AN EMPIRICAL ANALYSIS OF FATALITY RATES FOR LARGE TRUCK INVOLVED CRASHES ON INTERSTATE HIGHWAYS

Mouyid Bin Islam
Research Assistant, Department of Civil Engineering,University of Texas at El Paso
El Paso, TX, USA, email: mouyidbin.islam@gmail.com.

Salvador Hernandez, PhD*
Assistant Professor, Department of Civil Engineering, University of Texas at El Paso
El Paso, TX, USA, email: shernandez43@utep.edu.

Submitted to the 3rd International Conference on Road Safety and Simulation,
September 14-16, 2011, Indianapolis, USA

**Corresponding author*

ABSTRACT

Few studies have analyzed the impacts of freight movements (large truck) on crash rates. This study explores a novel application of a method to large truck movements, namely the random parameters tobit regression model, by examining crash rates (instead of frequencies) in truck-miles traveled and ton-miles of freight in the US as continuous censored variables. Using a nationwide crash database, the empirical results illustrate that the random-parameters tobit regression model provides an increase understanding of the factors determining large truck crash rates.

Keywords: large trucks, freight transportation, vehicle-mile travelled, ton-mileage, random parameters tobit model.

INTRODUCTION

As the national economy continues to recover, the volume of large trucks (i.e., having a gross vehicle weight rating of more than 10,000 pounds) present on the nation's highway system will also experience slow but consistent growth. This increased growth in large truck volume poses many challenges for transportation organizations that operate, maintain, and construct the transportation system. One example is the presence of increased safety hazards due to large trucks on highways—that is, the dangers associated with large trucks when mixed with passenger vehicles (Douglas, 2003). Recent statistical data have shown that large trucks have been responsible for more fatalities in the United States (US) than passenger vehicles based on the number of registered vehicles and vehicle-miles traveled (VMT) (FHWA, 2010; NHTSA, 2008). For example, large trucks accounted for roughly four percent of registered vehicles and about eight percent of VMT in 2008, but eleven percent of motor vehicle involved crash deaths in 2008 were due to large trucks (FHWA, 2010). To further illustