2000, U.S. Department of Transportation, Federal Highway Administration. Adapted and modified from* (*22*).

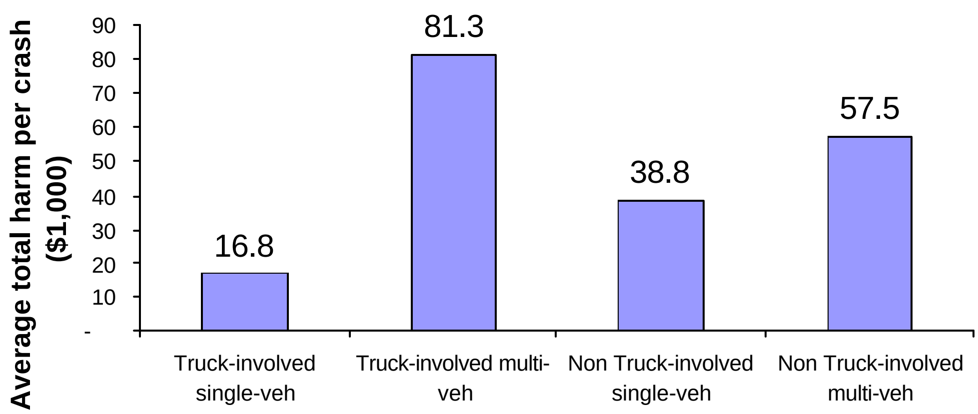
**FIGURE 1:** Location of the crash in a work zone.




Notes: Data come from HSIS 2000, which contained four new work zone-related variables 1) location of the crash, 2) work zone type, 3) work zone activity, and 4) work zone marked with sign/cones.
Non truck-involved collisions must involve a Car, SUV, or Van.
* Additional work zone-related data were coded from police officer accident reports to create 1) construction effect on the roadway and 2) type of work being done.

**FIGURE 2:** Conceptual and data structure.




Cost: $2,925,100 = Fatality; $144,796 = Severe injury; $37,486 = Moderate injury; $17,916 = Minor injury; $3,904 = No Injury (PDO)

**FIGURE 3:** Most severe injury and total crash harm distribution by type of collision.

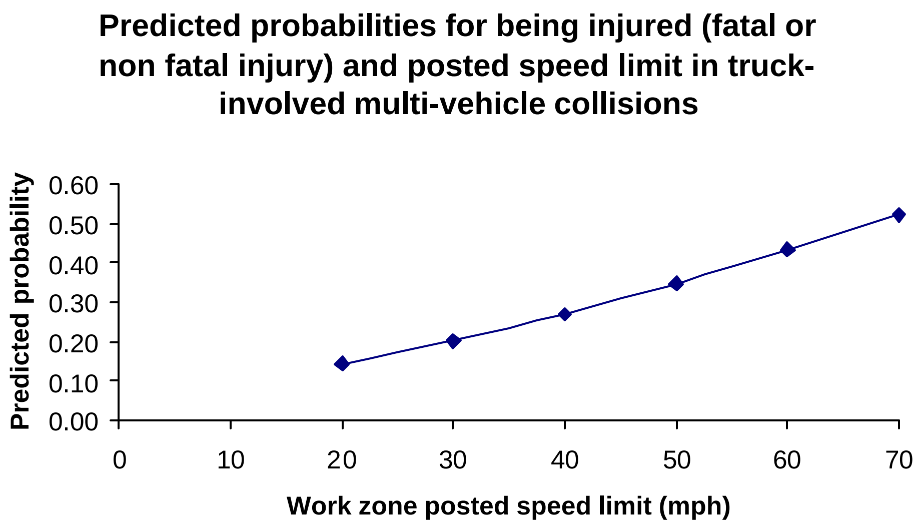



**FIGURE 4:** Predicted probabilities of injury severity for different posted speed limit in work zones.

**INJURY SEVERITY AND TOTAL HARM IN TRUCK-INVOLVED WORK ZONE CRASHES**

PART 2: VEHICLE LEVEL ANALYSIS

**Asad J. Khattak & Felipe Targa**
University of North Carolina at Chapel Hill

**Abstract—**This paper builds on the previous work by empirically examining truck-involved and non truck-involved collisions at the vehicle level. The paper helps understand which work zone attributes are statistically associated with the most seriously injured occupant and total harm in the vehicle, but it also allows us to explore and control for the effects of driver factors in multi-vehicle collisions. Controlling for these factors is important since driver actions and decisions contribute substantially to crash occurrence and injury severity. Using North Carolina work zone data for vehicle-collisions during the year 2000 (N=6,672), we estimated ordered probit models for the most seriously injured occupant in the vehicle and linear regression models for the "total harm" in the vehicle. The models account for the correlation among vehicles involved in the same collision. Modeling results suggest that the relationships found in the previous paper also hold at the vehicle level, e.g., the most injurious work zone vehicle-level crashes were those involving a truck when the roadway was closed, requiring a detour on the opposite side. Other crash, environmental, vehicle and driver-related effects are discussed in the paper. Additionally, vehicle-segmented models provide a deeper understanding of work zone related risk factors that can influence injury severity differently among various types of vehicles. The results are valuable for understanding work zone risks and potential strategies intended to improve vehicle occupants' safety by reducing their injury severity.

**Keywords—**safety, work zones, collisions, injury, large trucks

## INTRODUCTION

To discern the role of vehicles and drivers in work zone crashes, we analyzed the work zone crash data at the vehicle-level. The role of driver attributes such as age and gender and behaviors such as excess speed and violation of traffic control as well as vehicle factors such as single-unit vs. tractor trailer truck is likely to be very important in work zone crashes. However, analyzing data at the crash level, especially when multiple vehicles are involved makes it difficult to capture the complexity and effects of various vehicles types and individual drivers of those vehicles towards injury and costs. Specifically, the problem with studying multi-vehicle collisions at the crash level is that injury severity or harm for those involved in the same collision will be correlated, given that the unit of analysis is the vehicle. This violates an important assumption of parametric models, which are needed to understand the effects of work zone variables while controlling for all other factors. For this reason, we estimate the robust standard errors for all models reported in this paper. We use the Huber/White/Sandwich estimator of variance. Thus the paper builds on the previous research by allowing us to explore and control for the effects of driver and vehicle factors on work zone injury risk and costs in more detail, using vehicle-segmented models. (The literature review of the previous paper is relevant, though it is not reproduced to avoid redundancy.)

## RESULTS

Figure 1 shows a typical work zone and it is reproduced for completeness. Figure 2(a) shows the conceptual structure which is similar to the one proposed previously, but it also identifies the role of driver and vehicle factors.

### Descriptive Statistics

During 2000, 6,672 vehicles were involved in North Carolina work zone collisions; 675 (10.1%) were single-vehicle collisions and 5,997 (89.9%) involvements were multi-vehicle (see Figure 2(b)). Out of the 6,672 vehicle-involvements, 1,369 (20.5%) involved at least one truck—88 were single-vehicle collisions and 1,281 were multi-vehicle involvements; and 5,229 (78.4%) did not involve any truck but involved a car, a Sports Utility Vehicle (SUV), or a van—553 were single-vehicle collisions and 4,676 multi-vehicle involvements. Though most vehicle-involvements resulted in no injuries (PDO vehicle-involvements ranged from 77.8% in truck-involved multi-vehicle crashes to 59.4% in non truck-involved single-vehicle crashes), multi-vehicle truck-involved crashes appear to be the second most injurious at the vehicle-level after non truck-involved single-vehicle crashes (Figure 3). Specifically, 0.48% vehicle-crashes turned out to be fatal and 2.10% severe (injury type A) in truck-involved multi-vehicle collisions. Moreover, if we account for all injuries in the vehicle by assigning a dollar value as shown above, truck-involved multi-vehicle collisions in work zones are also the second most costly, on average, costing $37,400 per vehicle involved in the crash (Figure 3). However, on a per-vehicle basis, the costliest crashes are single-vehicle non-truck collisions ($38,800).

A brief description of the coding scheme for the variables and descriptive statistics for various types of work zone collisions is presented in Table 2. Compared with non truck-involved collisions, truck-involved collisions occurred more frequently when work zone activity was on-going and there was no traffic control device present. Moreover, truck-involved collisions occurred more frequently on two-way divided and protected roadways, as opposed to two-way undivided configurations. These relationships are investigated more rigorously by estimating regression models.

Although truck-involved vehicle-collisions were 675, truck-involved single-vehicle vehicle-collisions were only 88, precluding any modeling of single-vehicle collisions. Given that single-vehicle truck-involved vehicle-collisions are the least costly ($16,800 per vehicle-crash; see Figure 3), they seem to be the least problematic (from a safety perspective) among work zone vehicle-crashes.

### Modeling Results

Table 3 reports estimates from six models for multi-vehicle collisions at the vehicle-level: Three ordered probit models for injury severity and three OLS log-transformed models for total harm in the vehicle. Each three-set of models is estimated for the pooled data, truck-involved collisions and non truck-involved collisions. This allows the non truck-involved collision models to serve as the comparison group. Model interpretation for truck-involved collisions is also based on marginal effects and predicted probabilities for the ordered probit model (Table 4). Two additional tables show segmented models by vehicle type for both injury severity (Table 5) and total harm in the vehicle (Table 6). The model segmentation by vehicle type is based in the following categories:

- Trucks in truck-involved multi-vehicle work zone collisions;

- Cars, SUVs, or Vans in truck-involved multi-vehicle work zone collisions;
- Cars in non truck-involved multi-vehicle work zone collisions;
- SUVs or Vans in non truck-involved multi-vehicle work zone collisions.

We focus on discussing model results for statistically significant coefficients at the customary 99%, 95% and 90% confidence levels. Because the unit of analysis is vehicle-collision, injury severity and total harm may be independent across collisions, but this may not hold within collisions. For example, injury severity or total harm in a vehicle is expected to be correlated with the injury severity or total harm of another vehicle involved in the same crash. For this reason, robust standard errors are estimated for all models—i.e., the Huber-White-Sandwich estimator of variance is used.

All models are statistically significant as indicated by F-tests for the OLS models and Likelihood-Ratio tests for the probit models significant at the 99% confidence level. Adjusted $R^2$s range from 0.242 (all work zone involvements, pooled data) to 0.407 (truck-involved multi-vehicle collisions, segmented by trucks only) for the OLS log-transformed models; and McKelvey & Zavoina pseudo-$R^2$s range from 0.145 (non truck-involved collisions) to 0.870 (truck-involved multi-vehicle collisions segmented by trucks only) for the ordered probit models. These indicate reasonable goodness-of-fit.

*Pooled Models*

Among the work zone/roadway-related variables coded from the DMV crash report forms (only for truck-involved multi-vehicle collisions), *Roadway closed* and *Shoulder/Median closed* (construction effect on the roadway) are significantly associated with higher injury (Table 3). In addition to these two variables, *Lanes shift/become narrow* is also associated with higher cost in the total harm model. Roadway closed increases the cost per vehicle-involved by a factor of 80.3% ($e^{0.590}$-1) compared with other crashes. The marginal effects for the injury severity model (Table 4) suggest that if a crash occurs in a work zone where the roadway is closed, requiring a detour on the opposite side, then the chances of injuries in the vehicle increase by 32.0%, i.e., the chances of minor injuries are 17.3% higher, moderate injuries are 9.3% higher, severe injuries are 4.6% higher, and there is a small increase of 0.7% in the chances of fatalities. Similarly, Shoulder/Median closed increases the cost per vehicle-involved by a factor of 19.8% ($e^{0.181}$-1), while a shift in lanes decreases the cost by a factor of 31.6% ($e^{-0.380}$-1) compared to the "base." These results are consistent with the crash-level analysis.

For truck-involved collisions, the crash location in the work zone is a significant factor in injury severity and total harm at the vehicle level. Truck-involved crashes that occurred before the work area or in the transition area, where traffic moves out of its normal path, or adjacent to the work area are associated with higher total harm in the vehicle than crashes in the approaching taper or advance warning area (see Figure 1). The harm model suggests that the cost per vehicle-involved in the "before work area" is 20.3% ($e^{0.185}$-1) more while those in the activity/work area cost 12.8% ($e^{0.121}$-1) more than vehicles involved in crashes that occurred in the advance warning area. Table 4 suggests that the chances of injuries in the vehicle are higher by 7.9%, (i.e., the chances of minor injuries are 5.5% higher, moderate injuries are 1.8% higher, severe injuries by about 0.6% higher and fatalities by 0.1%) if the truck-involved crash occurs adjacent to the actual activity/work area compared with a crash in the advance warning area. Surprisingly, this association with injury and total harm is negative for non truck-involved collisions. This warrants greater attention to the activity/work area as a risk factor in truck-car collisions.

Work zones located on certain roadways are associated with higher injury severity and total harm in the vehicle. Specifically, work zones located on two-way not divided roadways are associated with 44.5% ($e^{0.368}$-1) higher costs per vehicle compared with work zones on two-way divided and protected roadways. This association is stronger for truck-involved than non truck-involved collisions, as indicated by the parameter estimates for the injury severity model. The chances of injuries in the vehicle are 16.3% higher, if a truck-involved crash occurs in a work zone located on a two-way undivided roadway. For truck-involved crashes only, a two-way divided but unprotected configuration was also more injurious than a two-way divided and protected configuration, increasing the chances of injuries in the vehicle by 9.8%. Thus to reduce vehicle injury severity and harm in truck-involved collisions, work zones located on two-way undivided and two-way divided but unprotected roadways need particular attention.

Higher posted speed limit in the work zone is associated with higher injury severity and total harm in the vehicle, given a crash. In terms of harm, every 10 mph increase in the authorized speed limit increases the total cost of the vehicle-crash by 12.3% ($e^{0.123}$-1). For the same 10 miles per hour increase in the authorized speed limit, the chances of injuries in a vehicle increase by 7.0%. This increase is somewhat uniform across truck-involved and non truck-involved crashes. Therefore, careful consideration needs to be given to setting up the speed limits in work zones.

The presence of stop/yield/warning flashing signs (traffic control devices) is associated with higher injury severity and harm in the vehicle, particularly for truck-involved collisions. However, this effect may not be causal, i.e., these devices typically prevent even more severe crashes from occurring. The presence of these devices is often associated with larger work zones and higher levels of hazardous activity at work zones. Nevertheless, the coefficient estimates suggest that crashes occurring in the presence of traffic control devices are more injurious and harmful in the vehicle than no control device present, so the use of these control devices in work zones needs to be further investigated.

In addition to work zone attributes, several other variables were included in the model specification. Crash-related factors such as collisions involving rollover were associated with higher injury severity and total harm in the vehicle, as expected. Head-on and angle collisions are also more severe than other type of collisions. Crash exposure measures, such as number of people in the vehicle is significantly associated with higher injury severity and total harm in the vehicle. Environmental factors, such as ambient light is significant but only for non truck-involved collisions.

Among driver contributing factors, excess speed is significantly associated with higher injury severity and total harm for truck-involved collisions. However, this effect is the opposite (and counter-intuitive) for non truck-involved collisions. The estimated speed at the impact is also associated with higher injury severity and total harm in the vehicle, as expected. Driver's race and gender are associated with higher injury severity in the vehicle. Particularly, being a non-white or female driver increases the chances of injuries in a vehicle by 5.5% and 7.0%, respectively. Driver's age is also associated with higher injury severity and total harm in the vehicle. For example, every 10 years increase the cost per vehicle by 4.5% ($e^{0.045}$-1). For the same 10 years increase in the driver's age, the chances of injuries in a vehicle increase by 2.0%. Interestingly, if (at least one of) the driver condition is (coded as) not normal, then the non-truck occupants' injury costs are high.

Collisions with vehicles parked or stopped in the work zone are more injurious and costly than other multi-vehicle collisions. Finally, and perhaps more importantly, passenger vehicles are associated with higher injury severity and total harm than trucks in truck-involved crashes. This is clearly due to the mass differential of the two vehicle types and it confirms for work zones what we know is true in general, i.e., that passenger vehicle occupants are more likely to be injured in collisions with large trucks. The harm model suggests that the cost associated to single unit trucks and large trucks (Truck/Trailer/Tractor) in multi-vehicle truck-involved collisions is 32.9% ($e^{0.400}$-1) and 36.5% ($e^{0.454}$-1) lower than the cost associated with passenger cars. The marginal effects for the injury severity model in Table 4 suggest that the chances of injuries in single unit trucks or tractor-trailers are lowered by 20.3%, and 22.8%, respectively, compared with passenger vehicles. Moreover, for non truck-involved crashes, passenger cars are associated with higher injury severity and total harm in the vehicle than SUVs, Pickups, or Vans. Clearly larger mass of vehicles is a factor in reducing injuries and harm in work zone crashes. These results motivate a further analysis of injury severity and total harm for truck-involved and non truck-involved collisions segmenting by vehicle type. The vehicle-segmented models presented in the next section provide further evidence related to the differential effects of work zone risk factors as they relate to various types of vehicles (given a crash).

*Segmented Models*

A set of segmented models were estimated. Some of the parameter estimates were not estimable due to lack of data in the category. The results shed light on the differential effects of various factors across trucks and passenger vehicles. Among the work zone/roadway-related variables coded from the DMV crash report forms, the significant association between injury and harm, and *Roadway closed, Shoulder/Median closed,* and *Lanes shift/become narrow* variables is attributable to passenger vehicles (passenger cars, SUVs, Pickups, and Vans) and not to trucks in truck-involved collisions. Therefore, passenger vehicle occupants are more seriously injured than truck occupants when they are involved in crashes with these work zone/roadway-related characteristics.

The crash location in the work zone is a significant factor in injury severity and total harm in the vehicle. Previously, we found that the truck-involved multi-vehicle crashes which occurred before or adjacent the work area were associated with higher total harm in the vehicle. When segmented by vehicle type, the association is significant for trucks adjacent the work area and for passenger vehicles before the work area. In particular, the harm model (in Table 6) suggests that passenger vehicles colliding with large trucks before the work area cost 33.0% ($e^{0.285}$-1) more than crashes in the advance warning area. Interestingly, for non truck-involved crashes (segmented by passenger vehicles) the coefficient of before the work area is negative. This result suggests again that passenger vehicle occupants are being seriously injured when they collide with trucks before the work area or in the transition area.

One work zone-related variable that was not statistically significant in the pooled models is significant for truck-involved crashes segmented by trucks. Work zone activity at the time of the crash is significantly associated with lower injury severity and total harm to truck occupants. The harm model (in Table 6) suggests that crashes that occurred during work zone activity cost 14.7% ($e^{-0.159}$-1) less than crashes occurring when there was no work activity at the time of the crash. Perhaps this captures the more cautious driving behavior when drivers observe work zone activity.

Work zones located on certain roadways are also associated with higher injury severity and total harm in the vehicles for all segmented models. The harm to passenger vehicle occupants is higher than large truck occupants when the crash occurred in work zones located on two-way undivided and two-way divided but unprotected roadways.

Higher posted speed limit in the work zone is associated with higher injury severity and total harm in all segmented models, as expected. For truck-involved crashes, the presence of stop/yield/warning flashing signs is associated with higher injury severity in truck and passenger vehicle occupants and higher harm only in truck occupants. Crash-related factors such as collisions involving rollover, head-on and angle collisions were also associated with higher injury severity and total harm for truck-involved crashes. Head-on crashes are more injurious to passenger vehicle occupants, while angle collisions seem more injurious to truck occupants. Crash exposure measures, such as number of people in the vehicle is significantly associated with higher injury severity and total harm in occupants for all vehicle types. Environmental factors, such as whether (not clear) is significantly associated with higher injury severity and total harm in truck occupants. Among driver factors, driver's physical condition (not normal), race (non white), and gender (female) are associated with higher injury severity for truck occupants, while driver's age with higher injury severity and total harm for passenger vehicle occupants.

## LIMITATIONS

If non-reporting of crashes varies systematically for work zone collisions, or any other type of collision such as truck-involved or non truck-involved, then it can bias the results. One could hypothesize both higher reporting in work zones (due to potentially higher police presence and less room to move damaged vehicles to the shoulders), and lower reporting (if vehicles in non-injury crashes are encouraged to leave the scene in order to decrease queues). Despite this potential bias, the accident data comes from a relatively clean, federally maintained Highway Safety Information System, there was very little missing data and the descriptive statistics and model results are quite reasonable.

## CONCLUSIONS

The purpose of this portion of the study was to investigate large truck-involved crashes in work zones at the vehicle level. Using a unique dataset that contains new work zone-related variables, not available before in police crash reports and coding additional variables from reports, we present empirical evidence regarding injury risks and harm in vehicles associated with work zone crashes. Given work zone crashes, we found that multi-vehicle truck-involved collisions were the most injurious and harmful, as expected.

Rigorously modeling injury and harm provided new insights into the role of work zones, while controlling for driver, vehicle, roadway and environmental factors. Confirming our earlier findings, truck-involved multi-vehicle crashes were most injurious and harmful in vehicles when the roadway was closed, requiring a detour on the opposite side. Furthermore, work zones located on a two-way undivided roadway substantially increased the chances of injuries in the vehicle compared with other roadway configurations. Within the work zone, truck-involved crashes occurring immediately before the actual work area or adjacent to the activity/work area were more injurious and harmful, compared to those occurring in the advance warning area. Higher speed limits were associated with higher injury and harm levels as was the presence of traffic control devices.

We found that truck-involved crashes in North Carolina are more injurious in work zones than non-truck involved crashes. So work zones represent a high-risk injury-inducing environment to motorists and strategies that can reduce work zone durations need to be explored.

Overall, these results provide valuable information for developing work zone strategies intended to improve vehicle occupants' safety. Clearly reducing these high-risk factors will be beneficial. Specifically, we need a careful examination and countermeasure development for 1) optimizing physical layout in work zones with special attention given to lane closings/detour opposite side, 2) work zones that are located on two lane undivided roadways where truck traffic is allowed, 3) posting speed limits in work zones and managing speed, 4) transition areas within work zones and 5) the use of traffic control devices in work zones.

## ACKNOWLEDGEMENT

Financial support was provided in part by the Southeastern Transportation Center. The study was conducted by the Carolina Transportation Program, University of North Carolina at Chapel Hill, with help from the UNC Highway Safety Research Center. We thank Dr. Herman Huang for his help with the data.

## LIST OF TABLES AND FIGURES

**TABLE 5:** Descriptive Statistics for Year 2000 NC Work Zone Crashes Summarized at Vehicle Level

| | **All collisions** | | **Truck-involved collisions** | | **Non truck-involved collisions** | |
|---|---|---|---|---|---|---|
| **Variable** | Valid N | Mean or Proportion | Valid N | Mean or Proportion | Valid N | Mean or Proportion |
| Roadway contributing circumstances | | | | | | |
| Work zone | 6,642 | 22.3% | 1,361 | 25.5% | 5,207 | 21.5% |
| Other | 6,642 | 5.4% | 1,361 | 5.3% | 5,207 | 5.3% |
| None(B) | 6,642 | 72.3% | 1,361 | 69.2% | 5,207 | 73.2% |
| Location of the crash | | | | | | |
| Before work area | 6,672 | 21.5% | 1,369 | 20.7% | 5,229 | 21.9% |
| Adjacent to actual work area | 6,672 | 45.5% | 1,369 | 46.3% | 5,229 | 45.2% |
| In work area approach taper(B) | 6,672 | 33.0% | 1,369 | 32.9% | 5,229 | 32.9% |
| Work zone type | | | | | | |
| Construction | 6,672 | 81.6% | 1,369 | 82.8% | 5,229 | 81.3% |
| Maintenance(B) | 6,672 | 9.3% | 1,369 | 8.8% | 5,229 | 9.4% |
| Utility(B) | 6,672 | 5.4% | 1,369 | 3.1% | 5,229 | 6.0% |
| Intermittent/moving(B) | 6,672 | 3.7% | 1,369 | 5.3% | 5,229 | 3.4% |
| Work zone activity (1 = On-going) | 6,672 | 46.3% | 1,369 | 37.3% | 5,229 | 48.6% |
| Work zone marked with sign/cones (1 = Yes) | 6,672 | 4.6% | 1,369 | 4.3% | 5,229 | 4.6% |
| Type of traffic control device | | | | | | |
| Stop/Yield/Warning sign | 6,672 | 24.5% | 1,369 | 19.1% | 5,229 | 26.1% |
| Stop/Yield/Warning flashing sign | 6,672 | 1.3% | 1,369 | 1.3% | 5,229 | 1.3% |
| Human control | 6,672 | 6.7% | 1,369 | 5.9% | 5,229 | 6.9% |
| Other | 6,672 | 11.6% | 1,369 | 11.1% | 5,229 | 11.7% |
| No control present(B) | 6,672 | 55.9% | 1,369 | 62.5% | 5,229 | 54.1% |
| Dummy for missing cases | 6,672 | 6.5% | 1,369 | 7.3% | 5,229 | 6.1% |
| Roadway configuration | | | | | | |
| One-way, not divided | 6,645 | 3.5% | 1,365 | 3.1% | 5,209 | 3.6% |
| Two-way, not divided | 6,645 | 45.4% | 1,365 | 30.8% | 5,209 | 49.0% |
| Two-way, divided, unprotected | 6,645 | 17.0% | 1,365 | 16.6% | 5,209 | 17.2% |
| Two-way, divided, median barrier(B) | 6,645 | 34.1% | 1,365 | 49.6% | 5,209 | 30.2% |
| Construction effect on the roadway(T) | | | | | | |
| Lane closed | | | 1,369 | 15.7% | | |
| Shoulder/Median closed | | | 1,369 | 16.7% | | |
| Roadway closed, detour opposing side | | | 1,369 | 2.3% | | |
| Lanes shift/become narrow | | | 1,369 | 1.6% | | |
| Other/Unknown | | | 1,369 | 21.8% | | |
| None(B) | | | 1,369 | 41.9% | | |
| Type of work being done(T) | | | | | | |
| Repaving/Resurfacing | | | 1,369 | 8.8% | | |
| Shoulder/Median work | | | 1,369 | 13.5% | | |
| New roadway | | | 1,369 | 2.3% | | |
| Other | | | 1,369 | 6.1% | | |
| Unknown(B) | | | 1,369 | 69.3% | | |
| Most harmful event (vehicle) | | | | | | |
| Ran off road/other non collision | 6,665 | 2.2% | 1,369 | 2.8% | 5,222 | 2.0% |
| Rollover | 6,665 | 2.4% | 1,369 | 2.8% | 5,222 | 2.0% |
| Collision with pedestrian/bicyclist/animal | 6,665 | 2.0% | 1,369 | 1.4% | 5,222 | 1.9% |
| Collision with movable object | 6,665 | 1.6% | 1,369 | 2.6% | 5,222 | 1.3% |
| Collision with fixed object | 6,665 | 7.0% | 1,369 | 4.9% | 5,222 | 7.5% |
| Collision with 2+ motor vehicles (head on) | 6,665 | 1.1% | 1,369 | 0.9% | 5,222 | 1.2% |
| Collision with 2+ motor vehicles (angle) | 6,665 | 7.8% | 1,369 | 8.1% | 5,222 | 7.7% |
| Collision with 2+ motor vehicles (other) | 6,665 | 75.9% | 1,369 | 76.6% | 5,222 | 76.4% |
| Authorized speed limit (mph) | 6,513 | 49.94 | 1,332 | 52.50 | 5,119 | 49.32 |
| Number of vehicles involved in the crash | 6,672 | 1.61 | 1,369 | 1.66 | 5,229 | 1.60 |
| Number of persons in the vehicle | 6,672 | 1.53 | 1,369 | 1.34 | 5,229 | 1.57 |
| Weather (1 = Not Clear) | 6,672 | 27.4% | 1,369 | 26.2% | 5,229 | 27.7% |
| Ambient light (1 = Non Daylight) | 6,668 | 19.4% | 1,369 | 19.6% | 5,225 | 18.9% |
| Driver contributing circumstances | | | | | | |
| No contributing circumstances | 6,579 | 48.0% | 1,355 | 50.3% | 5,156 | 47.5% |

| | | | | | | |
|---|---|---|---|---|---|---|
| Disregarded signals | 6,579 | 2.2% | 1,355 | 1.5% | 5,156 | 2.4% |
| Speed excess | 6,579 | 25.5% | 1,355 | 19.3% | 5,156 | 27.2% |
| Improper maneuvers | 6,579 | 24.3% | 1,355 | 29.0% | 5,156 | 22.9% |
| Driver vision obstruction (1 = Someone) | 6,431 | 5.8% | 1,326 | 6.4% | 5,057 | 5.6% |
| Driver physical condition (1 = Not normal) | 6,431 | 5.2% | 1,326 | 3.0% | 5,057 | 5.7% |
| Driver race (1 = White) | 6,318 | 73.7% | 1,313 | 72.8% | 4,968 | 73.9% |
| Driver gender (1 = Female) | 6,327 | 35.6% | 1,313 | 18.4% | 4,976 | 40.2% |
| Driver age (years) | 6,336 | 38.92 | 1,315 | 41.24 | 4,983 | 38.30 |
| Vehicle maneuver/action | | | | | | |
| Parked or stopped | 6,588 | 18.6% | 1,358 | 15.8% | 5,165 | 19.5% |
| Going straight ahead | 6,588 | 52.3% | 1,358 | 50.9% | 5,165 | 52.6% |
| Turning/passing | 6,588 | 12.2% | 1,358 | 15.9% | 5,165 | 11.2% |
| Other | 6,588 | 16.8% | 1,358 | 17.4% | 5,165 | 16.7% |
| Estimated speed at impact (mph) | 6,101 | 23.18 | 1,253 | 25.70 | 4,792 | 22.46 |
| Vehicle type | | | | | | |
| Single unit truck | 6,658 | 4.1% | 1,369 | 18.1% | 5,222 | 0.4% |
| Truck/trailer/tractor | 6,658 | 7.6% | 1,369 | 35.9% | 5,222 | 0.2% |
| Passenger Car | 6,658 | 52.1% | 1,369 | 27.8% | 5,222 | 58.9% |
| SUV/Pickup/Van | 6,658 | 32.0% | 1,369 | 16.0% | 5,222 | 36.5% |
| Other | 6,658 | 4.2% | 1,369 | 2.3% | 5,222 | 3.9% |

(1) Coded only for truck-involved collisions (police officer accident reports).

**TABLE 6:** Coefficients for Injury Severity and Total Harm Models—Multi-Vehicle Collisions.

| **Variable** | **Injury (Ordered probit)** All(R) | **Injury (Ordered probit)** Truck(R) | **Injury (Ordered probit)** Non truck(R) | **Harm (Semi-log)** All(R) | **Harm (Semi-log)** Truck(R) | **Harm (Semi-log)** Non truck(R) |
|---|---|---|---|---|---|---|
| Roadway contributing circumstances | | | | | | |
| Work zone | -0.060 | -0.109 | -0.065 | -0.038 | -0.061 | -0.051 |
| Other | 0.034 | 0.022 | 0.059 | 0.024 | 0.012 | 0.031 |
| None(B) | | | | | | |
| Location of the crash | | | | | | |
| Before work area | -0.062 | 0.231 | -0.122* | -0.026 | 0.185** | -0.078* |
| Adjacent to actual work area | 0.003 | 0.336*** | -0.073 | -0.015 | 0.121* | -0.062* |
| In work area approach taper(B) | | | | | | |
| Work zone type | | | | | | |
| Construction | -0.054 | -0.077 | -0.057 | -0.038 | -0.010 | -0.043 |
| Maintenance(B) | | | | | | |
| Utility(B) | | | | | | |
| Intermittent/moving(B) | | | | | | |
| Work zone activity (1 = On-going) | -0.019 | -0.160 | 0.021 | 0.007 | -0.074 | 0.022 |
| Work zone marked with sign/cones (1 = Yes) | -0.036 | -0.288 | 0.033 | 0.012 | -0.118 | 0.043 |
| Type of traffic control device | | | | | | |
| Stop/Yield/Warning sign | 0.087 | 0.012 | 0.115* | 0.031 | -0.004 | 0.060* |
| Stop/Yield/Warning flashing sign | 0.384** | 0.788** | 0.337* | 0.315** | 0.599* | 0.240 |
| Human control | -0.024 | -0.185 | 0.018 | -0.007 | -0.121 | 0.031 |
| Other | 0.142** | 0.152 | 0.164** | 0.104** | 0.052 | 0.124** |
| No control present(B) | | | | | | |
| Dummy for missing cases | 0.046 | 0.089 | 0.004 | -0.004 | 0.038 | -0.042 |
| Roadway configuration | | | | | | |
| One-way, not divided | 0.045 | 0.161 | 0.052 | 0.007 | 0.117 | 0.010 |
| Two-way, not divided | 0.314*** | 0.626*** | 0.288*** | 0.164*** | 0.368*** | 0.144*** |
| Two-way, divided, unprotected | 0.113* | 0.415*** | 0.060 | 0.060 | 0.250*** | 0.029 |
| Two-way, divided, median barrier(B) | | | | | | |
| Construction effect on the roadway(T) | | | | | | |
| Lane closed | | 0.050 | | | 0.013 | |
| Shoulder/Median closed | | 0.330* | | | 0.181* | |
| Roadway closed, detour opposing side | | 0.989*** | | | 0.590** | |
| Lanes shift/become narrow | | -0.695 | | | -0.380* | |
| Other/Unknown | | 0.217 | | | 0.139* | |
| None(B) | | | | | | |
| Type of work being done(T) | | | | | | |
| Repaving/Resurfacing | | -0.166 | | | -0.040 | |
| Shoulder/Median work | | -0.102 | | | 0.010 | |
| New roadway | | -0.219 | | | -0.060 | |
| Other | | -0.001 | | | 0.112 | |
| Unknown(B) | | | | | | |
| Most harmful event (vehicle) | | | | | | |
| Ran off road | 0.446** | 0.041 | 0.511* | 0.283** | -0.009 | 0.364* |
| Rollover | 1.506*** | 1.446*** | 1.546*** | 1.243*** | 1.098*** | 1.265*** |
| Collision with pedestrian/bicyclist/animal | -0.564** | -0.591* | -0.709** | -0.314*** | -0.153 | -0.394*** |
| Collision with movable object | -0.482** | -0.119 | -1.261** | -0.394** | -0.123 | -0.408*** |
| Collision with fixed object | 0.290** | 0.242 | 0.309* | 0.195* | 0.119 | 0.227* |
| Collision with 2+ motor vehicles (head on) | 1.021*** | 0.837* | 1.079*** | 1.114*** | 0.937* | 1.155*** |
| Collision with 2+ motor vehicles (angle) | 0.201** | 0.237 | 0.185** | 0.154*** | 0.137 | 0.155** |
| Collision with 2+ motor vehicles (other)(B) | | | | | | |
| Authorized speed limit (mph) | 0.014*** | 0.028*** | 0.012*** | 0.009*** | 0.012*** | 0.008*** |
| Number of vehicles involved in the crash | 0.093*** | 0.043 | 0.083*** | 0.064*** | 0.023 | 0.049** |
| Number of persons in the vehicle | 0.037** | 0.163*** | 0.033* | 0.240*** | 0.545*** | 0.263*** |
| Weather (1 = Not Clear) | -0.031 | 0.198* | -0.100* | -0.038 | 0.101 | -0.075** |
| Ambient light (1 = Non Daylight) | 0.137** | 0.185 | 0.123* | 0.082** | 0.095 | 0.077 |
| Driver contributing circumstances | | | | | | |
| No contributing circumstances(B) | | | | | | |
| Disregarded signals | -0.108 | -0.418 | -0.114 | -0.035 | -0.141 | -0.046 |
| Speed excess | -0.223*** | 0.203* | -0.315*** | -0.140*** | 0.122* | -0.183*** |

| | | | | | | |
|---|---|---|---|---|---|---|
| Improper maneuvers | -0.217*** | 0.056 | -0.305*** | -0.132*** | 0.050 | -0.186*** |
| Driver vision obstruction (1 = Someone) | 0.142 | 0.185 | 0.103 | 0.102 | 0.117 | 0.081 |
| Driver physical condition (1 = Not normal) | 0.452*** | 0.189 | 0.501*** | 0.385*** | 0.390 | 0.391*** |
| Driver race (1 = White) | -0.119*** | -0.212** | -0.111** | -0.094*** | -0.072 | -0.096*** |
| Driver gender (1 = Female) | 0.228*** | 0.263** | 0.229*** | 0.148*** | 0.130 | 0.142*** |
| Driver age (years) | 0.003** | 0.008*** | 0.002 | 0.001 | 0.005** | 0.000 |
| Vehicle maneuver/action | | | | | | |
| Parked or stopped | 0.350*** | 0.453*** | 0.315*** | 0.218*** | 0.249** | 0.206*** |
| Going straight ahead(B) | | | | | | |
| Turning/passing | -0.007 | -0.372** | 0.048 | 0.005 | -0.144* | 0.026 |
| Other | 0.113* | 0.205 | 0.102 | 0.075** | 0.108 | 0.074** |
| Estimated speed at impact (mph) | 0.009*** | 0.009** | 0.009*** | 0.006*** | 0.005*** | 0.006*** |
| Vehicle type | | | | | | |
| Single unit truck | -0.517*** | -0.768*** | | -0.260*** | -0.400*** | |
| Truck/trailer/tractor | -0.757*** | -0.930*** | | -0.415*** | -0.454*** | |
| Passenger Car(B) | | | | | | |
| SUV/Pickup/Van | -0.154*** | 0.049 | -0.186*** | -0.057** | 0.000 | -0.072** |
| Other | 0.368** | -0.528 | 0.363 | -0.091 | -0.332 | -0.168 |
| Constant | | | | 7.892*** | 6.991*** | 8.035*** |
| Threshold 1 | 1.938*** | 3.342*** | 1.665*** | | | |
| Threshold 2 | 2.902*** | 4.170*** | 2.688*** | | | |
| Threshold 3 | 3.659*** | 4.794*** | 3.554*** | | | |
| Threshold 4 | 4.304*** | 5.664*** | 4.139*** | | | |
| **SUMMARY STATISTICS** | | | | | | |
| N | 5,133 | 1,105 | 4,116 | 5,133 | 1,105 | 4,016 |
| Log likelihood intercept only | -3,967.69 | -799.88 | -3,125.01 | | | |
| Log likelihood at convergence | -3,685.47 | -670.39 | -2,922.23 | | | |
| P> $\chi^2$: LR test ($\chi^2$) / P> $\chi^2$: F test ($\chi^2$) | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| Adjusted $R^2$ | | | | 0.242 | 0.385 | 0.248 |
| McKelvey & Zavoina pseudo $R^2$ | 0.170 | 0.362 | 0.145 | | | |

***, **, and * denote coefficient significantly different from zero at the 1%, 5%, and 10% level of significance (two-tail test), respectively.
(B) Base category for coefficient comparison.
(T) Coded only for truck-involved collisions (police officer accident reports).
(R) Robust standard errors (adjusted for clustering on each collision id).
Coefficients for dummy categorical variables in the Harm models are adjusted for interpretation according to Kennedy *(30)* for a semi-log specification.

**TABLE 7:** Marginal Effects for Truck-Involved Injury Model (Multi-Vehicle Collisions)

| **Variable** | **Injury (Ordered probit) N=1,145** | | | | |
|---|---|---|---|---|---|
| | No injury | Minor | Moderate | Severe | Fatality |
| Roadway contributing circumstances | | | | | |
| Work zone | 0.026 | -0.018 | -0.006 | -0.002 | 0.000 |
| Other | -0.005 | 0.004 | 0.001 | 0.000 | 0.000 |
| None(B) | | | | | |
| Location of the crash | | | | | |
| Before work area | -0.052 | 0.037 | 0.011 | 0.004 | 0.000 |
| Adjacent to actual work area | **-0.079** | **0.055** | **0.018** | **0.006** | **0.001** |
| In work area approach taper(B) | | | | | |
| Work zone type | | | | | |
| Construction | 0.020 | -0.013 | -0.005 | -0.002 | 0.000 |
| Maintenance(B) | | | | | |
| Utility(B) | | | | | |
| Intermittent/moving(B) | | | | | |
| Work zone activity (1 = On-going) | 0.039 | -0.027 | -0.009 | -0.003 | 0.000 |
| Work zone marked with sign/cones (1 = Yes) | 0.062 | -0.044 | -0.013 | -0.004 | 0.000 |
| Type of traffic control device | | | | | |
| Stop/Yield/Warning sign | -0.003 | 0.002 | 0.001 | 0.000 | 0.000 |
| Stop/Yield/Warning flashing sign | **-0.257** | **0.142** | **0.074** | **0.035** | **0.005** |
| Human control | 0.040 | -0.029 | -0.009 | -0.003 | 0.000 |
| Other | -0.039 | 0.026 | 0.009 | 0.003 | 0.000 |
| No control present(B) | | | | | |
| Dummy for missing cases | -0.022 | 0.015 | 0.005 | 0.002 | 0.000 |
| Roadway configuration | | | | | |
| One-way, not divided | -0.033 | 0.024 | 0.007 | 0.002 | 0.000 |
| Two-way, not divided | **-0.163** | **0.107** | **0.040** | **0.015** | **0.002** |
| Two-way, divided, unprotected | **-0.098** | **0.068** | **0.022** | **0.007** | **0.001** |
| Two-way, divided, median barrier(B) | | | | | |
| Construction effect on the roadway(T) | | | | | |
| Lane closed | -0.011 | 0.008 | 0.002 | 0.001 | 0.000 |
| Shoulder/Median closed | **-0.085** | **0.057** | **0.020** | **0.007** | **0.001** |
| Roadway closed, detour opposing side | **-0.320** | **0.173** | **0.093** | **0.046** | **0.007** |
| Lanes shift/become narrow | 0.099 | -0.077 | -0.018 | -0.004 | 0.000 |
| Other/Unknown | -0.053 | 0.036 | 0.012 | 0.004 | 0.000 |
| None(B) | | | | | |
| Type of work being done(T) | | | | | |
| Repaving/Resurfacing | 0.039 | -0.027 | -0.009 | -0.003 | 0.000 |
| Shoulder/Median work | 0.025 | -0.017 | -0.006 | -0.002 | 0.000 |
| New roadway | 0.050 | -0.035 | -0.011 | -0.003 | 0.000 |
| Other | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Unknown(B) | | | | | |
| Most harmful event (vehicle) | | | | | |
| Ran off road(B) | -0.010 | 0.007 | 0.002 | 0.001 | 0.000 |
| Rollover | **-0.511** | **0.199** | **0.166** | **0.117** | **0.029** |
| Collision with pedestrian/bicyclist/animal | **0.100** | **-0.075** | **-0.019** | **-0.005** | **0.000** |
| Collision with movable object | 0.026 | -0.019 | -0.006 | -0.002 | 0.000 |
| Collision with fixed object | -0.064 | 0.043 | 0.016 | 0.006 | 0.001 |
| Collision with 2+ motor vehicles (head on) | **-0.274** | **0.150** | **0.079** | **0.039** | **0.006** |
| Collision with 2+ motor vehicles (angle) | -0.063 | 0.042 | 0.015 | 0.005 | 0.001 |
| Collision with 2+ motor vehicles (other) | | | | | |
| Authorized speed limit (mph) | **-0.007** | **0.005** | **0.002** | **0.001** | **0.000** |
| Number of vehicles involved in the crash | -0.011 | 0.007 | 0.002 | 0.001 | 0.000 |
| Number of persons in the vehicle | **-0.040** | **0.027** | **0.009** | **0.003** | **0.000** |
| Weather (1 = Not Clear) | **-0.051** | **0.034** | **0.012** | **0.004** | **0.000** |
| Ambient light (1 = Non Daylight) | -0.048 | 0.032 | 0.012 | 0.004 | 0.000 |
| Driver contributing circumstances | | | | | |
| No contributing circumstances(B) | | | | | |
| Disregarded signals | 0.077 | -0.057 | -0.015 | -0.004 | 0.000 |
| Speed excess | **-0.052** | **0.035** | **0.012** | **0.004** | **0.000** |

| | | | | | |
|---|---|---|---|---|---|
| Improper maneuvers | -0.013 | 0.009 | 0.003 | 0.001 | 0.000 |
| Driver vision obstruction (1 = Someone) | -0.049 | 0.033 | 0.012 | 0.004 | 0.000 |
| Driver physical condition (1 = Not normal) | -0.050 | 0.033 | 0.012 | 0.004 | 0.000 |
| Driver race (1 = White) | **0.055** | **-0.037** | **-0.013** | **-0.005** | **0.000** |
| Driver gender (1 = Female) | **-0.070** | **0.046** | **0.017** | **0.006** | **0.001** |
| Driver age (years) | **-0.002** | **0.001** | **0.000** | **0.000** | **0.000** |
| Vehicle maneuver/action | | | | | |
| Parked or stopped | **-0.133** | **0.083** | **0.035** | **0.014** | **0.002** |
| Going straight ahead(B) | | | | | |
| Turning/passing | **0.073** | **-0.053** | **-0.015** | **-0.004** | **0.000** |
| Other | -0.054 | 0.036 | 0.013 | 0.005 | 0.000 |
| Estimated speed at impact (mph) | **-0.002** | **0.001** | **0.000** | **0.000** | **0.000** |
| Vehicle type | | | | | |
| Single unit truck | **0.203** | **-0.131** | **-0.050** | **-0.019** | **-0.002** |
| Truck/trailer/tractor | **0.228** | **-0.151** | **-0.055** | **-0.020** | **-0.002** |
| Passenger Car(B) | | | | | |
| SUV/Pickup/Van | -0.017 | 0.009 | 0.005 | 0.003 | 0.000 |
| Other | | | | | |

Note: For dummy variables the marginal effect corresponds to the discrete change from 0 to 1.
(B) Base category for coefficient comparison. (T) Coded only for truck-involved collisions (police officer accident reports).

**TABLE 8:** Coefficients for Injury Severity Models Segmented by Vehicle Type (Multi-Vehicle Collisions)

| **Variable** | **Injury (Ordered probit)** | | **Injury (Ordered probit)** | |
|---|---|---|---|---|
| | Truck (trucks) (R) | Truck (Car/SUV/ Van) (R) | Non truck (Cars) (R) | Non truck (SUV/Van) (R) |
| Roadway contributing circumstances | | | | |
| Work zone | -0.134 | -0.181 | -0.032 | -0.120 |
| Other | 0.263 | -0.001 | 0.026 | 0.199 |
| None (B) | | | | |
| Location of the crash | | | | |
| Before work area | 0.107 | 0.282 | -0.129 | -0.133 |
| Adjacent to actual work area | 0.790*** | 0.109 | -0.105 | -0.032 |
| In work area approach taper (B) | | | | |
| Work zone type | | | | |
| Construction | -0.032 | -0.148 | -0.040 | -0.069 |
| Maintenance (B) | | | | |
| Utility (B) | | | | |
| Intermittent/moving (B) | | | | |
| Work zone activity (1 = On-going) | -0.607*** | -0.010 | -0.017 | 0.077 |
| Work zone marked with sign/cones (1 = Yes) | -8.903*** | -0.174 | 0.074 | -0.078 |
| Type of traffic control device | | | | |
| Stop/Yield/Warning sign | 0.034 | -0.066 | 0.126* | 0.068 |
| Stop/Yield/Warning flashing sign | 1.065** | 0.853* | 0.299 | 0.346 |
| Human control | -0.278 | -0.125 | 0.077 | -0.120 |
| Other | 0.460** | 0.115 | 0.134 | 0.113 |
| No control present (B) | | | | |
| Dummy for missing cases | 0.261 | -0.108 | -0.086 | 0.169 |
| Roadway configuration | | | | |
| One-way, not divided | 0.225 | 0.135 | 0.171 | -0.242 |
| Two-way, not divided | 0.833*** | 0.645*** | 0.361*** | 0.210* |
| Two-way, divided, unprotected | 0.526** | 0.358** | 0.077 | 0.032 |
| Two-way, divided, median barrier (B) | | | | |
| Construction effect on the roadway (T) | | | | |
| Lane closed | 0.273 | -0.119 | | |
| Shoulder/Median closed | 0.321 | 0.352* | | |
| Roadway closed, detour opposing side | 0.553 | 1.115** | | |
| Lanes shift/become narrow | 0.176 | -1.068*** | | |
| Other/Unknown | 0.227 | 0.258 | | |
| None (B) | | | | |
| Type of work being done (T) | | | | |
| Repaving/Resurfacing | -0.402 | -0.015 | | |
| Shoulder/Median work | -0.452* | 0.108 | | |
| New roadway | 0.144 | -0.291 | | |
| Other | -0.395 | 0.241 | | |
| Unknown (B) | | | | |
| Most harmful event (vehicle) | | | | |
| Ran off road (B) | -0.714 | 0.392 | 0.650* | 0.010 |
| Rollover | 1.559*** | 1.864*** | 1.229*** | 1.848*** |
| Collision with pedestrian/bicyclist/animal | -8.237*** | -0.475 | -8.848*** | -0.551 |
| Collision with movable object | -8.662*** | 0.125 | -1.066* | -8.138*** |
| Collision with fixed object | 0.585 | 0.021 | 0.318* | 0.217 |
| Collision with 2+ motor vehicles (head on) | 0.925 | 1.099* | 1.130*** | 0.976*** |
| Collision with 2+ motor vehicles (angle) | 0.783*** | -0.064 | 0.170 | 0.152 |
| Collision with 2+ motor vehicles (other) | | | | |
| Authorized speed limit (mph) | 0.051*** | 0.022*** | 0.014*** | 0.006 |
| Number of vehicles involved in the crash | 0.292*** | -0.069 | 0.109*** | 0.039 |
| Number of persons in the vehicle | 0.531*** | 0.138** | 0.149*** | 0.050* |
| Weather (1 = Not Clear) | 0.380** | 0.136 | -0.036 | -0.227** |
| Ambient light (1 = Non Daylight) | 0.056 | 0.144 | 0.106 | 0.143 |
| Driver contributing circumstances | | | | |
| No contributing circumstances (B) | | | | |

| | | | | |
|---|---|---|---|---|
| Disregarded signals | -7.192*** | -0.332 | 0.060 | -0.650** |
| Speed excess | 0.066 | 0.244 | -0.267*** | -0.467*** |
| Improper maneuvers | -0.491** | 0.252 | -0.276*** | -0.400*** |
| Driver vision obstruction (1 = Someone) | 0.181 | 0.228 | 0.168 | 0.027 |
| Driver physical condition (1 = Not normal) | 0.914* | -0.148 | 0.540*** | 0.576*** |
| Driver race (1 = White) | -0.301* | -0.187 | -0.059 | -0.195** |
| Driver gender (1 = Female) | 0.738** | 0.107 | 0.239*** | 0.217*** |
| Driver age (years) | 0.006 | 0.008** | 0.002 | 0.004* |
| Vehicle maneuver/action | | | | |
| Parked or stopped | 0.439 | 0.479** | 0.360*** | 0.167 |
| Going straight ahead(B) | | | | |
| Turning/passing | -0.075 | -0.436** | 0.057 | -0.013 |
| Other | 0.179 | 0.332* | 0.230*** | -0.154 |
| Estimated speed at impact (mph) | 0.012** | 0.007* | 0.010*** | 0.008*** |
| Vehicle type | | | | |
| Single unit truck | | | | |
| Truck/trailer/tractor | | | | |
| Passenger Car(B) | | | | |
| SUV/Pickup/Van | | | | |
| Other | | | | |
| Constant | | | | |
| Threshold 1 | 6.504*** | 2.585*** | 2.157*** | 1.304*** |
| Threshold 2 | 7.386*** | 3.471*** | 3.216*** | 2.301*** |
| Threshold 3 | 8.115*** | 4.114*** | 4.059*** | 3.329*** |
| Threshold 4 | 9.082*** | 5.021*** | 4.699*** | 3.733*** |
| N | 593 | 506 | 2,440 | 1,537 |
| Log likelihood intercept only | -237.79 | -500.98 | -2,009.87 | -1,055.59 |
| Log likelihood at convergence | -174.23 | -457.19 | -1,877.67 | -969.24 |
| P> $\chi^2$: LR test ($\chi^2$) / P> $\chi^2$: F test ($\chi^2$) | 0.001 | 0.001 | 0.001 | 0.001 |
| Adjusted $R^2$ | | | | |
| McKelvey & Zavoina pseudo $R^2$ | 0.870 | 0.231 | 0.340 | 0.338 |

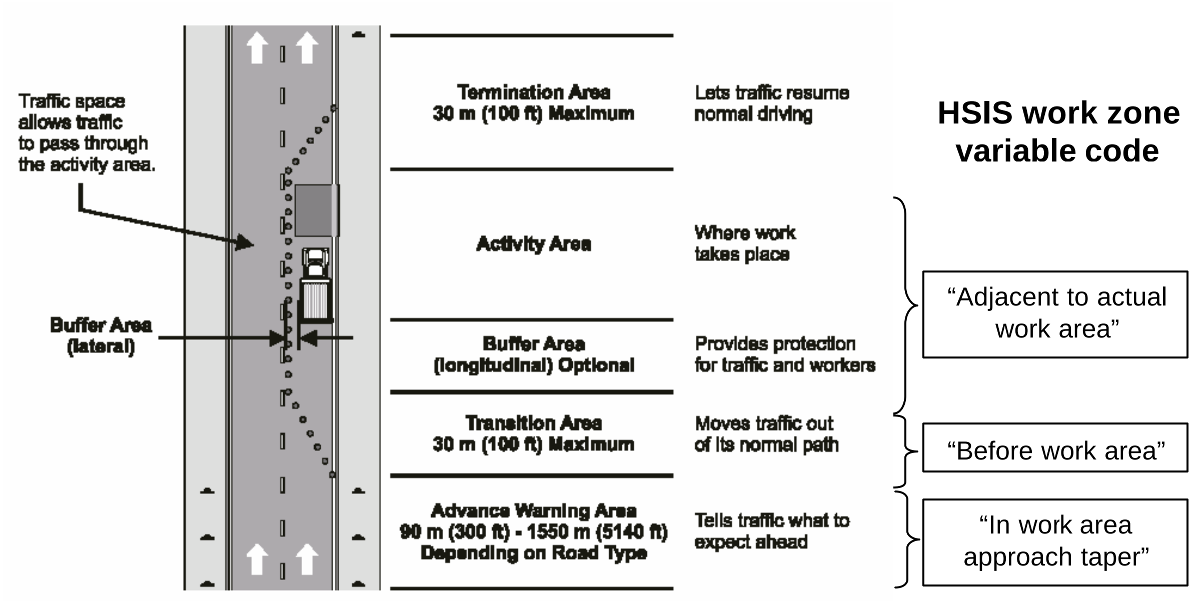
***, **, and * denote coefficient significantly different from zero at the 1%, 5%, and 10% level of significance (two-tail test), respectively.
(B) Base category for coefficient comparison.
(T) Coded only for truck-involved collisions (police officer accident reports).
(R) Robust standard errors (adjusted for clustering on each collision id).

**TABLE 9:** Coefficients for Total Harm Models Segmented by Vehicle Type (Multi-Vehicle Collisions)

| Variable | Harm (Semi-log) | | Harm (Semi-log) | |
|---|---|---|---|---|
| | Truck (trucks) (R) | Truck (Car/SUV/Van) (R) | Non truck (Cars) (R) | Non truck (SUV/Van) (R) |
| Roadway contributing circumstances | | | | |
| Work zone | -0.066 | -0.110 | -0.037 | -0.082* |
| Other | 0.061 | 0.003 | 0.060 | 0.148 |
| None(B) | | | | |
| Location of the crash | | | | |
| Before work area | 0.057 | 0.285* | -0.102** | -0.047 |
| Adjacent to actual work area | 0.174*** | 0.004 | -0.070 | -0.029 |
| In work area approach taper(B) | | | | |
| Work zone type | | | | |
| Construction | -0.023 | -0.057 | -0.048 | -0.025 |
| Maintenance(B) | | | | |
| Utility(B) | | | | |
| Intermittent/moving(B) | | | | |
| Work zone activity (1 = On-going) | -0.159*** | -0.015 | -0.016 | 0.040 |
| Work zone marked with sign/cones (1 = Yes) | -0.055 | -0.112 | 0.032 | -0.049 |
| Type of traffic control device | | | | |
| Stop/Yield/Warning sign | 0.011 | -0.064 | 0.069 | 0.036 |
| Stop/Yield/Warning flashing sign | 0.723* | 0.713 | 0.176 | 0.252 |
| Human control | -0.010 | -0.222 | 0.040 | -0.022 |
| Other | 0.113 | 0.017 | 0.074 | 0.063 |
| No control present(B) | | | | |
| Dummy for missing cases | 0.110 | -0.101 | -0.075 | 0.065 |
| Roadway configuration | | | | |
| One-way, not divided | 0.041 | 0.178 | 0.112 | -0.127 |
| Two-way, not divided | 0.234*** | 0.539*** | 0.238*** | 0.094* |
| Two-way, divided, unprotected | 0.198** | 0.321** | 0.064 | 0.021 |
| Two-way, divided, median barrier(B) | | | | |
| Construction effect on the roadway(T) | | | | |
| Lane closed | 0.080 | -0.048 | | |
| Shoulder/Median closed | 0.117 | 0.300* | | |
| Roadway closed, detour opposing side | 0.051 | 1.116** | | |
| Lanes shift/become narrow | -0.049 | -0.584** | | |
| Other/Unknown | 0.108 | 0.230 | | |
| None(B) | | | | |
| Type of work being done(T) | | | | |
| Repaving/Resurfacing | -0.157* | 0.052 | | |
| Shoulder/Median work | -0.111 | 0.157 | | |
| New roadway | 0.003 | -0.092 | | |
| Other | -0.059 | 0.308 | | |
| Unknown(B) | | | | |
| Most harmful event (vehicle) | | | | |
| Ran off road(B) | -0.067 | 0.135 | 0.410 | 0.012 |
| Rollover | 0.992** | 1.488*** | 0.956** | 1.324*** |
| Collision with pedestrian/bicyclist/animal | -0.157 | -0.466 | -0.643*** | -0.244 |
| Collision with movable object | -0.356*** | 0.093 | -0.508*** | -0.436*** |
| Collision with fixed object | 0.165 | 0.058 | 0.216 | 0.129 |
| Collision with 2+ motor vehicles (head on) | 0.417 | 1.531* | 1.201*** | 1.109*** |
| Collision with 2+ motor vehicles (angle) | 0.211* | 0.020 | 0.135* | 0.087 |
| Collision with 2+ motor vehicles (other) | | | | |
| Authorized speed limit (mph) | 0.010*** | 0.015** | 0.009*** | 0.003 |
| Number of vehicles involved in the crash | 0.100*** | -0.063 | 0.056** | 0.034 |
| Number of persons in the vehicle | 0.684*** | 0.521*** | 0.505*** | 0.407*** |
| Weather (1 = Not Clear) | 0.111** | 0.112 | -0.027 | -0.124*** |
| Ambient light (1 = Non Daylight) | 0.084 | 0.121 | 0.079 | 0.054 |
| Driver contributing circumstances | | | | |
| No contributing circumstances (B) | | | | |
| Disregarded signals | -0.019 | -0.273 | 0.067 | -0.312*** |

| | | | | |
|---|---|---|---|---|
| Speed excess | 0.057 | 0.137 | -0.166*** | -0.209*** |
| Improper maneuvers | -0.052 | 0.186 | -0.188*** | -0.185*** |
| Driver vision obstruction (1 = Someone) | 0.101 | 0.139 | 0.135 | 0.008 |
| Driver physical condition (1 = Not normal) | 0.769 | 0.090 | 0.480*** | 0.377** |
| Driver race (1 = White) | -0.048 | -0.080 | -0.032 | -0.107** |
| Driver gender (1 = Female) | 0.301 | 0.053 | 0.134*** | 0.159*** |
| Driver age (years) | 0.003 | 0.006* | 0.001 | 0.002** |
| Vehicle maneuver/action | | | | |
| Parked or stopped | 0.149 | 0.390** | 0.231*** | 0.117* |
| Going straight ahead(B) | | | | |
| Turning/passing | -0.013 | -0.274* | 0.013 | 0.011 |
| Other | 0.019 | 0.293* | 0.126*** | -0.046 |
| Estimated speed at impact (mph) | 0.004** | 0.006* | 0.006*** | 0.005*** |
| Vehicle type | | | | |
| Single unit truck | | | | |
| Truck/trailer/tractor | | | | |
| Passenger Car(B) | | | | |
| SUV/Pickup/Van | | | | |
| Other | | | | |
| Constant | 6.654*** | 6.896*** | 7.447*** | 8.029*** |
| Threshold 1 | | | | |
| Threshold 2 | | | | |
| Threshold 3 | | | | |
| Threshold 4 | | | | |
| N | 593 | 506 | 2,440 | 1,537 |
| Log likelihood intercept only | | | | |
| Log likelihood at convergence | | | | |
| P> $\chi^2$: LR test ($\chi^2$) / P> $\chi^2$: F test ($\chi^2$) | 0.001 | 0.001 | 0.001 | 0.001 |
| Adjusted $R^2$ | 0.407 | 0.323 | 0.307 | 0.380 |
| McKelvey & Zavoina pseudo $R^2$ | | | | |

***, **, and * denote coefficient significantly different from zero at the 1%, 5%, and 10% level of significance (two-tail test), respectively.
(B) Base category for coefficient comparison.
(T) Coded only for truck-involved collisions (police officer accident reports).
(R) Robust standard errors (adjusted for clustering on each collision id).
Coefficients for dummy categorical variables in the Harm models are adjusted for interpretation according to Kennedy *(30)* for a semi-log specification.




*Source: Manual on Uniform Traffic Control Devices* (*MUTCD*) *2000, U.S. Department of Transportation, Federal Highway Administration. Adapted and modified from* (*22*).

**FIGURE 5:** Location of the crash in a work zone.

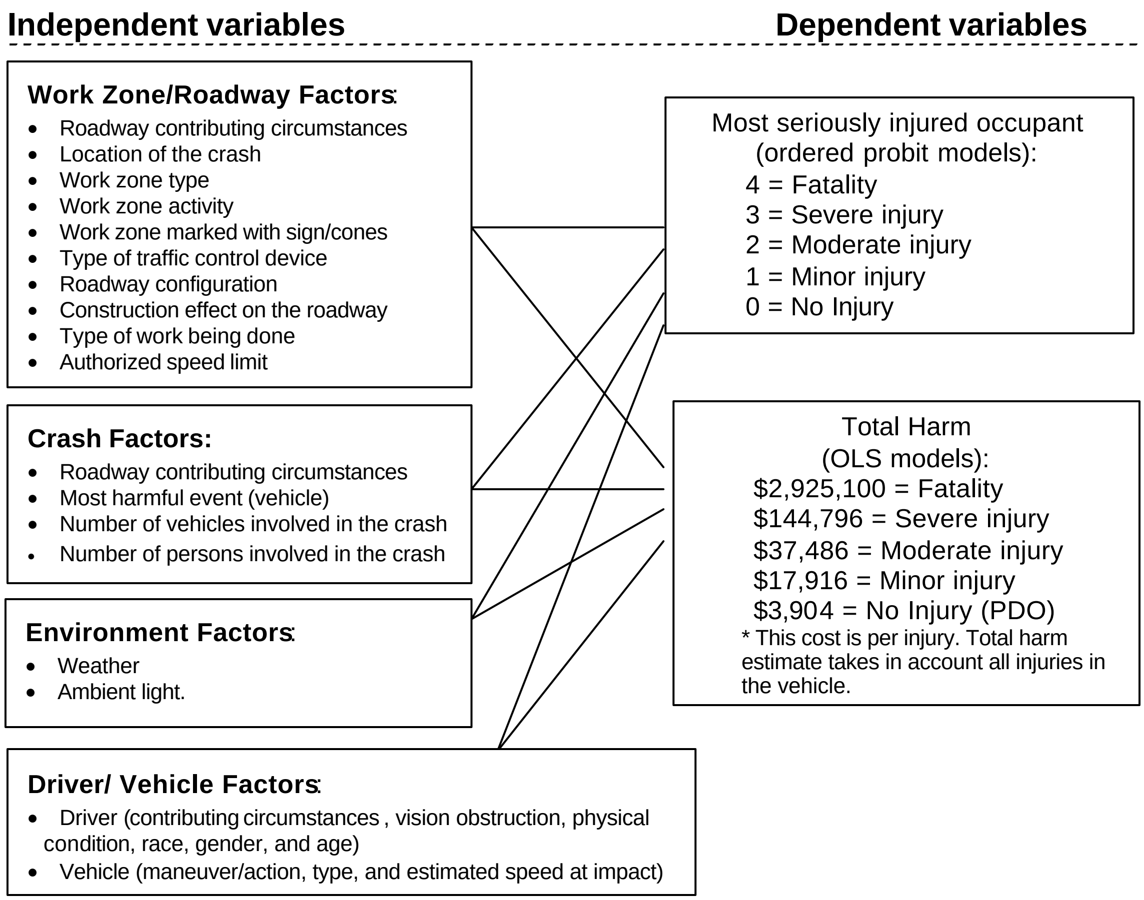
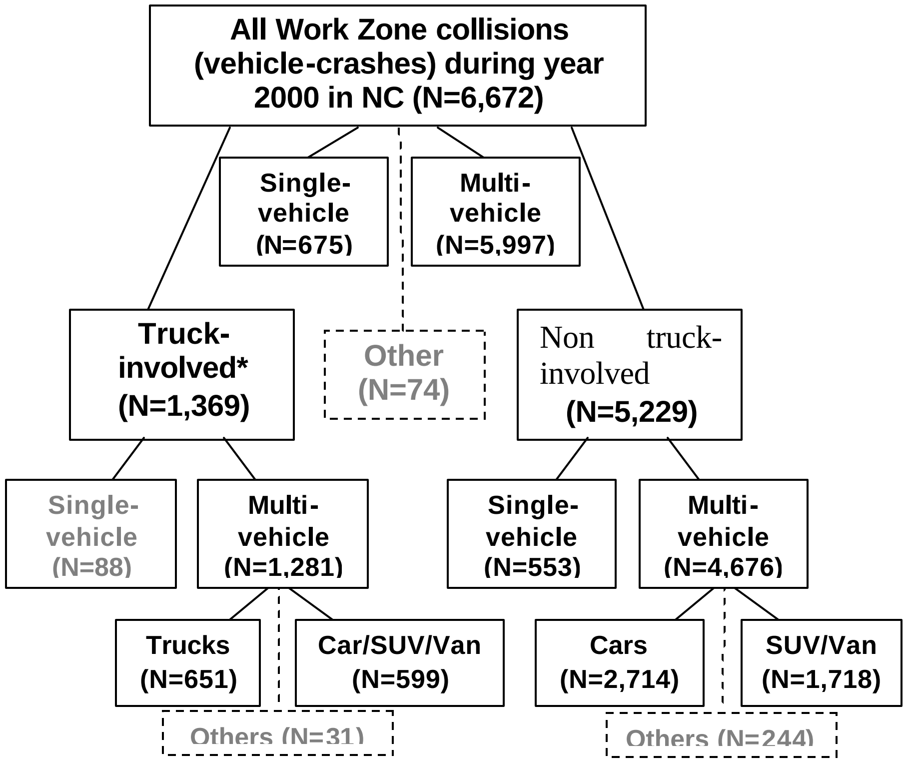
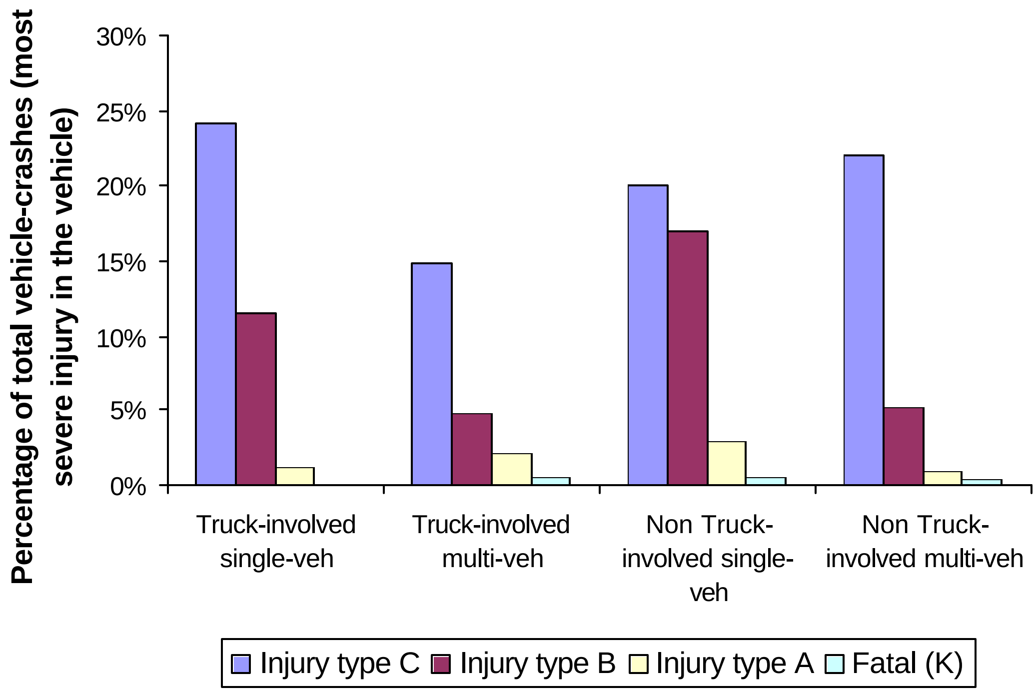


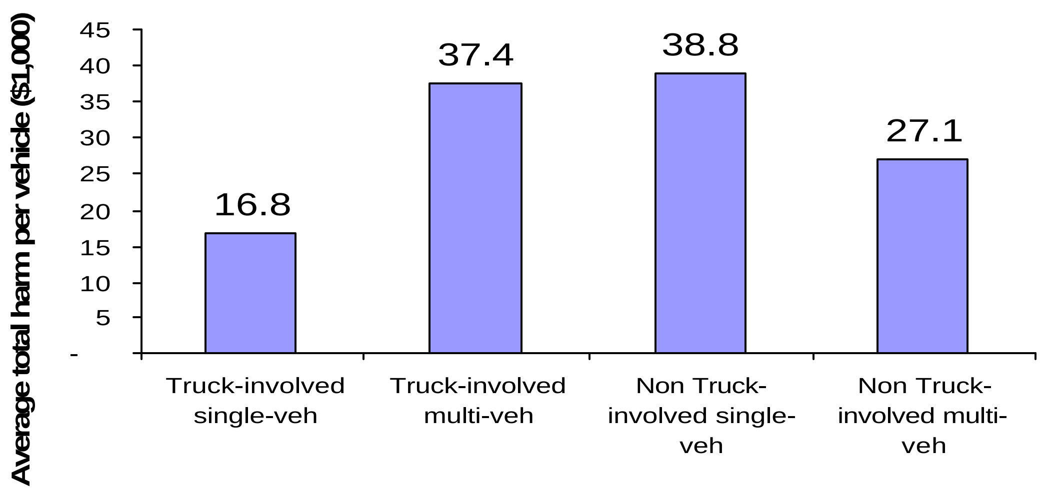

**FIGURE 6 (a):** Conceptual structure.




Notes: Data come from HSIS 2000, which contained four new work zone-related variables (location of the crash, work zone type, work zone activity, and work zone marked with sign/cones). Non truck-involved collisions must involve a Car, SUV, or Van.
* Additional work zone-related data were coded from police officer accident reports (construction effect on the roadway and type of work being done).

**FIGURE 7 (b):** Data structure.




Cost: $2,925,100 = Fatality; $144,796 = Severe injury; $37,486 = Moderate injury; $17,916 = Minor injury; $3,904 = No Injury (PDO)

**FIGURE 8:** Most severe injury and total vehicle harm distribution by type of collision.

# THE EFFECTS OF TRUCK DRIVER WAGES AND WORKING CONDITIONS ON HIGHWAY SAFETY

Daniel A. Rodríguez[a], Marta Rocha[a], Asad Khattak[a] & Michael Belzer[b]
[a]University of North Carolina at Chapel Hill
[b]Wayne State University

**Abstract—**The role that human capital and occupational factors play in influencing driver safety outcomes has gained increased attention from trucking firms and policy-makers. This paper examines the role of these factors, in addition to demographic factors, in influencing the crash frequency of truck drivers. A unique driver-level dataset from a large truckload firm collected over a period of 26 months is used for estimating regression models of crash counts. Based on estimates from a zero-inflated Poisson regression model, results suggest that human capital and occupational factors, such as pay, tenure at the job, and percent of miles driven during winter months, have a significantly better explanatory power of crash frequency than demographic factors. Taking into account both the zero-inflation and the count model, results suggest that higher pay rates and experiencing a pay increase are related to lower expected crash counts and to a higher probability of having no crashes, all else held equal. Although the data for the study come from a single firm, the evidence provided is a first step in examining the structural causes of unsafe driving behavior, such as driver compensation. These results can motivate other firms in modifying operations and driver hiring practices. They also support the need for a broader examination of the relationship between driver compensation and driver safety.

**Keywords:** truck driver safety, compensation, economic rewards, count model

## INTRODUCTION

Although the involvement of large trucks in fatal crashes in the U.S. has dropped substantially over the last decade when measured per unit of travel, the public health burden of large truck crashes, as measured by deaths per 100,000 population, has not improved over time because of the large increase in truck mileage (1). Crashes involving trucks impose costs on truck drivers, road users, trucking firms, shippers, and the public. In 2000 5,211 people died and about 140,000 were seriously injured in large truck-related crashes (2). With trucking operations accounting for almost one third of the total freight ton-miles traveled, and expected to grow in the future, trucking safety continues to demand heightened attention from researchers and policy-makers.

The role that truck driver occupational and behavioral factors play in crash involvement is a particular area that has received increased attention over the last decade. Research examining potential modifications to the hours-of-service regulations and their enforcement (3, 4), detecting and measuring driver alertness and fatigue (5, 6), and understanding driver speeding behavior (7-9), underscore the increasing awareness regarding the importance of driver behavioral factors for trucking safety. Even though an acute research focus on particular truck driver behaviors that increase crash risk is useful, it is at least equally important to confront the factors that motivate such behaviors. At the individual level, such factors include scheduling and operational pressures, pay rate, pay method, and personal characteristics, among others. Indeed, several studies (3, 10-12) have raised questions about the role that occupational and human capital factors play in truck crashes. Therefore, improving our understanding of the importance of the structural causes of certain behaviors, such as compensation and driver economic rewards, and their relationship to driver safety outcomes can lead to appropriate policy responses at the firm and government levels. Insurance companies, trucking firms, shippers, and regulators have an interest in developing such understanding in order to improve the safety outcomes of trucking operations.

In this paper, we examine how compensation and work conditions are associated with frequency of truck crashes at the driver level by estimating count regression models while controlling for driver socio-demographic characteristics and individual exposure. As such, the research focuses on better understanding the occupational factors that can lead to reducing the risk of truck involved collisions as well as in quantifying the relationships between crash frequency and available explanatory variables. To this end, we use a proprietary, driver-level dataset from J.B. Hunt, one of the largest truckload firms in the U.S. Even though the results are not expected to be fully representative of the population of for-hire truckload drivers, we find that J.B. Hunt drivers are comparable to other drivers in terms of demographic and occupational characteristics. Given this, the relationships among the variables may be representative of similar relationships in other firms. Results can be of immediate use to individual trucking firms and can stimulate policy-maker's interest in extending this single-firm study to a more general case. The next section presents a review of the literature focusing on occupational factors and driver safety, followed by a detailed explanation of the data, the crash modeling performed, and the results and implications.

## Literature Review

Human capital theory suggests that variations in human capital across individuals and firms explain differences in labor force outcomes, such as productivity and safety (for the theoretical foundations see 13, 14). Greater job experience, for example, is expected to be related to greater safety. Similarly, because pay can be considered partly a proxy for different levels of human capital, it is expected to be related to better employee outcomes. In a competitive market, higher pay would allow firms to attract and retain drivers with certain characteristics, which will lead to better safety records.

The association between driver behavior, driver pay, and driver characteristics (education, skills, and other experience) suggested by human capital theory has been tested empirically in several studies of the trucking industry. Krass (15) detects a significant inverse relationship between wages and crash risk for the period after economic deregulation of the trucking industry. The work of Hirsch (16) suggests that a substantial fraction of driver wage differences may account for human capital differences among drivers. A recent study concludes that trucking industry compensation and human capital characteristics appear to be more significant determinants of safety than demographic variables (17). Even though other studies have also supported a connection between driver safety and human capital characteristics and driver compensation (10-12, 18), only a handful have examined this relationship explicitly and, to our knowledge, none has focused on crash frequency.

Other research has also examined the link between compensation and occupational factors such as working conditions and driver fatigue. Fatigue is arguably one of the most important risk factor that emerges from analyzing the role of occupational factors in driver safety (4, 6, 8, 19, 20). Most studies of fatigue have examined the causes and the extent of fatigue in truck drivers (for current reviews see 21-23) and more recently the link between fatigue and crash risk (10, 24). McCartt et al. (23, 25) find that drivers perceive the scheduling of loads (measured as driving hours and waiting time for loads) as a significant factor that contributes to fatigued driving. Similarly, after conducting focus groups to examine the factors related to truck crashes, Chatterjee et al. (26) conclude that direct pressure from dispatchers forces drivers to work long hours under unsafe conditions. Lin et al. (27) rely on operational data from another large national less-than-truckload (LTL) carrier to find that total driving time has a greater effect on crash risk than either time of day or driving experience. Not surprisingly, occupational factors also have been associated with illegal substance use (11, 24, 28) and a higher propensity to speed (7, 29).

Related to fatigue is driver (and firm) non-compliance with established hours-of-service regulations, which limit the amount of driving time of truck drivers. Using self-reported data of 498 long-distance drivers, Beilock (3) estimates that 26% of schedules given to drivers result in violations to existing service hours regulations assuming that the average speed limits do not exceed legal limits. In a 1992 survey, Braver et al. (12) found that drivers who violated the hours-of-service rules were more likely to report that they have fallen asleep at the wheel. Clearly, non-compliance with hours of service regulations is also related to negative safety outcomes.

In summary, the literature suggests that certain occupational and human capital factors are related to unsafe driving behaviors and crash outcomes. Tight schedules, fatigue, increasing demands on drivers, and low pay are positively correlated to crash occurrence, although it is unclear if these correlations are carried through to crash frequency. Given the apparent link between level of driver pay and driver safety, one expects that firms would raise pay in order to skim the cream of the trucking labor market. However, this does not seem to be occurring in most cases. Furthermore, on average, earnings of truck drivers and the quality of driving jobs continue to erode, especially among non-union drivers (30). Although speculative, one explanation may be that motor carriers do not perceive that the safety benefits of higher pay offset the increased costs to firms. This points to our limited knowledge about the relationship between pay and driver behavior, and underscores the identified need to develop relevant research that can inform policy-makers and firms. We address the paucity of empirical work regarding the study of crash frequency for truck drivers, while contributing to the understanding of the complex relationship between driver pay and driver safety.

**Data**

A unique longitudinal dataset was used to analyze the association between truck drivers' socio-economic and occupational factors and crash involvement and frequency. The dataset is rich because it contains human resources, operations, and safety data for 11,540 unscheduled over-the-road dry-van tractor-trailer drivers of J.B. Hunt, a major U.S. for-hire truckload company, over a period of 26 calendar-months. The drivers were observed for two periods of 13 months each beginning in September of 1995 and ending in March 1998, with an interval of five months between October 1996 and February 1997 during which no data were collected. Some drivers are observed for a single month while others are observed for the entire 26 months, a characteristic accounted for in the modeling approach discussed below. On average, each driver is observed for 9.2 months.

The end of the first time period (October 1996) coincides with the announcement of changes in the firm's human resource practices designed to improve driver safety, which was then implemented at the beginning of the second period (February 1997). Of particular interest to this study are significant increases in driver per-mile compensation. Only the subset of drivers who were hired before the pay increase announcement and remained with the firm until the pay raise became effective experienced a pay increase. Pay for new hires was substantially higher after J.B. Hunt implemented the new pay policy, so drivers who joined the firm after the pay raise were hired at a higher base pay but did not experience a pay increase.

The longitudinal nature of the data allows us to study the dynamics of truck driver involvement in crashes. The dataset also is unique because it relies on data recorded by the firm, and not on driver recall, a known source of bias present in prevailing survey data. The richness of the data therefore facilitates an in-depth investigation of driver human capital and occupational factors for crash-involved *and* crash-uninvolved drivers, a characteristic commonly unavailable in other crash databases. Despite the longitudinal nature of the data at the driver-month level, we aggregate it at the driver level and conduct the analysis at such level for two reasons. First, we are mainly interested in driver-level attributes such as human capital, which vary more from subject to subject than from month to month. Second, unobserved variables such as vehicle and environmental factors would likely bias the coefficients estimated if we performed this analysis at the crash-level. This would occur if unobserved variables influencing driver safety were correlated with demographic and occupational variables. A driver's age and rate of pay, for example, might correlate with exposure to interstate highways or with the type of vehicle driven. Because type of vehicle and exposure to interstate highways may influence a driver's crash risk, not accounting for their effect can yield biased coefficients for age, pay, or both. The failure to include roadway and environmental factors might also be a source of bias.

**Data validity**

A comparison with other sources of information about the TL sector provides information regarding the degree to which the dataset allows cautious generalizations to other firms in this sector (Table 1). The first source of data for comparison is a survey conducted by the University of Michigan Trucking Industry Program (UMTIP) (for details, see 31). The numbers reported cover 233 full-time drivers who are employees and are paid by the mile. Owner operators and those drivers who are paid hourly are excluded from the figures presented. The second data source is a survey of firms included in the National Survey of Driver Wages published by Signpost, Inc. Signpost surveys approximately 200 truckload firms of various sizes. Most major TL carriers are represented and the set includes a sample of medium sized and smaller carriers. The figures presented are for 102 firms with mileage-paid employee drivers and which responded to the UMTIP survey of Signpost respondent firms regarding their pay practices for non-driving time (for a description of this survey, see 31). The third source of data presented includes figures estimated from a 1999 survey conducted for the Truckload Carriers Association (32).

One major difference between J.B. Hunt and the first two sources of data is in the average length of each dispatch. This may be the result of the firm's reliance on rail transportation for hauling freight over long distances or due to particular characteristics of this individual firm's freight business. However, J.B. Hunt's figures are more similar to those reported by the Truckload Carriers Association (TCA) survey. The other major difference is the average tenure of each driver at the firm. J.B. Hunt's average tenure during the time the data was collected is significantly less than what is suggested by the UMTIP survey or the TCA survey (although this may be a measurement artifact, as measurement methods differ across samples). Compensation or demographic characteristics are very similar.

It is important also to highlight that using firm-specific data has some shortcomings. Most prominent is that the results apply exclusively to the population of drivers belonging to the firm. As a result, any inferences about other truckload drivers are limited. This limitation should be viewed in the context of the relative unavailability of driver-level demographic and occupational data to researchers, which may explain the paucity of research on this topic. Researchers have relied on primary data in a limited number of studies (for example see 24, 33, 34) or have collaborated with firms to examine their human resources and operations data (for example see 27, 35). Even when following the former approach, the ability to make general statements remains an issue. Truck stop surveys, for example, may oversample truckload for-hire carriers and over-the-road drivers. Similarly, self-reports about illegal behaviors such as speeding behavior or violation of the hours-of-service rules can result in known response bias in those surveys.

**Variables observed**

The outcome variable for our analysis is the total number of crashes (at fault/not at fault) for each driver recorded during the period of observation and that involved $500 or more of actual or estimated damages. We tested other crash cost cutoff points for the outcome variable such as $200, $1000 and $2000, and found no significant changes in the results. The dollar figure for the crash is the firm's estimated or actual cost associated with each crash (including bodily injury, property damage, and recording costs to all parties involved, but excluding potential changes in insurance costs) or the firm's actuarial estimates of the cost based on data for past crashes with similar characteristics. Non-casualty costs to the firm (such as productivity losses or the cost of recruiting or hiring new employees) and social costs such as losses to third parties (negative externalities) are excluded from these figures. Drivers have an average of 0.38 crashes during the observed period. Of course, this masks the fact that the majority of drivers (77%) do not record crashes during the period in which they were observed.

Independent variables include measures of working conditions, driver's demographic characteristics, an explicit measure of human capital, and compensation variables, which we use as proxies for unobserved human capital characteristics. The human capital variable explicitly included in the model is driver tenure with the firm when first observed (years). Because tenure is not allowed to vary, the effect of "learning on the job" is not captured by this variable. Instead, we argue that the total number of miles driven captures such effect because the higher the number of miles driven, the higher the driving experience acquired. The compensation variables included are driver pay rate when hired (cents per mile) and the percent pay raise received at the beginning of the second time period, if applicable. Based on the evidence provided by the review of the literature, we expect drivers' rate of pay to be negatively associated with their expected crash count, as higher pay will attract drivers with higher human capital. We also hypothesize that the higher the percent raise, the lower the expected crash count. This incentive effect is due to the anticipated impact of higher pay on individual behavior, effectively making unsafe behavior more costly to the individual.

It is possible that the coefficient for percent pay raise also captures a second effect associated with the specific drivers who get a pay increase. To illustrate this potential bias, consider the drivers who experienced a pay raise. These drivers remained with the firm until the announcement of the pay raise and thus they were the only ones able to enjoy such raise. If drivers who tended to remain with the firm also happen to be safer drivers – a perfectly plausible supposition – their safety would in part be responsible for the fact that they got a pay increase. Thus, the causality for the percent pay raise variable would be muddled for this subset of drivers. We address this potential shortcoming by including two additional dummy variables. The first variable (*Cross*) indicates that a driver was hired before the pay raise and effectively received a pay increase thus capturing the potential influence of a driver's safety record on getting the pay increase. The second variable (*After)* indicates drivers who were hired after the pay raise. This variable measures the effect of human capital characteristics not captured by pay or other observed driver characteristics, and assuming that the firm is able to observe such characteristics. The default category corresponds to drivers hired and who left that company before the pay raise.

Variables that measure working conditions include total number of miles driven during the time each driver is observed (in millions), percentage of total miles driven during winter months (defined as December through March), and the total number of dispatches recorded. All else held equal, we expect that a higher number of dispatches involves a higher expected crash count due in part to the fact that each dispatch may be associated with more unpaid and unproductive waiting time and more frequently pulling-in and out of traffic conflict zones such as docks and urban areas. "Miles driven" is treated as an exposure variable, at the same time that the percentage of miles driven during the winter captures possible seasonal effects of the weather on crash risk.

Demographic characteristics we include are age when first observed, race, sex, and marital status. Descriptive summaries of the data show that the average driver age is 39.69 years and 48% of drivers are married (Table 2). Mirroring the industry, drivers tend to be mostly male (96%) and white (77%). The average pay rate at the time of hire was $0.30 per mile and the pay increase averaged across all drivers is 9%. The latter figure substantially understates the pay raise because only the drivers working with the firm when the wage raise went into effect (24%) received the pay increase. Neither drivers who were hired during the first period and left before the pay raise became effective, nor drivers who were hired at a higher rate during the second period receive a pay raise. Among drivers receiving a pay raise, the average pay increased by 39.5%. Finally, the average miles driven for each driver is 70,000 miles although these vary considerably depending on how long a driver is observed. Finally, 38% of all the miles occurred during a winter month.

## Crash Counts MODELS

To examine the impact of compensation, work conditions, and driver demographic characteristics on crash frequency while controlling for driver exposure, we estimated a model where crash frequency is the dependent variable, and demographic and occupational variables are the independent variables. When "count" dependent variables are treated as continuous variables, estimates derived using ordinary least squares regression can be inefficient, inconsistent, and biased (36). We therefore applied regression models such as Poisson and negative binomial models which are the most appropriate modeling technique for data that have a large number of zeroes and a lower number of positive integer variables (37, 38). In our case, we observed zero crashes for 77.2% of drivers, one crash for 12.9%, two crashes for 6.4%, three crashes for 2.7%, and four or more crashes for less than 1 percent of drivers.

Poisson regression and negative binomial models arguably are the most popular count regression models. The key distinction between the two is that the Poisson model requires that the mean of crash counts equal its variance, while the latter allows for differences between the mean and the variance of crash counts. Such difference in the mean and variance may be the result of unobserved heterogeneity across drivers. Because heterogeneity can also cause excess zeros, and in certain cases will always do so (39), we also examine the appropriateness of using zero-inflated models. Indeed, Lee et al. (37) encourage the application of zero-inflated models in the presence of unobserved heterogeneity. Zero inflated models assume that there are drivers that will always have a crash zero count and other drivers for whom the crash frequency can vary. Thus the zero inflated models identify both processes separately, with a binary function such as a logit or probit equation for determining always-zero cases and a Poisson or negative binomial equation for modeling the counts. This added flexibility of modeling crash counts may assist in crash prediction, and their results can suggest many specific relationships between independent variables and crash rates. For details on the derivation of these count models, see Long (36).

**Results**

Our research approach was to select the best fit among a number of count regression models as determined by visual inspection of predicted versus observed counts, likelihood ratio tests, non-parametric tests, and prior theory (where appropriate). Once we selected a preferred model, we evaluated the contribution of demographic and occupational characteristics to crash frequency. In addition to the different model specifications, we explored alternatives to incorporating the impact of miles driven (exposure) on crash rates.

**Preferred model selection**

Following the suggestion from Lee et al. (37) an examination of the empirical frequency distribution of crashes suggests that the variance of crashes exceeded the mean and that there were a relatively high number of zero cases. This evidence was indicative, although not conclusively, of a poor fit for the Poisson model. Using the independent variables described in Table 2, we fit to the data four models: Poisson, negative binomial, and their zero-inflated counterpart models. Although we have no structural reason to believe that a zero inflation process occurs in the data (i.e., that certain truck drivers are risk-free of crash involvement while driving at least one mile), we specified models that included a zero-inflation process. This is because Cameron and Trivedi (39) suggest that in many cases unobserved heterogeneity can result in excess zeroes in addition to overdispersion. Thus, specifying the model with a zero-inflation process was the result of practical rather than theoretical considerations. In addition, we included the natural log of the exposure variable (miles driven) in each model with its coefficient constrained to one. This implied an assumption that the estimated crash rate increases linearly with exposure.

To determine the best-fitting model we estimate a Poisson model to use as a baseline against which other models could be tested. The three additional count models (negative binomial, zero-inflated Poisson, and zero-inflated negative binomial) were examined against this baseline model. Likelihood ratio tests were used to determine the preferred model when these models were nested (as with the Poisson and the negative binomial model). When one model is not nested within the other, we use the non-parametric statistical test for comparing their fit as proposed by Vuong (for details see 40).

Consistent with our expectations, results suggested that the zero-inflated Poisson model was the preferred model. The preferred model was identified in two steps. First, the overdispersion parameter in the zero-inflated negative binomial specifications was not significantly different from zero (0.468) leading to its rejection over the Poisson model (results not shown). Second, the Vuong statistic comparing the zero-inflated Poisson model to the Poisson model (15.45) suggested that the former had better fit than the latter. A visual comparison between observed crash frequency and predicted crash frequency for each model specification suggested that the zero inflated Poisson model fit the data best. Both the negative binominal and the Poisson models tended to underpredict zero counts and overpredict counts greater than zero. Fitting the same models for different crash cost cutoff points (no cutoff point, $200, $1000 and $2000) instead of the $500 initially specified also resulted in the selection of the zero-inflated Poisson model as the preferred regression model for this data. Furthermore, the estimated coefficients for the results using different crash cost cutoff points for the dependent variable did not vary significantly from those discussed in the next section.

**Estimated influence of human capital and driver occupational factors on crash frequency**
Results for the preferred count model and the always-zero logit equations accompanying each count model are provided in Table 3. For completeness, the estimated coefficients for the inflation equation cannot be interpreted separately from the coefficients of the count equation. The practical implication is that the net effect of both equations on crash count should be considered at all times because failing to do so would have resulted in biased estimates. The left set of columns in Table 3 (model 1) shows the preferred model specification with only demographic characteristics as independent variables and with total miles driven as an exposure variable. The model explains 0.99% of the log-likelihood of the constant-only model (1- (-9,347.292/ -9,437.056)). The importance of the unobserved heterogeneity created by excluding the human capital and occupational variables is reflected in the bias implicit in the estimated demographic variable coefficients.

Inclusion of the occupational and human capital variables (model 2) confirms the relevance of human capital and occupational factors in explaining crash frequency at the driver level, and indicates that information collected regarding crashes should include some level of occupational and human capital data for the involved driver. Such variables increase the share of the log-likelihood explained to 7.84%, all else held equal. Despite the statistical significance of model 2, it has a substantial amount of unexplained variance. This is not entirely surprising given the stochastic nature of crash involvements and the relatively large sample size. However, several additional factors may contribute to improving the explanatory power of the model, such as additional driver-level factors (e.g. driving ability and ability to tolerate fatigue), vehicle factors (vehicle condition), other occupational factors (time spent waiting for loads, or loading and unloading, regularity of schedule, and hours worked/awake), and environmental factors (weather and quality of roads).

In Table 3, the right-hand side column for the preferred Poisson model, labeled factor change, shows the change in the expected crash count given a change of one unit in the independent variable, holding all other variables constant. The factor change in expected crash count is calculated as the exponential of the estimated coefficient for each variable. This column is useful because, unlike linear regression, the coefficients estimated in count regression models do not indicate the effect of a unit change in the $j$th independent variable. Similarly, the partial derivative of the function with respect to the $j$th independent variable (known popularly as the marginal effect) cannot be used to estimate such an effect because the function is not linear.

Instead, the factor change provides information about how to relate each variable with the expected crash count.

The coefficient estimated for the pay rate variable in the count equation for Model 2 suggests that for every additional cent per mile a driver is paid, the expected crash count decreases by 8.15% (1-exp(-0.085)). Evaluated at the mean pay rate of $0.30, this translates into an elasticity of crash count with respect to pay rate of –2.47. However, because the pay coefficient in the always-zero equation has a positive sign, the estimated effect of the count equation underestimates the importance of pay rate. Taking into account both equations jointly, the coefficients suggest that one-cent higher pay increase is related to a 2.22 % lower probability of observing one or more crashes. Figure 1 (left panel) shows how the probability of having a zero crash count varies as driver pay rate increases. The probability in the y-axis accounts for the probability resulting from the inflation model and the count model. As a proxy for unobserved human capital characteristics of drivers, this result implies that higher pay is associated with better driving records. Unobserved human capital characteristics include driving experience, other work experience, and driver character and disposition, among others. Interestingly, the coefficient for the dummy variable *After* indicates that it is not statistically significant. This suggests that drivers who were hired after the pay raise were not inherently safer than drivers hired before the pay raise, once we control for the human capital, occupational, and demographic variables noted before.

Similar to results for pay rate, the coefficient for the percent pay raise variable in the count model and the inflation model jointly suggest that for every additional percentage point in the percent pay raise variable, drivers have a 0.23 percent lower expect crash frequency. The coefficient for percent pay raise suggests that there appears to be a relevant motivational effect related to the pay increase that led drivers to have better safety records. The coefficient for the dummy variable *Cross*, which accounts for the causality issue of who experienced a pay raise, is positive but not statistically significant. Although speculative, it is possible that the link between percent pay raise and crash risk may be mediated by a variable such as intent to quit. At higher pay levels, drivers may be less likely to want to quit and therefore this accumulated experience may reflect positively in their driving record.

For driver occupational factors, we find that the higher the miles driven during winter months the higher the expected crash count. This probability, however, is moderated by the coefficient of the zero-inflation equation. A similar moderation effect is detected for dispatches, where the zero-inflation equation suggests that more dispatches are associated with a higher probability of a crash, but the count equation suggests the opposite. The effect captured by the zero-inflation equation is stronger than the effect captured by the count equation, suggesting overall that higher dispatches are associated with a lower probability of remaining crash-free.

**Estimated influence of demographic factors on crash frequency**

While our primary interest lies with the impacts of human capital and occupational factors on driver safety, it is also useful to examine the estimated influence of control variables on crash frequency. Obtaining reasonable results with respect to other variables lend credence to this exercise. Consistent with prior literature (41-43), the coefficients estimated for the age variable suggest that the expected crash count diminishes as driver age increases, but at a decreasing rate.

Taking into account both equations simultaneously suggests that married individuals are 7.07 percent less likely to have any crashes than non-married individuals. The count-only model equation coefficient suggests that the expected crash count for married drivers is 0.89 times the count of non-married drivers, although this is an underestimate because the zero-always equation is not being taken into account. Finally, looking at both models simultaneously we find that the probability of having no crashes for females is 6.9% lower than for males. This contrasts with recent research (44) finding that, for the population at large, there is no difference between crash involvement rates by sex after controlling for annual miles driven.

As with the age variable, the impact of tenure at the firm when first observed on expected crash count is also quadratic, a result that is consistent with the prior research (e.g., 27). The estimated coefficient for tenure and its square suggest that the probability of having a zero count is highest when the driver has been with the firm for 5.81 years (Figure 1, right panel). In contrast, the net crash count effect of a unit increase from the mean for dispatches or for the percent miles driven during winter is ambiguous because the effect of these two variables in the count equation and the always-zero equation are in opposing directions. Finally, a somewhat surprising result relates to the two control variables introduced to address potential causality problems with the percent raise variable. Coefficients for both dummy variables suggest that there is no difference in the expected crash count between drivers hired before and after the pay raise, all else held equal. This does not mean, however, that there was no safety improvement before and after the pay raise. The coefficients for the compensation variables suggest the opposite: that being hired at a higher pay or getting a pay raise contributed to improving the overall crash record of drivers.

## Conclusions

Recent research has shown that human capital and occupational factors are important predictors of driver crash involvement as well as of unsafe behaviors such as speeding and violation of the hours-of-service rules. Using a different methodological approach and a unique dataset, the results of this research support such evidence by showing that such factors are also important predictors of frequency of crash involvement. In particular, higher pay rates and pay raises are related to lower expected crash counts and to a higher probability of zero crash counts, all else held equal.

The results strengthen the limited empirical evidence linking structural occupational factors of drivers – such as economic rewards – with safety outcomes, and extend it by examining their effects on crash occurrence and frequency of occurrence. The effect of tenure and age on expected crash count exhibits the expected nonlinear form, thereby suggesting that factors that keep drivers at jobs can also contribute to better safety outcomes. Although the evidence provided here is not definitive, it may suffice to motivate changes in human resources and hiring practices for some firms. Policy-makers can use this research to support efforts to investigate trucking labor markets using a broader framework that traces the impact of individual carrier compensation on the truck driver labor market and other related labor markets.

Potential limitations of this study include the firm-specific nature of the data, which constrains the extent to which inferences about the truckload sector can be made, and that several variables, such as driving experience and prior driving record, are not available for all observations. However, comparisons between the current data and three surveys of the truckload sector suggest that the average characteristics of the drivers in this study seem comparable with those of the sector. More importantly, the unavailability of certain variables can introduce bias to the coefficients of certain observed variables. In practical terms, however, variables such as age, marital status, and tenure at the firm are expected to be reasonable proxies for unobserved variables such as driving experience. A natural extension of this study is to include disaggregate data from other firms in order to understand the unique contribution of firm characteristics, such as financial performance and size, to driver safety. Such an approach will also provide the possibility of determining the extent to which pay solely is a proxy for human capital characteristics or if it plays a broader role in motivating employees. At the driver level, other extensions of this study include marrying these types of disaggregate, driver-level data sources with other driver-related variables such as driving hours and non-driving work hours. Similarly, the combination of driver-level data and firm-level data to study the effects of policies such as driver education, compensation, or operating policy that varies by firm.

From the modeling perspective, zero-inflated models improve model fit but reduce the interpretability of the results. Unbiased estimates of expected crash count and of the probabilities of particular crash counts other than zero are not readily available for these models because of the nature of the zero-always inflation equation. Although Poisson and negative binomial models are more practical and provide conceptual simplicity, we find that the more sophisticated zero-inflated models provide enough information to yield useful results while reducing the possibility of introducing bias in the estimated coefficient. The amount of unexplained variation points to the fact that other factors than the ones included in these models are important to explain truck crashes. A comprehensive model should include firm characteristics, vehicle factors, environmental factors and roadway factors along with driver factors. The lack of available data on all of these variables remains an important problem. Despite this, the results provide continued indication of the relevance of accounting for occupational and human capital variables in the study of truck driver safety. Failing to do so may results in biased results of limited value to policy-makers and researchers.

## References


1. Lyman, S. and E. Braver, Occupant Deaths in Large Truck Crashes in the United States: 25 Years of Experience, *Accident Analysis and Prevention*, forthcoming, pp. 1-9.

2. US Department of Transportation, *Large Truck Crash Facts*, Federal Motor Carrier Safety Administration, 2000.

3. Beilock, R., Schedule-Induced Hours-of-Service and Speed Limit Violations Among Tractor Trailer Drivers, *Accident Analysis and Prevention*, vol. 27, 1995, pp. 33-42.

4. Arnold, P. K., L. R. Hartley, A. Corry, D. Hochstadt, F. Penna, and A. M. Feyer, Hours of Work and Perceptions of Fatigue Among Truck Drivers, *Accident Analysis and Prevention*, vol. 29, 1997, pp. 471-477.

5. Wierwille, W. W. and L. A. Ellsworth, Evaluation of Driver Drowsiness by Trained Raters, *Accident Analysis and Prevention,* vol. 26, 1994, pp. 571-581.

6. Wylie, C. D., J. C. Shultz, M. M. Miller, and R. R. Mackie, *Commercial Motor Vehicle Driver Fatigue and Alertness Study Project Report*. Report No. FHWA-MC-97-002, October 1996.

7. Beilock, R., R. Capelle, and E. Page, Speed and Training Factors Associated with Heavy Truck Accidents, *Transportation Quarterly*, vol. 43, 1989, pp. 571-589.

8. Golob, T., "Driver Behavior of Long Distance Truck Drivers: The Effects of Schedule Compliance on Drug Use and Speeding Citations," presented at 74th Transportation Research Board Meeting, Washington, D.C., 1995.

9. Walton, D., Examining the Self-enhancement bias: Professional Truck Driver's Perceptions of Speed, Safety, Skill and Consideration, *Transportation Research Part F: Traffic Psychology and Behavior,* vol. 2, 1999, pp. 91-113.

10. National Transportation Safety Board, *Factors That Affect Fatigue in Heavy Truck Accidents Vol. 1: Analysis*. NTSB/SS-95/01. National Transportation Safety Board, Washington, D.C. 1995.

11. GAO, *Freight Trucking: Promising Approach for Predicting Carriers' Safety Risks*. GAO/PEMD-91-13. United States General Accounting Office (US GAO), Washington, D.C. April, 1991.

12. Braver, E. R., C. W. Preusser, D. F. Preusser, H. M. Baum, R. Beilock, and R. Ulmer, Long Hours and Fatigue: A Survey of Tractor-trailer Drivers, *Journal of Public Health Policy*, 1992, pp. 341-366.

13. Becker, G., *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: Columbia University Press, 1964.

14. Becker, G., Investment in Human Capital: A Theoretical Analysis, *Journal of Political Economy*, vol. 70 (supplement), 1962, pp. 9-44.

15. Krass, A., The Impact of the US Motor Carrier Act of 1980 on Road Safety in California: An Econometric Policy Evaluation, *Logistics and Transportation Review*, vol. 29, 1993, pp. 179-92.

16. Hirsch, B., Trucking Deregulation and Labor Earnings: Is the Union Premium a Compensating Differential, *Journal of Labor Economics*, vol. 11, 1993, pp. 279-301.

17. Monaco, K. and E. Williams, Assessing the Determinants of Safety in the Trucking Industry, *Journal of Transportation and Statistics*, vol. 3, 2000, pp. 69-80.

18. Beilock, R., "Schedule Tightness and Work Hours of Long-Distance Tractor Trailer Drivers in the United States," presented at Canadian Transportation Research Forum, Victoria (BC), 1994.

19. Sucharov, L. J., F. Saccomanno, M. Yu, and J. H. Shortreed, "Effect of Driver Fatigue On Truck Accident Rates.," in *Urban Transport and the Environment for the 21st Century*, L. J. Sucharov, Ed. Boston: Computational Mechanics Publications, 1995, pp. 439-446.

20. Feyer, A.-M., A. M. Williamson, R. Jenkin, and T. Higgins, *Strategies to Combat Fatigue in the Long Distance Road Transport Industry - Bus and Coach Perspective.* CR 122. Australian National Occupational Health and Safety Commission, Sydney., Canberra 1993.

21. Lyznicki, J., T. Doege, R. Davis, and M. Williams, Sleepiness, Driving and Motor Vehicle Crashes, *Journal of the American Medical Association*, vol. 279, 1998, pp. 1908-1913.

22. Hakkanen, H. and H. Summala, Sleepiness at Work Among Commercial Truck Drivers, *Sleep*, vol. 23, 2000, pp. 49-57.

23. McCartt, A., J. W. Rohrbaugh, M. C. Hammer, and S. Z. Fuller, Factors Associated With Falling Asleep at The Wheel Among Long-Distance Truck Drivers, *Accident Analysis and Prevention*, vol. 32, 2000, pp. 493-504.

24. Hensher, D., R. Daniels, and H. Battellino, "Safety and Productivity in the Long Distance Trucking Industry," presented at Proceedings, 16th Australian Research Board Conference, Perth, 1992.

25. McCartt, A., M. C. Hammer, and S. F. Fuller, "Understanding and Managing Fatigued Driving: Study of Long-Distance Truck Drivers in New York State," presented at International Large Truck Safety Symposium, Knoxville, Tennessee, 1997.

26. Chatterjee, A., E. Cadotte, H. Sink, M. Venigalla, and G. Gaides, *Driver-Related Factors Involved With Truck Accidents*. Institute for Transportation Research and Education, Raleigh, NC 1994.

27. Lin, T. D., P. Jovanis, and C. Z. Yang, Modeling the Safety of Truck Driver Service Hours Using Time-Dependent Logistic Regression, *Transportation Research Record 1407*, 1993, pp. 1-10.

28. Hensher, D. and H. C. Battelino, "Long-Distance Trucking: Why Do Truckers Speed?," presented at Australasian Transport Research Forum, 1990.

29. Hensher, D., H. Battellino, J. Gee, and R. Daniels, *Long Distance Truck Drivers On-road Performance and Economic Reward*. Australia, Institute of Transport Studies, Sydney University, December 1991.

30. Belzer, M., *Sweatshops on Wheels: Winners and Losers in Trucking Deregulation*. New York: Oxford University Press, 2000.

31. Belzer, M., D. Rodriguez, and S. Sedo, *Paying for Safety: An Economic Analysis of the Effect of Compensation on Truck Driver Safety*. Report to the Federal Motor Carrier Safety Administration under Prime Contract Number DTFH 61-98-C-0061. Wayne State University, Washington D.C. April 2002.

32. Martin Labbe Associates, *1999 Dry Van Driver Survey*. Prepard for Truckload Carriers Association. Martin Labbe Associated, Ormond Beach, FL 1999.

33. Hertz, R. P., Hours of Service Violations Among Tractor-trailer Drivers, *Accident Analysis and Prevention*, vol. 23, 1991, pp. 29-36.

34. Belman, D. L. and K. A. Monaco, The Effects of Deregulation, De-unionization, Technology, and Human Capital on the Work and Work Lives of Truck Drivers, *Industrial & Labor Relations Review*, vol. 54, 2001, pp. 502-524.

35. Jovanis, P. and H. Chang, Disaggregate Model Of Highway Accident Occurrence Using Survival Theory, *Accident Analysis and Prevention*, vol. 21, 1989, pp. 445-458.

36. Long, S., *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage, 1997.

37. Lee, A. H., M. R. Stevenson, K. Wang, and K. K. W. Yau, Modeling Young Driver Motor Vehicle Crashes: Data With Extra Zero, *Accident Analysis and Prevention*, vol. 34, 2002, pp. 515-521.

38. Miaou, S. P., The Relationship Between Truck Accidents and Geometric Design of Road Sections: Poisson Versus Negative Binomial Regressions, *Accident Analysis and Prevention*, vol. 26, 1994, pp. 471-482.

39. Cameron, A. C. and P. K. Trivedi, *Regression Analysis of Count Data*. Cambridge: Cambridge University Press, 1998.

40. Vuong, Q. H., Likelihood Ratio Tests For Model Selection and Non-nested Hypotheses, *Econometrica*, vol. 57, 1989, pp. 307-333.

41. Campbell, K., Fatal Accident Involvement Rates by Driver Age for Large Trucks, *Accident Analysis and Prevention*, vol. 23, 1991, pp. 287-295.

42. Blower, D., *The Accident Experience of Younger Truck Drivers*. GLCTTR 81-96/01. University of Michigan Transportation Research Institute/Great Lakes Center for Trucking and Transit Research, Ann Arbor, MI 1996.

43. National Highway Traffic Safety Administration, *Large-Truck Accident Causation*. United States Department of Transportation, Research and Development, National Center for Statistics and Analysis, NHTSA, Washington D.C. 1982.

44. Lourens, P. F., J. A. M. M. Vissers, and M. Jessurun, Annual Mileage, Driving Violations and Accident Involvement in Relation to Drivers' Sex, Age and Level of Education, *Accident Analysis and Prevention*, vol. 31, 1999, pp. 593-597.

## List of Tables



**List of Figures**

**TABLE 1 Comparison between J.B. Hunt data and other data sources for the truckload sector**

| Variable | J.B. Hunt | UMTIP driver survey | Signpost and UMTIP firm survey | Truckload Carriers Association |
|---|---|---|---|---|
| Age (years) | 39.69 | 42.18 | n.a. | 41.0 |
| Race (1 = Non-white) | 22.7% | 14% | n.a. | n.a. |
| Married | 48% | 69% | n.a. | n.a. |
| Tenure at firm when first observed (yrs) | 1.20 | 3.46 | n.a. | 4.2 |
| Base pay (cents/mi) | 30.28 | 28.6 | 28.6 | n.a. |
| Miles per dispatch | 575.8 | 858.0 | 905.9 | 686.0 |

**TABLE 2 Summary statistics and variable explanation (N=11,540)**

| Variable label | Explanation | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| **DEMOGRAPHIC** | | | | | |
| Age | Mean driver age | 39.69 | 10.14 | 20 | 76 |
| Female | = 1 if female driver, = 0 otherwise | 0.04 | 0.19 | 0 | 1 |
| Married | = 1 if married, = 0 otherwise | 0.48 | 0.50 | 0 | 1 |
| **COMPENSATION** | | | | | |
| Pay | Pay (cents/mile) when hired | 30.28 | 6.73 | 16 | 49 |
| %Raise | Percentage pay raise | 9.16 | 19.45 | 0 | 123.53 |
| **HUMAN CAPITAL** | | | | | |
| Race | = 1 if non-white, = 0 otherwise | 0.23 | 0.42 | 0 | 1 |
| Tenure | Tenure at firm when first observed (years) | 1.20 | 2.16 | 0.08 | 19.17 |
| **OCCUPATIONAL** | | | | | |
| Miles | Million miles driven during observed period | 0.07 | 0.08 | 0 | 0.5 |
| Mile_win | Percentage of miles driven during winter season (December – March) | 0.38 | 0.33 | 0 | 1 |
| Dispatch | Total number of dispatches during observed period | 128.63 | 138.86 | 1 | 940 |
| **OTHER CONTROL VARIABLES** | | | | | |
| Cross | =1 if driver was hired before the pay raise period and received a pay raise, = 0 otherwise | 0.24 | 0.42 | 0 | 1 |
| After | = 1 if driver was hired after the pay raise occurred, = 0 otherwise | 0.36 | 0.48 | 0 | 1 |
| **DEPENDENT VARIABLE** | | | | | |
| Crash_ct | Crash count per driver | 0.38 | 0.81 | 0 | 8 |

## TABLE 3 Zero-inflated Poisson crash models

| Variable | Model 1 | | | | Model 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Poisson regression | | Logit zero-always (1) | | Poisson regression | | | Logit zero-always (1) | |
| | **Coefficient** | **T-stat** | **Coefficient** | **T-statistic** | **Coefficient** | **T-stat** | **Factor change** | **Coefficient** | **T-stat** |
| Constant | 4.761*** | 19.660 | -0.117 | -0.770 | 6.087*** | 18.150 | | -1.073*** | -2.660 |
| **DEMOGRAPHIC** | | | | | | | | | |
| Age | -0.103*** | -8.650 | -0.002 | -0.660 | -0.038*** | -3.100 | 0.963 | -0.010*** | -2.450 |
| $Age^2$ | 0.001*** | 7.530 | | | 0.001*** | 4.020 | 1.001 | | |
| Female | 0.571*** | 5.460 | 0.487*** | 2.690 | 0.189* | 1.690 | 1.208 | 0.494*** | 2.550 |
| Married | -0.197*** | -4.550 | 0.257*** | 3.220 | -0.112*** | -2.430 | 0.894 | 0.239*** | 2.850 |
| **COMPENSATION** | | | | | | | | | |
| Pay | | | | | -0.085*** | -8.890 | 0.919 | 0.043*** | 2.820 |
| %Raise | | | | | -0.008*** | -3.940 | 0.992 | 0.004 | 1.050 |
| **HUMAN CAPITAL** | | | | | | | | | |
| Race | | | | | 0.160*** | 3.310 | 1.173 | -0.296*** | -3.140 |
| Tenure | | | | | -0.056* | -1.720 | 0.946 | 0.142*** | 2.370 |
| $Tenure^2$ | | | | | 0.006* | 1.860 | 1.006 | -0.011* | -1.710 |
| **OCCUPATIONAL** | | | | | | | | | |
| Miles | 1.000 | | | | 1.000 | | | 6.812*** | 4.140 |
| Mile_win | | | | | 0.417*** | 3.570 | 1.517 | 0.805*** | 4.230 |
| Dispatch | | | | | -0.003*** | -13.350 | 0.997 | -0.005*** | -4.780 |
| **OTHER CONTROL VARIABLES** | | | | | | | | | |
| Cross | | | | | 0.136 | 1.190 | 1.146 | -0.225 | |
| After | | | | | 0.144 | 1.050 | 1.155 | 0.158 | 0.730 |
| | | | | | | | | | |
| Log-L full model | | -9347.292 | | | | | -8,394.167 | | |
| Log-L constant-only (2) | | -9437.056 | | | | | -9,052.593 | | |
| LR test (model 2 vs. 1) | | | | | | | -1906.25 | | |
| Rho-square | | 0.99% | | | | | 7.84% | | |
| N | | 11,540 | | | | | 11,540 | | |
| Vuong statistic | | 16.650*** | | | | | 15.450*** | | |

***, **, and ** denote significance at a 99%, 95% and 90% levels of confidence

(1) The inflation process models the probability of remaining in non-crash state (zero-always), which is the opposite of the count model. As such, the coefficients are expected, although not required, to have different signs in the inflation model than in the count model.

(2) Log-likelihood of constant-only model varies from model 1 to model 2 because the likelihood at convergence of the zero-always logit model differs for both models.

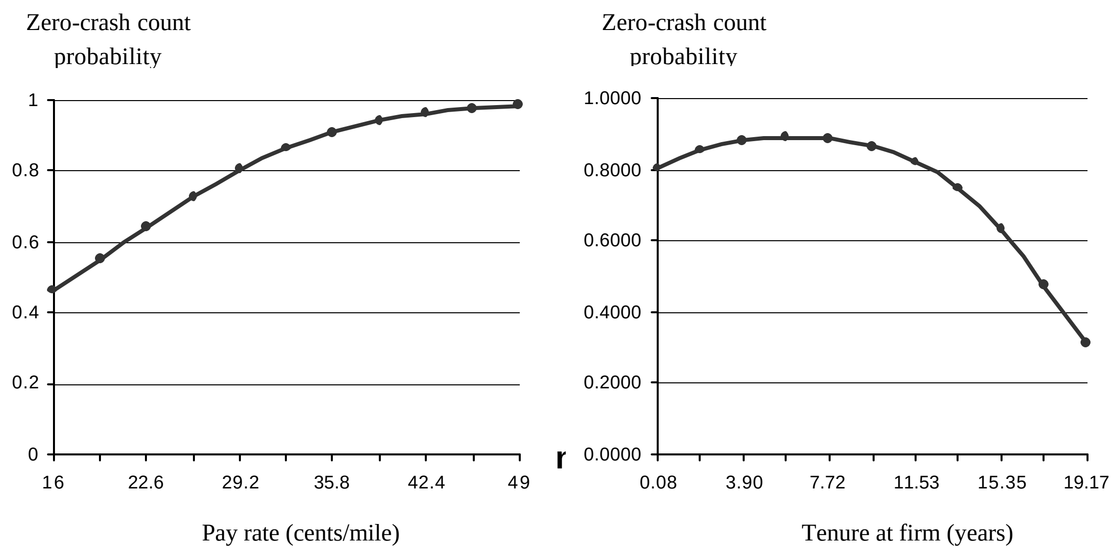



**FIGURE 1 Estimated variations in the probability of having no crashes with changes in pay rate (left panel) and tenure (right panel)**

# PAY INCENTIVES AND TRUCK DRIVER SAFETY

A CASE STUDY

Daniel Rodriguez[a], Felipe Targa[a] & Michael Belzer[b]
[a]University of North Carolina at Chapel Hill
[b]Wayne State University

**Abstract—**This paper provides an empirical examination of the safety consequences of increasing truck driver pay by estimating discrete duration models of driver separation from-the firm and driver crash probability using data from a large-over-the road truckload firm that in 1997 raised wages an average of 33.4%. A two-stage approach allows the isolation of the direct influence of demographic, pay, and operational factors on crash probability from their indirect effect through the probability of leaving the firm. Results suggest that, for a given set of drivers, a pay raise resulted in better crash records, controlling for demographic and operational factors, including prior driving experience and experience acquired on the job. Higher pay rate and lower separation probability are also associated with lower crash probability. To the extent that generalizations about the truckload sector can be made from this study, our findings suggest that human capital characteristics are important predictors of driver safety, but that motivational and incentive factors also play an important role in determining the safety outcomes of truck drivers.

---

**Keywords:** truck driver safety, compensation, economic rewards, hazard model

## INTRODUCTION

Trucking safety has become an increasingly important transportation policy concern in recent years. Between 1992 and 1999 there was a 19% increase in the number of persons who died in crashes involving large trucks. Although such an increase might be expected given the simultaneous 32% increase in the annual number of miles traveled by large trucks, a General Accounting Office (GAO) study points out that fatality levels continue to exceed national goals for reducing fatalities by 50 percent by the year 2009 (1). While trucks experience fewer crashes per mile than passenger cars, trucks have a 68% higher rate of crashes with fatal injuries. Although trucks represent 4% of registered highway vehicles and about 7% of vehicle miles traveled, 12% of all people killed in motor vehicle incidents are involved in a crash with a large truck (2). Clearly, there is both a solid rationale for public concern and a strong impetus for improved data collection and research on the causes of trucking crashes.

Despite growing awareness of the complexity of firms' operating environments and the stochastic nature of crashes, studies of large truck safety continue to focus on human factors, load characteristics, vehicle characteristics and maintenance, and roadway and environmental conditions. However, there is increasing interest in the economic conditions of the trucking industry, and how market pressures may manifest themselves in declining real freight rates, tightening of schedules to meet shipper demands, increased inter-firm competition, and negative safety outcomes (3-6). In particular, changes in wage structures and increased competition caused by economic deregulation in the industry have heightened researchers' interest in the role that employee compensation and industrial relations play in the trucking industry (e.g., 7-9).

Linking wage structures and the occurrence of crashes, in 1990 the National Transportation Safety Board called for a review of trucking industry structure and conditions that may create incentives for unsafe driving behaviors (10). Recent studies have shown that low pay is inversely related to a driver's probability of crash involvement (11) and crash frequency (12). In addition, low levels of pay have been associated with long driving hours and illegal substance use (3, 4, 13), the onset of fatigue (14, 15), and higher propensity to speed (16), and to violate hours-of-service regulations (17, 18). Accordingly, a national survey of for-hire truckload (TL) dry van, flatbed, refrigerated, and tank drivers showed that compensation emerged as a key dimension for human resources improvement (19). At the firm level, a survey of 148 trucking company personnel managers found that managers believe that pay level is the most important factor in drivers' choice for employment with motor carriers (20). By contrast, other studies have suggested that truck driver compensation level plays a less important role in determining crash involvement and driver satisfaction than commonly thought (21-23).

All of these studies have brought issues of compensation and driver behavior to the forefront of the trucking policy debate. However, their identification as important factors does not make clear the causal pathways through which pay influences safety outcomes. Does higher pay elicit desirable (safer) behavior from current drivers? Alternatively, does higher pay attract drivers with human capital characteristics that lead to improved performance, such as better safety outcomes? Although increased safety is a common outcome shared by both causal paths, the policy implications may differ depending on the paths' relative importance. In this paper we focus uniquely on the impact of a pay increase on the safety outcomes of a group of drivers. We use panel data from J.B. Hunt, a large-over-the road truckload firm, to observe the same set of drivers before and after they receive a pay increase. By estimating discrete duration models of crash probability at the driver level, we examine the impact of the pay increase on crash likelihood per month. In doing so, we are able to isolate the influence of the pay increase on current drivers separate from the effect that such an increase may have in attracting different drivers.

The remainder of this paper is divided into 7 sections. By summarizing the literature on the expected safety and separation behavior consequences of changes in driver pay, the following two sections present the questions that animate this study. The third section introduces the research design and data characteristics for the case study. The fourth and fifth sections describe the modeling of driver separation and crash probabilities. The sixth section presents and interprets the study's results, while the final section offers conclusions and implications for further research.

## Expected safety consequences of changes in driver pay level

Human capital theory suggests that variations in human capital across individuals and firms explain differences in labor force outcomes (24), such as productivity and safety. For example, we expect greater job experience to be related to greater safety. At the same time, because pay can be considered partly a proxy for different levels of human capital, it is expected to be related to better employee outcomes (25). In the case of trucking firms, where the labor supply appears to be highly elastic (26, 27), high pay is likely to attract drivers with desirable human capital characteristics such as high driving experience and a low number of prior moving violations. This occurs if drivers with superior human capital receive better compensation packages or if those with higher human capital are more able to obtain well-paying positions in the industry. In either case, certain human capital characteristics of drivers are expected to be related to better driving outcomes, such as being on time, being courteous with customers, and being safer. By ‘skimming the cream’ of the labor market, a firm’s increase in wages is expected to improve safety outcomes.

The association between safety outcomes, driver pay, and driver characteristics suggested by human capital theory has been tested empirically in several studies of the trucking industry. The work of Hirsch (28) suggests that human capital differences among drivers – including education, skills, and other driving experience – may explain a substantial fraction of driver wage differences. Krass (29) detects a significant inverse relationship between wages and crash risk for the period after economic deregulation of the trucking industry. A recent study concludes that trucking industry compensation and human capital characteristics appear to be more significant determinants of safety than demographic variables (11).

In addition to influencing the quality of drivers attracted to the firm, a pay increase is also expected to influence the behavior of current drivers. It may influence driver retention rates, allowing drivers to accumulate general and firm-specific human capital. As such, higher pay may influence driver safety outcomes indirectly. Improvements in safety records can occur in other ways, too. Using a labor-leisure framework where drivers are assumed to trade off leisure time for work time (and earnings), the safety impact of a pay increase for a given set of drivers will depend on the pay rate prior to the increase, current driving hours, the extent of the raise, and the shape of the labor supply curve. This is because at different points on the labor supply curve, higher pay creates substitution effects between work time and leisure time and income effects. The income effect says that if drivers work the same number of hours, they will have more income and therefore they will choose to consume more of all (normal) goods, including leisure. This indicates a reduction in hours worked. By contrast, the substitution effect says that since the mileage rate is higher, leisure is more expensive, and less will be purchased. As a result, this stylized labor-leisure framework suggests that *a priori* the safety impacts of an increase in driver pay are ambiguous.

The labor-leisure framework assumes that drivers are free to select the amount of work they desire. Although this assumption holds more directly for owner-operators and leased owner-drivers than for employees of for-hire truckload firms, the latter also have a say in accepting or rejecting loads. For example, employee drivers can quit a job that gives them an unsatisfactory labor-leisure package. Additionally, the desire for more or less driving may also be reflected in the potential violation of federally-mandated hours of service regulations.

By applying the labor-leisure framework described, Belzer et al. (30) estimate an empirical labor supply curve for truck drivers that conforms to expectations. At low pay levels (less than $0.313 per mile), the curve has a positive slope and thus the substitution effect offsets the income effect. Since most for-hire truckload drivers are paid only for driving time, the implication is that drivers will desire more driving hours and higher earnings. By contrast, at higher pay levels (greater than $0.313 per mile), in the backward-bending part of the labor supply curve, the substitution effect between desired hours driven and desired leisure time devoted to consumption will complement the income effect, thereby leading to a reduction in the desired number of hours worked.

Under current conditions, where hours-of-service rules are commonly violated (see for example 11) and truck driver fatigue is prevalent (see for example 14), we can assume that, at the margin, changes in driving activity by current drivers may lead to measurable changes in safety outcomes. Belman and Monaco (7) confirm this assumption by showing that, for their national sample of drivers, annual miles driven are positively related to having violated the ten-hour driving rule in the last thirty days.

Other explanations suggest that pay raises are expected to improve safety outcomes because this will discourage workers from shirking, since losing a good job imposes a cost on the worker (31). If the cost of monitoring workers is higher than that of the increased wages, Yellen (31) argues that this can be an efficient way for the employer to elicit additional effort from workers. At the core of this argument is an increase in the opportunity cost of engaging in "bad" driving behavior, in part because bad behavior eventually leads to a dismissal (32). Other compensation theories that incorporate notions of fairness and similar social norms to explain the better outcomes from employees, such as improved safety records, include rent-sharing (33) and reciprocal-gift exchange models (34, 35). In sum, although many theories support the hypothesis that higher pay will elicit improved performance, including better safety outcomes of current employees, other theories suggest that the expected safety impact is not clearly identifiable partly because driver preferences for time off and work vary as pay varies. This ambiguity, in addition to societal concerns about trucking safety and the prospect of increasing freight movements by truck, are fundamental motivators of this research.

**Connection between pay, driver separations, and driver safety**

Certain segments of the trucking industry tend to exhibit very high turnover rates (36). Belman and Monaco (7) estimate for their sample that while the median truck driver has been driving for twelve years, the median tenure with the current employers was only eighteen months. These high rates of turnover tend to result from drivers continually looking for better job opportunities, or from disciplinary actions from firm management. Human capital theory also can be used to explain individuals' job separation probability, and to link this probability with driver safety outcomes because the probability of leaving a given firm is expected to decrease as drivers accumulate firm-specific skills (37). If tenure with the firm is low and the probability of separation is high, firm-specific skills are expected to be low and crash probability is expected to be high. By contrast, when separation probability is low, the relationship is reversed unless low separation probability is an indicator of depreciated firm-specific human capital, in which case crash probability would also be high.

Empirically, high employee pay has been correlated with a low probability of separation (e.g., 38-41). Similar results have been reproduced for truck drivers (36, 42-44). Previous research also suggests that truck driver attitudes and perceptions, such as satisfaction with the job, organizational commitment, and sense of trust in the firm, are important mediating variables in the relationship between pay and intent to quit (43, 45). At the individual level, Osterman (46) concluded that a negative association between individual-level job mobility and work performance was well established, and Alexander et al. (47) found that nursing turnover decreased hospital efficiency. At the same time, high labor turnover rates have been associated with negative safety outcomes. For example, the Bureau of Labor Statistics (48) has found that workers were approximately three times more likely to be injured during the first month of employment than during their ninth month of employment.

For trucking firms, the relationships between compensation, turnover, and safety outcomes also have been analyzed (36, 49, 50). Recently, Corsi et al. (51) used the Federal Motor Carrier Safety Administration's SafeStat data to examine the connection between a firm's financial and safety performance for major segments of the trucking industry. They conclude "those [firms] with satisfactory carrier reviews devote a higher percentage of their operating expenses to driver wages than do carriers without positive reviews." Other studies performed at the firm level have shown that turnover rates are positively correlated with injury rates and injury costs (52). Using USDOT's MCMIS crash file, and portraying driver tenure at the firm as an indicator of turnover rates, Feeny (53) shows a significant and inverse relationship between length of service and crash rates, with over half of nearly 200,000 DOT crashes involving drivers with less than a year of experience with the firm. Bruning (54) also finds that drivers with less than one year with a reporting carrier incurred over 50% of crashes in a similarly-sized database.

Correlations between observed driver turnover rates and driver safety outcomes have led to recommendations regarding strategies to reduce driver turnover rates that are also expected to increase driver safety. However, the extent to which the association between pay and turnover rates may be spurious is cause for concern and may result in equivocal prescriptions. If human capital underpins the observed connection between turnover rates and safety, then higher pay may lead to lower separation probabilities and better safety records of newly hired drivers. For current drivers, higher pay increases the opportunity costs of losing the job but its expected impact on crash rates is ambiguous. Furthermore, if individuals with high propensity to leave the firm are at higher risk of crash involvement than individuals with low propensity, then the probability of separation is a source of heterogeneity across those drivers who should be accounted for in understanding driver safety. Failing to do so may result in biased results. Therefore, in this study we also estimate a model of separation probability based on each driver's demographic and occupational characteristics. The coefficients estimated are used to develop a point prediction of separation probability over time for each driver. Such a time-varying estimate is then included as an independent variable in the duration model of crash probability. In addition to providing useful information about the impact of a pay increase on separation probabilities for a given set of drivers, including a predicted separation probability term in the crash model allows us to examine the degree to which separations are associated with safety, while controlling for a wealth of demographic and individual occupational characteristics.

**Data description**

Demographic, operations, compensation, and crash data for 2,715 unscheduled over-the-road drivers of J.B. Hunt are observed over a total of 26 months, in two periods of 13 months each, beginning in September 1995 and ending in March 1998. No data are available for the five months between October 1996 and February 1997.This represents a challenge and a motivation for the methodology selected, which we describe in the next section. The end of the first period coincides with the announcement of changes in the firm's human resource practices designed to improve driver safety and reduce driver turnover rates. Of particular interest to this study are significant increases in driver per-mile compensation. The pay increase was implemented by assigning raises of different percentages to drivers at different pay scales. Drivers at the low end of the pay scale got a larger percent increase in pay than drivers at the high end. Thus, only drivers who were hired before the announcement and remained with the firm until the pay raise became effective experienced a pay increase. These drivers are the ones included in this study.

In adopting a firm-specific focus, we trade off the generality of results obtained from intra and inter-industry, multi-occupation research for the better definition and data resolution provided by this narrow focus. To examine potential limits to our ability to generalize from the data to other truckload industry sub-sectors, we compare the current dataset with three additional sources of information for the TL sector (Table 2). The first source of information is a survey conducted by the University of Michigan Trucking Industry Program (UMTIP) (for details, see 7, 30). Figures reported for UMTIP cover 233 full-time drivers who are employees and paid by the mile. Owner-operators and those drivers who are paid hourly are excluded. The second data source is a survey of firms included in the National Survey of Driver Wages published by Signpost, Inc. Signpost surveys approximately 200 truckload firms of various sizes, including most major TL carriers and a sample of medium-sized and smaller carriers. The figures presented are for 102 firms, with mileage-paid employee drivers, who responded to the UMTIP survey of Signpost respondent firms (see 30). The third source of information is the Federal Motor Carrier Safety Administration's form MC-150. Data on exposure and crash involvement are extracted from these forms for all truckload firms, excluding J.B. Hunt, having more than 200 million miles driven per year. Data for eighteen firms other than J.B. Hunt were collected. These eighteen firms account for more than 35% of all the vehicle miles of the truckload sector, and thus can be a reasonable comparison with J.B. Hunt's safety data.

The major difference between J.B. Hunt and the other sources of data is in the category of miles per dispatch. Although speculative, this may be a result of the firm's reliance on rail transportation for hauling freight over long distances. However, figures from a 1999 survey conducted for the Truckload Carriers Association (55) are close to J.B. Hunt's figures, showing an average of 550 miles per dispatch. Differences in crash rates using SafeStat's reported figures are not notable, with J.B. Hunt having about the same crash rates as the average firm. Furthermore, no meaningful differences are detected in terms of compensation or demographic characteristics, suggesting that the data may be more representative than originally expected.

The potential limitations of using a single firm's data should not obscure the relevance of conducting a partial equilibrium analysis of the relationship between pay changes and safety outcomes. Such analysis may be indicative of the usefulness of engaging in a general equilibrium analysis. Furthermore, because human capital and occupational characteristics are being controlled for experimentally or quasi-experimentally, it may be that the relationship between pay changes and safety examined is representative of similar relationships for other trucking firms. Other strengths of the dataset include its reliance on the firm's database, and not on driver recall, a known source of bias present in prevalent survey data.

*Dependent variable and driver compensation variables*

On average, drivers observed during both time periods experienced a pay increased of 33%, or 9.66 cents per mile driven. Simultaneous changes in the firm's human resource policies other than the pay increase may also influence the propensity of being involved in a crash. The longitudinal nature of the data allows us to control for possible confounding effects between the pay increase and other policy changes by including a variable that distinguishes driving activity occurring before the pay raise, from driving activity that occurred after the pay raise. Drivers in the sample are observed for an average of 14 months, with 12% of the drivers being observed less than 6 months and 59% observed more than 12 months. These figures confirm the importance of better understanding driver turnover in the truckload segment of the trucking industry in the U.S. Furthermore, they indicate the existence of at least two groups of drivers: one that engages in high turnover, often leaving the firm for jobs in other firms or occupations, and another that remains with firms for a longer term.

The dependent variable in the analysis is the occurrence of a crash (regardless of fault) involving $1,000 or more of actual or estimated damages. The dollar figure is the firm's short-term actual costs associated with each crash (including bodily injury, property damage, and recording costs to all parties involved) or the firm's actuarial estimates of the cost based on data for past crashes with similar characteristics. These costs exclude insurance costs. There were 806 crashes during the period observed, corresponding to an average of 0.3 crashes per driver, but 75% of drivers did not have a reported crash costing $1,000 or more. Results from models using different crash cost cutoff points, such as $200 and $500, are very similar to those presented below. However, a higher cost threshold was preferred to emphasize costly crashes, which tend to be more policy relevant than minor crashes.

The driver compensation information collected includes the base pay rate (cents/mile) for each driver when hired by the firm and the percent pay increase for each driver implemented on March 1, 1997. The figure for base pay for each driver is constant during the entire period drivers are observed. In contrast, the percent pay increase is zero up to the point when the pay raise becomes effective. Even though the firm also changed its compensation policy to provide additional income to drivers from safety and productivity bonuses, the vast majority of a driver's income results from miles driven. Because data regarding safety and productivity bonuses are not available to us, we expect that the before-and-after indicator variable will control for such changes in policy. This before-and-after variable will account for the influence of changes in firm policy on driver safety as long as they were concurrent with changes in drivers' pay.

In the context of the empirical labor supply curve estimated by Belzer et al. (30), the average J.B. Hunt driver before the pay raise was on the positive-sloping part of the curve, and thus desired additional driving time with marginal increases in pay. By contrast, the substantial pay increase shifted the average driver to the part of the curve with negative slope, in which marginal increases in pay would result in less desired driving time. For the average driver, the difference between the desired driving time before and after the pay raise is about 1 hour per week.

*Other relevant variables*

Control variables include driver activity data and driver demographic characteristics. Driver activity data include average miles driven up to the beginning of each month, miles driven during each month, month of the year when the activity was recorded, and the total number of dispatches per month. The coefficient for the 'average monthly miles driven up to the beginning of each month' variable can be interpreted as a proxy for driver productivity and earnings. We have no prior expectations regarding the expected sign of the coefficient for this variable, in part because higher earnings may be the result of both regular working hours at high rates of pay and long driving hours at a low pay rate. The anticipated relationship between monthly miles driven during each month and crash likelihood is also ambiguous. On the one hand, higher miles increase crash exposure, all else held equal. On the other hand, high monthly mileage probably is correlated with increased use of interstate highways, which are considered safer than other roads (56). The sign expected for the coefficient of the dispatch variable is positive. Drivers with a higher number of dispatches are also less likely to make long runs and are expected to be in conflict zones more often than drivers with a lower number of dispatches.

We also control for possible seasonal effects of the weather on crash probability. December through March were classified as winter months for this purpose, with all other months as the baseline category. For the variable measuring the number of dispatches, we expect the coefficient to be negative in sign because dispatches are used as a proxy for unpaid and unproductive waiting time and more pulling in and out of traffic conflict zones such as docks and urban areas.

Demographic data used in the analysis include age, race, marital status, sex and driving experience. Because driving experience varies with time, it simultaneously captures overall driving experience and the experience acquired as time on the job increases. Such a distinction is critical in interpreting the coefficient of percent pay raise as the net incentive or motivational effect of the pay increase, net of the experience provided by time on the job. Age is included with both linear and quadratic terms to account for potential nonlinearities with crash probability. Although the dataset is rich in details, the presence of children in the household is not observed. For this variable, marital status is used as a proxy. Table 2 provides descriptive statistics of key information summarized at the individual level.

**Crash probability model**

We apply a multivariate methodology to examine crash probability based on semi-parametric hazard modeling techniques (see 57, 58). Specifically, we define $T_i$ as a discrete random variable representing the duration of stay in a non-crash state for driver *i*. The distributions of durations in a non-crash state are modeled as transition probabilities between a non-crash state and a crash state. The calendar time is not the same for all drivers, and therefore we measure duration on person-specific clocks that each are set to zero when we begin to observe each individual.

The recorded duration is the interval $[t_i\text{-}1, t_i)$ for truck driver $i = 1,\ldots,N$, all of whom are initially in the non-crash state at time *0*. Drivers also are recorded as either having a crash during the interval or as remaining in the non-crash state ('censored cases'). The probability of moving to a crash state having survived until $t$ in a non-crash state (hazard rate $h_{it}$), can be expressed as

$$h_{it} = \Pr(T_i = t \mid T_i > t; X_{it}) \qquad [1]$$

where $X_{it}$ is a vector of covariates summarizing observed differences between individuals, which may vary with time. For cases without crashes the conditional probability of a crash occurring given survival up to time $T_i = t$ is

$$\Pr(T_i > t + s_i \mid T_i > t - 1) = \prod_{t=t}^{t+s_i} (1 - h_{it}) \qquad [2]$$

and for non-censored cases is

$$\Pr(T_i = t + s_i \mid T_i > t - 1) = \left[\frac{h_{it+s_i}}{(1 - h_{it+s_i})}\right] \prod_{t=t}^{t+s_i} (1 - h_{it}) \,. \qquad [3]$$

Given the conditional probabilities, the likelihood for the entire group is expressed as

$$\prod_{i=1}^{n} \left[ \left[ \frac{h_{it+s_i}}{(1-h_{it+s_i})} \right] \prod_{t=t}^{t=s_i} (1-h_{it}) \right]^{d_i} \left[ \prod_{t=t}^{t=s_i} (1-h_{it}) \right]^{1-d_i} , \qquad [4]$$

where $\delta_i = 1$ for non-censored cases and 0 otherwise. The log-likelihood function is thus

$$\sum_{i=1}^{n} d_i * \log\left[ \frac{h_{it+s_i}}{(1-h_{it+s_i})} \right] + \sum_{i=1}^{n} \sum_{t=t}^{t=s_i} \log(1-h_{it}) \,. \qquad [5]$$

In order to specify fully the likelihood function, it is necessary to identify an expression of the hazard rate ($h_{it}$) for this particular process. This specification will have a substantial impact on the inferences made about the process, since the interpretation of the covariates varies according to the hazard specification selected. For this paper a complementary log-log specification was used for the crash hazard rate, which results in a model that is the discrete time counterpart of the continuous time proportional hazards model (58, 59). A proportional hazards specification refers to the influence of any covariate having a multiplicative effect on the baseline hazard function. Such specification has been used elsewhere in other transportation safety research (see 60-62).

Even though there is no compelling reason to support specifying proportional hazards vis-à-vis non-proportional hazards, all model specifications tested here assume proportionality. We test the validity of this assumption by interacting each time invariant explanatory variable with a time variable measuring the length of the spell. This test is performed by re-estimating the models with an additional covariate $X_m$ such that $X_m = X_i \cdot f(t)$, where $X_i$ is an explanatory variable already in the model and $f(t)$ is a particular functional spell length distribution. The subscript for time has been dropped because by definition $X_i$ does not vary with time. A test of the hypothesis that the coefficient for $X_m = 0$ is a test of the proportional hazards assumption for $X_i$.

The duration dependence of the hazard rate is accounted for semiparametrically with dummy variables for time periods during which drivers are observed (63). The underlying assumption is that the probability is constant during the time period captured by each dummy variable. Thus, dummy variables provide information on how the baseline probability rate increases or decreases across time periods, thereby explicitly allowing for occurrences of periodic heterogeneity

**Model for separation from the firm**

Although the estimator used in the previous section was derived with respect to crash probability, we applied a similar reasoning to the derivation of the estimator for examining separation probability (e.g. quits and discharges) with one exception. There are structural reasons suggesting that a proportion of the population is not at risk of leaving the firm during the time observed. For example, certain drivers with high pay and long tenure with the firm might not be willing to leave the firm despite changes in operating conditions. By contrast, the crash probability model estimator presented above assumes that everyone is at risk of a crash as long as they had driving activity during that month—a more plausible assumption for crashes than for individual separations from the firm. As a result, equation [4] is modified for the separation probability estimator to include the probability of never leaving the firm $c$ as follows

$$\prod_{i=1}^{n}\left[c\left[\frac{h_{it+s_i}}{(1-h_{it+s_i})}\right]\prod_{t=t}^{t=s_i}(1-h_{it})\right]^{d_i}\left[\left(c+(1-c)\right)\prod_{t=t}^{t=s_i}(1-h_{it})\right]^{1-d_i} . \qquad [6]$$

Accordingly the separation probability estimator assumes that the population is homogenous with respect to $c$. Although incorporating heterogeneity for $c$ at the driver level is theoretically plausible, such models are difficult to fit. Further, the estimate $c$ provides a testable hypothesis regarding the existence of a split population. If $c = 0$, the model reduces to the standard discrete-time proportional hazards survival model. Derivation of log-likelihood function follows from [5] above. The resulting estimator of the model is therefore what economists commonly refer to as a split population survival model (64).

For model estimation, one modification to the estimation sample and several exclusions of variables in the model specification are necessary. First, drivers who left the firm and who had a crash recorded during the month they left or during the previous month are excluded from the estimation sample. This ensures that drivers who leave the firm do so for reasons other than being at fault for a crash and thus guarantees that the timing of events between separations and crashes conforms to the causal direction implicit in the crash models. As a result, 72 drivers

(2.65 percent of all drivers) and 905 driver-months (2.34 percent of all driver-months) were excluded from the estimation sample of the driver separation model. Because it is not possible to isolate the drivers who leave the firm due to a crash from those that leave for other reasons and happen to have had a recent crash, exclusion of these drivers may sacrifice the efficiency of the estimation, but coefficient estimates will remain unbiased.

Second, four independent variables related to monthly activity are excluded from the separation model specification. These include monthly dispatches, miles driven during the month, whether or not the month falls in winter, and the dummy variable indicating whether the driving activity occurred before or after the compensation policy changes. Since the first two variables vary from month to month, it would be inappropriate to include them in the driver separation model, because they will tend to have low values for drivers who leave the firm during a given month. This is true in part because drivers can quit or be terminated at any time during a month, not only at the end.

Excluding the winter month variable is justified on theoretical grounds because there is no particular expectation that the weather will have an effect on individual separation probabilities. Instead, we include a term capturing peak driving activity, taking the value of one if driving activity occurs between September and December. It is during these months that the driving activity is at its peak, and thus demands on drivers (and earnings) are at their highest level. The final variable excluded is the indicator of driving activity occurring before the policy change, which is not identifiable because the data cover only drivers who were present when the compensation policy changes occurred. Therefore, none of the drivers in the dataset left the firm before the compensation was changed.

**Results and Discussion**

We report estimates from three different models in Table 3: the split population model of driver separation probability (model 1), the proportional hazards model of driver crash probability (model 2), and the same proportional crash hazards model including predicted separation probability as an independent variable (model 3). Appendix 1 contains the coefficients of the baseline hazard estimated for the three models.

*Driver separation model*

Results suggest that the driver separation model is statistically significant at a 99 percent level of confidence. However, driver-level independent variables add limited explanatory power to the model beyond the baseline hazard. For example, the pay variables are not statistically significant and have unexpected signs. Similarly, most demographic variables are not statistically significant except for age, which suggests that as age increases, separation probability decreases but at a decreasing rate. Driving activity variables appear to be more relevant than demographic factors in predicting an individual's separation from the firm. The estimated coefficient for average miles driven per month suggests that for every additional one thousand miles per month driven, separation probability decreases by 38.2 percent, which evaluated at the average miles per month suggests an elasticity of –3.07. A quadratic term for the 'average miles driven per month' variable was used to test nonlinearities in the relationship, but the estimated effects were not statistically significant. The coefficient for the dummy variable for activity occurring during the peak driving season suggests that drivers are 56.6 percent less likely to leave the firm during that time. Other occupational and demographic variables are statistically insignificant. However, the statistical significance of the parameter $c$

indicates the existence of a split population in this sample, with some drivers having a very low likelihood of leaving the firm.

The baseline hazard for this model (Appendix 1) suggests unexpectedly that the longer a person is observed the higher the separation probability, all other variables held equal. To examine potential explanations for this, we estimate the Kaplan Meier or empirical separation hazard for the data. This estimator summarizes the fraction of drivers ongoing at the beginning of a time period who leave the firm during that period, without controlling for other covariates (figure 1). The graph of the empirical separation hazard confirms that the probability of separation is higher the longer a driver is observed. Thus, this effect is carried over to the baseline hazard of the separation model. These two results suggest that the longer a driver is observed, the higher the separation likelihood, controlling for other observed factors. We speculate that this effect may be the result of unrealized expectations by drivers. The prospect of a significant pay raise may have raised individual expectations about work conditions and future earnings. Because these expectations may have not been realized, separations increased several months after the pay raise. Finally, this result should also be understood in the context of the population of drivers with the firm. For the population of drivers, monthly turnover rates decreased from 8.7 percent for twelve months before the pay increase to 3.3 percent for the twelve months after the pay raise.

*Driver crash involvement model*

The driver crash models have substantially better fit than the separation model, with model 2 explaining 53.4% and model 3 explaining 53.6% of the log-likelihood share of a model estimated only with the semiparametric baseline hazard. The likelihood ratio statistic for testing model 2 versus model 3 significantly exceeds the chi-square test statistic with one degree of freedom at a 95.6 percent level of confidence. Thus, introduction of the predicted separation probability term improves the fit of the model. Coefficient signs and statistical significance are consistent across the two crash models except for the age and experience variables, which become statistically insignificant. This is the result of the very high degree of colinearity among independent variables in model 3 created by the introduction of the predicted separation probability term. Nonetheless, the following discussion emphasizes the coefficient estimates from model 3, unless otherwise noted, with the understanding that inferences regarding the age and the experience coefficients are limited due to the colinearity detected.

Driver compensation and driver safety

Examination of the estimated coefficient for percent pay increase contributes to the main objective of the paper – to examine the influence of pay raises on current drivers separate from its effect through newly attracted drivers to the firm. The coefficient suggests that a one percent increase in pay, *ceteris paribus*, is related to a decrease of two percent in crash probability. At the average driver pay raise of 33.4 percent, this translates into a crash probability decrease of 66.8 percent. By contrast, the coefficient for the dummy variable indicating whether the driving activity occurred after operations and compensation policies exhibits a positive sign. This dummy variable captures effects other than the pay raise that occurred simultaneously with the pay change but that we do not observe. The estimate suggests that driving activity after these changes was less safe than activity before the changes. Omitted confounders that may bias the coefficient for driving activity include changes in firms' operating policies, such as honoring requests to return home, and in the assignment of dispatches among drivers. Nonetheless, these results support the idea that pay is not uniquely a proxy for human capital but that it also has an important motivational component for drivers in this sample. Future research is needed to examine the particular causal mechanisms through which the pay raise resulted in improved safety outcomes.

Similarly, by measuring the effect of individual separation probability on driver crash probability, the estimated coefficient for the predicted separation probability variable allows us to isolate the indirect impact of the pay raise on crash probability through separation, from its direct impact. The coefficient confirms expectations created by theory and prior empirical evidence, suggesting that the higher the probability of separation, the higher will be the crash probability. This suggests that strategies that keep drivers for longer periods of time also have the added benefit of decreasing their crash involvement likelihood, and explains firm-level correlations between driver turnover rates and crash rates found by other researchers (36, 65). Furthermore, interpreted jointly with the percent-pay-raise coefficient, these results suggest that an increase in pay is associated with a decreased crash likelihood, while controlling for the raise's influence on the likelihood of leaving the firm.

The coefficient for the base pay rate variable has the hypothesized negative sign. Evaluated at the margin, the base pay coefficient suggests that an increase of one cent per mile from the mean (roughly a three percent higher pay) is associated with a 15.5 percent lower crash probability. At the mean pay of 29 cents per mile driven, this translates into an elasticity of –4.5. Tests for the proportionality of hazards assumption reject it for the coefficients of both compensation variables (base pay rate and percent pay increase). Therefore, the final model specification includes an interaction term between length of observed time (which increases by a unit for every additional month observed for each driver) and base pay rate, and an interaction term between length of observed time and percent pay increase. Because the estimated coefficients for these two interaction terms are positive, they suggest that the negative association between crash probability and the two compensation variables moderates over time. Thus, for example, the longer a driver is observed after a pay raise, the smaller the decrease in crash probability. Insofar as lower crash probability may be a proxy for driver quality more generally, these results suggest that a human resource strategy based on an efficiency wage model, which reduces turnover, may have positive effects on firm performance. Further research would be necessary to confirm the relationship between these policies and firm outcomes.

Driver safety and demographic and occupational factors

While our primary interest lies with the relationship between pay and crash outcomes, it is also useful to examine the estimated influence of control variables used in the study. Results for these control variables can provide further insight regarding the connection between demographic, compensation and occupational factors and driver safety. The coefficients for demographic variables tend to be consistent with existing empirical evidence (66, 67). A non-linear association is detected between truck driver age and crash probability. Driving experience shows a quadratic relationship with respect to crash probability similar to age. With the estimated coefficients, we calculate that a driver with 28 years of driving experience has a crash probability similar to an individual with no truck driving experience. Figure 2 depicts the relationship between crash probability, driver pay (including its interaction with time), and driving experience, holding all other variables constant at their mean. White drivers have a crash likelihood that is 31.7 percent lower than the crash likelihood of drivers of other races, while being unmarried decreases the likelihood of having a crash by 16.7 percent. We speculate that these effects may be confounded by differences in human capital characteristics across races and their impact on labor market outcomes. Regarding sex, we detect no differences between males and females, which is consistent with recent research for the population at large (68).

Two of the four occupational measures are statistically significant. For reasons discussed in the data section, we expected the number of dispatches to be related to higher crash probability. Accordingly, the coefficient of dispatches has the expected positive sign but it is not statistically significant. We did not have prior expectations regarding the expected effects for other occupational variables due to potential omitted variable biases. The coefficient for the 'miles driven during each month' variable suggests that as the number of miles driven during every month increases from the current mean, crash probability decreases. Specifically, for every additional one thousand miles driven during a given month, crash probability decreases by 11.3 percent, all else held equal, indicating a -1.13 elasticity if evaluated at the mean miles driven per month. We believe this result captures a consequence, not a cause of crash involvement. The artifact is that after a crash, the firm might suspend a driver, a driver might take time off voluntarily, or might wait to get another truck, which results in lower miles for that month.

Increases in the average monthly miles driven up to each month, an indicator of overall miles driven, are related to increases in crash likelihood for the given month. There are several speculative explanations for this result. For example, it may be that driver fatigue accumulates over time, and thus individuals driving longer, on average, are more likely to be crash involved in the future. Alternatively, this variable may be measuring scheduling constraints on drivers on long runs. As such, we expect increases in driving distance to be related to lower safety outcomes. Regardless, these results appear to merit additional research. Finally, the coefficient sign for the variable measuring whether the driving activity took place during a winter month suggests a statistically significant decrease in crash likelihood.

Additional information of interest is contained in the baseline hazard function resulting from the time dummy variable coefficients for models 2 and 3 (Appendix 1). The coefficients show negative duration dependence in the underlying hazard after controlling for other covariates (the period between 0 and 2 months is taken as the base category). This means that the longer a driver is observed, the lower the crash probability, and further supports the view that strategies that retain drivers for longer periods appear to have a positive effect on safety.

Finally, it is possible that unobserved heterogeneity in the sample is biasing the coefficients estimated towards negative duration dependence (69). This would be the practical consequence if recruiters hire drivers based on traits not observed by us, such as prior moving violations, prior employment history, character, or disposition, which can explain crash outcomes of drivers. A solution to addressing the problem of heterogeneity, other than incorporating additional variables into the model, is to generalize the hazard rate specification to include an additive error-term $e_i$ at the individual level with mean zero and uncorrelated with the vector of explanatory variables, $X_{it}$ . The error term for the sample is then assumed to follow a parametric distribution and, by integrating it out of the likelihood function, model estimation is feasible (59). This requires imposing a distribution, such as normal, lognormal, or gamma, on failure-prone individuals, with a different distribution for those less "vulnerable" (70). We attempted to parameterize the heterogeneity term imposing two specific distributions, a gaussian mixture distribution and a gamma mixture distribution, as suggested by Meyer (57) and implemented by Jenkins (59). Unfortunately, neither of the parameterizations resulted in reliable estimates of the coefficients. It is possible that the numerical methods used in estimating the coefficients are not reliable due to the large number of individuals observed over time (2,715) and the high correlation that exists for each driver over time.

**Conclusions**

In this study we examine the safety implications of raising driver pay rates for a sample of drivers of a truckload, over-the-road firm whom we observe for time periods ranging between 2 and 26 months. A rich disaggregate dataset including crash involved and crash un-involved drivers is used to estimate duration models of crash probability. Results suggest that after receiving a pay raise drivers have better crash records, controlling for demographic, compensation, occupational factors, prior driving experience, and experience acquired on the job. Thus, in addition to perhaps shifting the average demographic and human capital characteristics of the firm by attracting drivers with different characteristics, these results suggest that the pay increase influenced safety by modifying the behavior of current drivers.

Although the causal chain through which such increases in safety occurred could not be isolated with the current dataset, the results suggest that additional attention to how drivers trade off labor and leisure at different wage rates can contribute to understanding the behavioral links between compensation and driver safety outcomes. Building on established economic theory and empirical evidence substantiating the labor-leisure tradeoff, the findings of this research with respect to pay and safety tend to support regulatory efforts to limit truck driver hours of work as argued by Saltzman and Belzer (71). More broadly, focusing on the structural factors that characterize the industry and their relationship with safety outcomes appears to be a useful area for future research. Not until the behavioral underpinnings of driver safety are better understood in the context of a highly competitive trucking market can firms and policy-makers begin to cope effectively with the motivators of unsafe driving behavior.

Results also suggest that the base pay rate at which drivers are hired also partly is related to drivers' future involvement in crashes. Higher pay is associated with lower crash involvement probability, which is consistent with viewing pay as a proxy for human capital characteristics. Through a model specification that allows for a dual route of influence of demographic and occupational characteristics on crash probability, through separation probability and directly, we detect a positive association between an individual's probability of leaving the firm and the likelihood of crash involvement.

Although these results apply only to J.B. Hunt, other firms may find the implications of this partial equilibrium analysis useful, in particular regarding their driver hiring practices. Comparisons between J.B. Hunt and other large firms in the sector suggest that J.B. Hunt is more representative of the average large firm than originally expected. Thus, to the extent that generalizations about the truckload sector can be made from this study, our findings suggest that human capital characteristics are important predictors of driver safety, but that motivational and incentive factors related to driver pay also play an important role in determining the safety outcomes of truck drivers.

**Appendix 1. Estimated baseline hazard**

| | Model 1[a] | | Model 2[b] | | Model 3[b] | |
|---|---|---|---|---|---|---|
| | Coefficient | Std. Error | Coefficient | Std. Error | Coefficient | Std. Error |
| time3 | 1.176*** | 0.446 | -0.508*** | 0.150 | -0.805*** | 0.211 |
| time4 | 1.167** | 0.551 | -0.624*** | 0.165 | -0.982*** | 0.244 |
| time5 | 1.062 | 0.674 | -0.611*** | 0.174 | -0.935*** | 0.238 |
| time6 | -0.256 | 0.668 | -0.956*** | 0.196 | -0.933*** | 0.197 |
| time7 | -0.756 | 0.655 | -1.308*** | 0.221 | -1.150*** | 0.235 |
| time8 | -1.081 | 0.777 | -1.648*** | 0.249 | -1.418*** | 0.274 |
| time9 | 0.908* | 0.465 | -1.323*** | 0.243 | -1.612*** | 0.283 |
| time10 | 1.196** | 0.481 | -1.790*** | 0.286 | -2.141*** | 0.336 |
| time11 | 1.744*** | 0.458 | -2.098*** | 0.324 | -2.579*** | 0.404 |
| time12 | 1.889*** | 0.470 | -2.474*** | 0.366 | -2.971*** | 0.442 |
| time13 | 2.007*** | 0.490 | -2.611*** | 0.401 | -3.042*** | 0.458 |
| time14 | 1.921*** | 0.534 | -3.069*** | 0.510 | -3.502*** | 0.554 |
| time15 | 2.658*** | 0.496 | -3.237*** | 0.551 | -3.842*** | 0.630 |
| time16 | 2.385*** | 0.548 | -3.987*** | 0.639 | -4.501*** | 0.690 |
| time17 | 2.631*** | 0.546 | -4.582*** | 0.710 | -5.139*** | 0.763 |
| time18 | 1.566*** | 0.609 | -4.533*** | 0.689 | -4.795*** | 0.702 |
| time19 | 2.710*** | 0.521 | -4.578*** | 0.733 | -5.121*** | 0.783 |
| time20 | 2.836*** | 0.531 | -4.594*** | 0.779 | -5.190*** | 0.834 |
| time21 | 2.956*** | 0.544 | -5.309*** | 0.842 | -5.921*** | 0.896 |
| time22 | 2.786*** | 0.576 | -5.731*** | 0.896 | -6.279*** | 0.937 |
| time23 | 2.251*** | 0.663 | -6.330*** | 0.961 | -6.720*** | 0.980 |
| time24 | 3.361*** | 0.576 | -6.711*** | 1.017 | -7.357*** | 1.067 |
| time25 | 2.557*** | 0.898 | -7.097*** | 1.080 | -7.330*** | 1.086 |

*** Significant at a 99% confidence level
** Significant at a 95% confidence level
* Significant at a 90% confidence level

Note: time 1 and 2 are the baseline category.
[a] Split population survival model. For details see Schmidt (64).
[b] Proportional crash probability model. For details see Meyer (57).

**References**


1. Scheinberg, P. F., *Truck Safety: Effectiveness of Motor Carriers Office Hampered by Data Problems and Slow Progress on Implementing Safety Initiatives. Testimony Before the Subcommittee on Ground Transportation, Committee on Transportation and Infrastructure, House of Representatives, March 17, 1999*. United States General Accounting Office, Washington, D.C. 1999.

2. US Department of Transportation, Large Truck Crash Facts, vol. 2002: Federal Motor Carrier Safety Administration, 2000.

3. GAO, *Freight Trucking: Promising Approach for Predicting Carriers' Safety Risks*. GAO/PEMD-91-13. United States General Accounting Office (US GAO), Washington, D.C. 1991.

4. Hensher, D., R. Daniels, and H. Battellino, "Safety and Productivity in the Long Distance Trucking Industry," presented at Proceedings, 16th Australian Research Board Conference, Perth, 1992.

5. Quinlan, M., *Report of Inquiry into Safety in the Long Haul Trucking Industry*. Motor Accidents Authority of New South Wales, Sydney 2001.

6. Traynor, -. T.-L. and -. P.-S. McCarthy, Economic Regulation and Highway Safety in the Trucking Industry: A Limited Dependent Variable Analysis, *Quarterly Review of Economics and Finance*, vol. 33, 1993, pp. 141-153.

7. Belman, D. L. and K. A. Monaco, The Effects of Deregulation, De-unionization, Technology, and Human Capital on the Work and Work Lives of Truck Drivers, *Industrial & Labor Relations Review*, vol. 54, 2001, pp. 502-524.

8. Belzer, M. H., Collective Bargaining in the Trucking Industry: Do the Teamsters Still Count, *Industrial and Labor Relations Review*, vol. 48, 1995, pp. 636-655.

9. Hunter, N. and S. Mangum, Economic Regulation, Employment Relations, and Accident Rates in the U.S. Motor Carrier Industry, *Labor Studies Journal*, vol. 20, 1995, pp. 48-63.

10. National Transportation Safety Board, *Safety study: fatigue, alcohol, other drugs, and medical factors in fatal-to-the-driver heavy truck crashes. Vols 1 and 2*. NTSB/SS-90/01, Washington D.C. 1990.

11. Monaco, K. and E. Williams, Assessing the Determinants of Safety in the Trucking Industry, *Journal of Transportation and Statistics*, vol. 3, 2000, pp. 69-80.

12. Rodriguez, D. A., M. Rocha, A. Khattak, and M. H. Belzer, The effects of truck driver wages and working conditions on highway safety: A case study, *Transportation Research Record*, vol. forthcoming, 2002, pp. 23 pp.

13. Hensher, D. and H. C. Battelino, "Long-Distance Trucking: Why Do Truckers Speed?," presented at 15th Annual Meeting Australasian Transport Research Forum, 1990.

14. Mitler, M. M., J. C. Miller, J. J. Lipsitz, J. K. Walsh, and D. Wylie, The Sleep of Long-Haul Truck Drivers, *The New England Journal of Medicine*, vol. 337, 1997, pp. 755-761.

15. Saccomanno, F., M. Yu, and J. H. Shortreed, Effect of Driver Fatigue On Truck Accident Rates, in *Urban Transport and the Environment for the 21st Century.*, L. Sucharov, Ed., 1995, pp. 439-446.

16. Hensher, D., H. Battellino, J. Gee, and R. Daniels, *Long Distance Truck Drivers On-road Performance and Economic Reward*. Sydney University, Australia, Institute of Transport Studies, Sydney 1991.

17. Braver, E. R., C. W. Preusser, D. F. Preusser, H. M. Baum, R. Beilock, and R. Ulmer, *Who violates work hour rules? A survey of tractor-trailer drivers.* Insurance Institute for Highway Safety, Arlington, VA 1992.

18. Hertz, R. P., Hours of service violations among tractor-trailer drivers, *Accident Analysis and Prevention*, vol. 23, 1991, pp. 29-36.

19. Stephenson, F. J. and R. J. Fox, Driver Retention Solutions: Strategies for For-Hire Truckload (TL) Employee Drivers, *Transportation Journal*, vol. 35, 1996, pp. 12-25.

20. Southern, R. N., J. P. Rakowski, and L. R. Goodwin, Motor Carrier Road Driver Recruitment in a Time of Shortages, *Transportation Journal*, vol. 28, 1989, pp. 42-48.

21. Beilock, R., R. Capelle, and E. Page, Speed and Training Factors Associated with Heavy Truck Accidents, *Transportation Quarterly*, vol. 43, 1989, pp. 571-589.

22. Beilock, R., Schedule-Induced Hours-of-Service and Speed Limit Violations Among Tractor Trailer Drivers, *Accident Analysis and Prevention*, vol. 27, 1995, pp. 33-42.

23. McElroy, J. C., J. M. Rodriguez, G. C. Griffn, P. C. Morrow, and M. G. Wilson, Career stage, time spent on the road, and truckload driver attitudes, *Transportation Journal*, vol. 33, 1993, pp. 5-14.

24. Becker, G., Investment in Human Capital: A Theoretical Analysis, *Journal of Political Economy*, vol. 70 (supplement), 1962, pp. 9-44.

25. Abowd, J., J. Haltiwanger, R. Jarmin, J. Lane, P. Lengermann, K. McCue, K. McKinney, and K. Sandusky, "The relationship between human capital, productivity and market value: building op from micro evidence," presented at Measuring Capital for the New Economy, Washington, D.C., 2002.

26. Hirsch, B., Trucking regulation, unionization, and labor earnings: 1973-1985, *Journal of Human Resources*, vol. 23, 1988, pp. 196-219.

27. Rose, N., Labor rent sharing and regulation: Evidence from the trucking industry, *The Journal of Political Economy,* vol. 95, 1987, pp. 1146-1178.

28. Hirsch, B., Trucking Deregulation and Labor Earnings: Is the Union Premium a Compensating Differential, *Journal of Labor Economics*, vol. 11, 1993, pp. 279-301.

29. Krass, A., The Impact of the US Motor Carrier Act of 1980 on Road Safety in California: An Econometric Policy Evaluation, *Logistics and Transportation Review*, vol. 29, 1993, pp. 179-92.

30. Belzer, M. H., D. A. Rodriguez, and S. A. Sedo, *Paying for Safety: An Economic Analysis of the Effect of Compensation on Truck Driver Safety.* Report to the Federal Motor Carrier Safety Administration under Prime Contract Number DTFH 61-98-C-0061. Wayne State University, Washington D.C. 2002.

31. Yellen, J. L., Efficiency Wage Models of Unemployment, *American Economic Review*, vol. 74, 1984, pp. 200-205.

32. Cappelli, P. and K. Chauvin, An Interplant Test of the Efficiency Wage Hypothesis, *The Quarterly Journal of Economics*, vol. 106, 1991, pp. 769-787.

33. Levine, D. I., Can Wage Increases Pay for Themselves? Tests with a Production Function, *Economic Journal*, vol. 102, 1992, pp. 1102-1115.

34. Burks, S. V., The Origins of Parallel Segmented Labor and Product Markets: A Reciprocity-Based Agency Model with an Application to Motor Freight. Amherst, MA: University of Massachusetts, 1999, pp. 1-24, I-XIV.

35. Milgrom, P. and J. Roberts, *Economics, Organization and Management*. Englewood Cliffs: Prentice Hall, 1992.

36. LeMay, S., T. Stephen, and G. Turner, Driver Turnover and Management Policy: A Survey of Truckload Irregular Route Motor Carriers, *Transportation Journal*, vol. 33, 1993, pp. 15-21.

37. Strober, M. H., Human capital theory: Implications for HR managers, *Industrial Relations*, vol. 29, 1990, pp. 214-239.

38. Cotton, J. L. and J. M. Tuttle, Employee Turnover: A Meta-Analysis and Review with Implications for Research, *Academy of Management Review*, vol. 11, 1986, pp. 55-70.

39. Holzer, H., Wages, Employer Costs, and Employee Performance in the Firm, *Industrial & Labor Relations Review*, vol. 43, 1990, pp. 147S-164S.

40. Leonard, J. S., Carrots and Sticks: Pay, Supervision, and Turnover, *Journal of Labor Economics*, vol. 5, 1987, pp. S136-S152.

41. Munansinghe, L., Wage growth and the theory of turnover, *Journal of Labor Economics*, vol. 18, 2000, pp. 204-220.

42. Gupta, N., D. J. Jenkins, and J. E. Delery, *Motor Carrier Effectiveness: A Study of Personnel Practices, Driver Turnover, and Company Effectiveness in the Trucking Industry*. Mack-Blackwell Transportation Center, Fayettevillle, AR 1996.

43. Richard, M. D. and S. LeMay, Factor-Analytic Logit Approach to Truck Driver Turnover, *Journal of Business and Logistics*, vol. 16, 1995, pp. 281-289.

44. Taylor, G. S., Using Performance Appraisals of Dispatchers to Reduce Driver Turnover, *Transportation Journal*, vol. 30, 1991, pp. 49-55.

45. Richard, M. D., S. A. LeMay, G. S. Taylor, and G. B. Turner, A Canonical Correlation Analysis of Extrinsic Satisfaction in a Transportation Setting, *Logistics and Transportation Review*, vol. 30, 1994, pp. 327-338.

46. Osterman, P., *Turnover, employment security, and the performance of the firm*. Madison, WI: Industrial Relations Research Association, 1987.

47. Alexander, J. A., J. Bloom, and B. A. Nichols, Nursing Turnover and Hospital Efficiency: An Organizational-Level Analysis, *Industrial Relations*, vol. 33, 1994, pp. 505-520.

48. Bureau of Labor Statistics, *Age and Experience Count*. Business Research Publications. Bureau of Labor Statistics, Washington, D.C. 1982.

49. Corsi, T. and P. Fanara, Driver Management Policies and Motor Carrier Safety, *Logistics and Transportation Review*, vol. 24, 1988, pp. 153-163.

50. Shaw, J. D., N. Gupta, and J. E. Delery, Voluntary turnover, work force performance, and financial performance, *Under review, Academy of Management Journal*, 2003, pp. 44 pp.

51. Corsi, T., R. Barnard, and J. Gibney, *Motor Carrier Industry Profile: Linkages Between Financial and Safety Performance Among Carriers in Major Industry Segments*. Report prepared for the Federal Motor Carrier Safety Administration. University of Maryland, College Park, MD 2002.

52. Rinefort, F. and D. D. Van Fleet, Work Injuries and Employee Turnover, *American Business Review*, vol. 16, 1998, pp. 9-13.

53. Feeny, J., *Driver turnover: the leading cause of truck accident frequency*. Motor Carrier Safety Services, Inc., Gladstone, MO 1995.

54. Bruning, E. R., The Relationship Between Profitability and Safety Performance in Trucking firms, *Transportation Journal*, vol. 28, 1989, pp. 40-49.

55. Martin Labbe Associates, *1999 Dry Van Driver Survey*. Prepared for Truckload Carriers Association. Martin Labbe Associated, Ormond Beach, FL 1999.

56. Insurance Institute for Highway Safety, Tractor-trailers, other commercial vehicles are in more multiple-vehicle crashes on most toll roads; single-vehicle crash rate is lower for commercials, in *Status Report*, vol. 34, 1999, pp. 4-5.

57. Meyer, B. D., Unemployment insurance and unemployment spells, *Econometrica*, vol. 58, 1990, pp. 757-782.

58. Prentice, R. and L. Gloeckler, Regression analysis of grouped survival data with application to breast cancer data, *Biometrics*, vol. 34, 1978, pp. 57-67.

59. Jenkins, S. P., Easy estimation methods for discrete-time duration models, *Oxford Bulletin of Economics and Statistics*, vol. 57, 1995, pp. 129-138.

60. Chang, H. and P. Jovanis, Formulating Accident Occurrence As A Survival Process., *Accident Analysis and Prevention*, vol. 22, 1990, pp. 407-419.

61. Jovanis, P. and H. Chang, Disaggregate Model Of Highway Accident Occurrence Using Survival Theory, *Accident Analysis and Prevention*, vol. 21, 1989, pp. 445-458.

62. Mannering, F., Male/female driver characteristics and accident risk: some new evidence, *Accident Analysis and Prevention*, vol. 25, 1993, pp. 77-84.

63. Box-Steffensmeier, J. M. and B. S. Jones, Time is of the essence: Event history models in Political Science, *American Journal of Political Science*, vol. 41, 1997, pp. 1414-1461.

64. Schmidt, P. and A. D. Witte, Predicting criminal recidivism using 'split population' survival time models, *Journal of Econometrics*, vol. 40, 1989.

65. Corsi, T. and P. Fanara, Deregulation, New Entrants, and the Safety Learning Curve, *Journal of the Transportation Research Forum*, vol. 29, 1988, pp. 3-8.

66. Blower, D., *The Accident Experience of Younger Truck Drivers*. GLCTTR 81-96/01. University of Michigan Transportation Research Institute/Great Lakes Center for Trucking and Transit Research, Ann Arbor, MI 1996.

67. Campbell, K., Fatal Accident Involvement Rates by Driver Age for Large Trucks, *Accident Analysis and Prevention,* vol. 23, 1991, pp. 287-295.

68. Lourens, P. F., J. A. M. M. Vissers, and M. Jessurun, Annual Mileage, Driving Violations and Accident Involvement in Relation to Drivers' Sex, Age and Level of Education, *Accident Analysis and Prevention*, vol. 31, 1999, pp. 593-597.

69. Heckman, J. and B. Singer, Econometric Duration Analysis, *Journal of Econometrics,* vol. 24, 1985, pp. 63-132.

70. Lancaster, T., *The Econometric Analysis of Transition Data*. Cambridge: Cambridge University Press, 1990.

71. Saltzman, G. M. and M. H. Belzer, The Case for Strengthened Motor Carrier Hours of Service Regulations, *Transportation Journal*, vol. 41, 2002, pp. 51-71.

**Table 1. Comparison between current dataset and selected truckload sector datasets**

| Variable | JB Hunt | UMTIP driver survey | Signpost and UMTIP firm survey | Safest at |
|---|---|---|---|---|
| Age (years) | 41.86 | 42.18 | n.a. | n.a. |
| Race (1 = White) | 72.7% | 86% | n.a. | n.a. |
| Marital status | 48.7% | 31% | n.a. | n.a. |
| Driving experience (yr.) | 4.29 | 3.46 [b] | n.a. | n.a. |
| Base pay (cents/mi) | 28.95 | 28.6 | 28.6 | n.a. |
| Miles per dispatch [a] | 582 | 858.0 | 905.9 | n.a. |
| Crashes per million miles [c] | 0.29 | n.a. | n.a. | 0.32 |
| Injury or fatal crashes per million miles | 0.14 | n.a. | n.a. | 0.15 |
| Fatal crashes per million miles | 0.014 | n.a. | n.a. | 0.013 |

[a] Average at the person-month level, not the individual level as in Table 1.
[b] Tenure at the firm, not driving experience.
[c] Only crashes involving a tow, injury, or fatality are included.
n.a. = not available

**Table 2: Descriptive statistics summarized at the individual level** [a]

| Variable | Mean or Percentage | Standard Deviation | Min | Max |
|---|---|---|---|---|
| Months observed | 14.3 | 6.4 | 1 | 21 |
| Age (years) | 41.9 | 9.6 | 20.6 | 71.3 |
| Sex (1 = Female) | 2.5% | | | |
| Race (1 = White) | 72.7% | | | |
| Marital status | 48.7% | | | |
| Base pay (cents/mi) | 29.0 | 5.2 | 17 | 47 |
| Percentage pay increase | 33.4% | 22.8 % | 0% | 123.5% |
| Miles driven during month observed | 9,113 | 2,979 | 6 | 21,391 |
| Avg. miles driven up to beginning of observed month | 8,030 | 3,220 | 6 | 22,751 |
| Dispatches per month | 16.0 | 4.7 | 1 | 40.7 |
| Percentage of driving during winter | 25.7% | | | |
| Percentage of driving activity after compensation policy change | 50.1% | | | |
| Total driving experience at t=0 (yr) | 4.3 | 4.7 | 0 | 36.5 |
| Proportion of drivers who left the firm | 10% | | | |
| Percentage of driving during peak season | 25.9% | | | |
| Crashes per driver (> $200) | 0.6 | 0.8 | 0 | 4 |
| Crashes per driver (> $1,000) | 0.3 | 0.6 | 0 | 4 |
| Crashes per driver (> $2,000) | 0.2 | 0.4 | 0 | 4 |

N = 2,715 (aggregated from 38,737 driver-month observations)
[a] For variables that change with time, such as miles driven, summary statistics at the individual level may provide a skewed depiction.

**Table 3: Driver level split population separation model and discrete time proportional crash probability models**

| | Model 1[(a)] (Dep. var: separation = 1) | | | Model 2[(b)] (Dep. var: crash = 1) | | | Model 3[(b)] (Dep. var: crash = 1) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | S.E | Percent change[(c)] | Coeff. | S.E | Percent change[(c)] | Coeff. | S.E | Percent change[(c)] |
| Age (years) | -0.167*** | 0.029 | | -0.084*** | 0.011 | | -0.039 | 0.025 | |
| $Age^2$ | 0.030*** | 0.006 | | 0.019*** | 0.002 | | 0.011** | 0.005 | |
| Sex (1 = Female) | 0.553 | 0.389 | 73.9 | -0.265 | 0.262 | -23.3 | -0.414 | 0.272 | -33.9 |
| Race (1 = White) | -0.120 | 0.179 | -11.3 | -0.382*** | 0.072 | -31.7 | -0.347*** | 0.074 | -29.3 |
| Marital status (1=single) | -0.090 | 0.162 | -8.6 | -0.182*** | 0.068 | -16.7 | -0.153** | 0.070 | -14.2 |
| Base pay (cents/mi) | 0.028 | 0.028 | 2.8 | -0.167*** | 0.015 | -15.4 | -0.168*** | 0.015 | -15.5 |
| % pay increase | 0.429 | 0.508 | 53.5 | -4.122*** | 0.961 | -98.4 | -3.913*** | 0.964 | -98.0 |
| Activity date (1=after pay increase) | | | | 0.589*** | 0.199 | 80.3 | 0.556*** | 0.199 | 74.4 |
| Avg. miles driven up to observed month (000s) | -0.482*** | 0.044 | -38.2 | 0.049*** | 0.018 | 5.1 | 0.180*** | 0.068 | 19.7 |
| Miles driven observed month (000s) | | | | -0.120*** | 0.018 | -11.3 | -0.122*** | 0.018 | -11.5 |
| Dispatches per month | | | | 0.008 | 0.008 | 0.8 | 0.009 | 0.008 | 0.9 |
| Total driving experience at t=0 (yr) | -0.054 | 0.065 | | -0.043* | 0.022 | | -0.028 | 0.024 | |
| Total driving $experience^2$ | 0.000 | 0.003 | | 0.002** | 0.001 | | 0.001* | 0.001 | |
| Driving activity during winter month (1=yes) | | | | 0.035 | 0.091 | 3.6 | -0.026 | 0.095 | -2.6 |
| Driving activity during peak season month (1=yes) | -0.835*** | 0.311 | -56.6 | | | | | | |
| Predicted separation probability | | | | | | | 0.258** | 0.129 | 29.4 |
| Interaction Base pay by time observed | | | | 0.013*** | 0.002 | | 0.012*** | 0.002 | |
| Interaction % pay increase by time observed | | | | 0.340*** | 0.080 | | 0.316*** | 0.081 | |
| Split population parameter ($c$) | 0.724*** | 0.034 | | | | | | | |
| Baseline hazard-only log-likelihood [(d)] | -1,230.0 | | | -8,259.1 | | | -8,259.1 | | |
| Log likelihood at convergence | -1,196.7 | | | -3,842.1 | | | -3,840.0 | | |
| Driver-months | 37,832 | | | 38,737 | | | 38,737 | | |
| Drivers | 2,643 | | | 2,715 | | | 2,715 | | |
| P> $X^2$: LR test of $c$ = 0 ($X^2$) | 0.001 | | | | | | | | |
| P> $X^2$: LR test of model 3 vs. model 2 ($X^2$) | | | | | | | 0.044 | | |

*** Significant at a 99% confidence level
** Significant at a 95% confidence level
* Significant at a 90% confidence level
[(a)] Split population survival model. For model details see Schmidt (64).
[(b)] Proportional crash probability model. For model details see Meyer (57).

[c] Percent change in predicted probability given a discrete unit change in the independent variable at the mean, calculated as (exp(b) - 1). Factor changes for Age and Total driving experience are omitted because they are specified as having a non-linear relationship with the dependent variable. The relationship between these two sets of variables and crash probability is depicted in Figure 1.

[d] Log-likelihood for split population model is the first iteration's likelihood. The model with the baseline-hazard only could not be estimated.

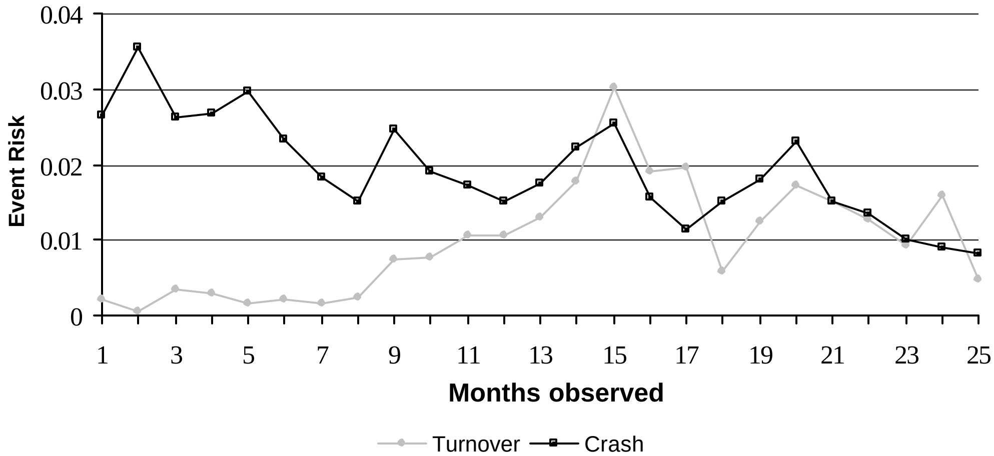



**Figure 1. Empirical crash and empirical separation hazards for JB Hunt drivers who received a pay increase**

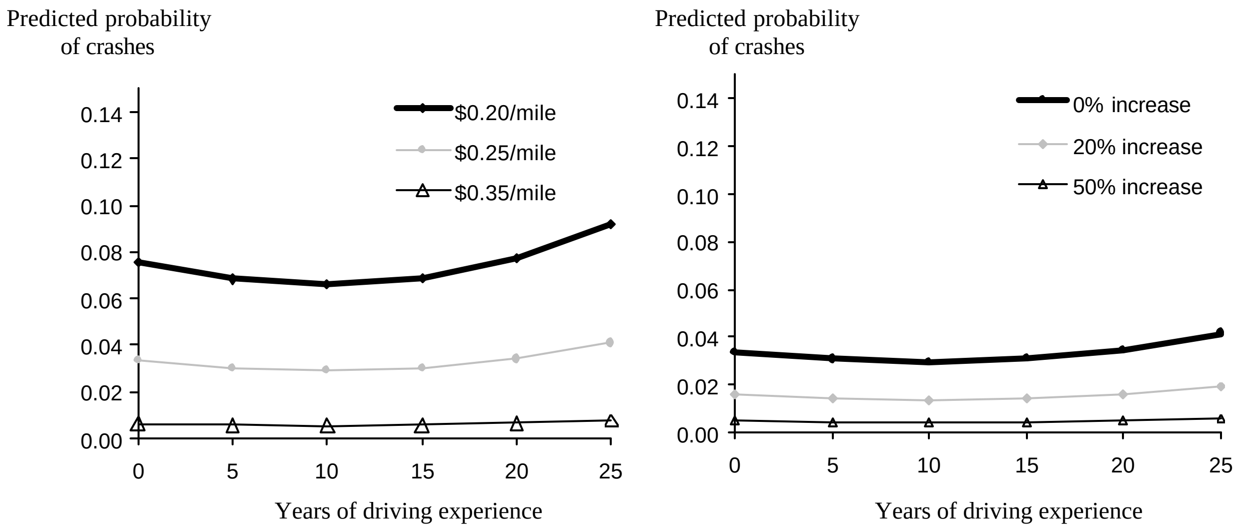



**Figure 2. Predicted probability of crashes, driving experience, and pay rate (left panel) and percent pay increase (right panel)**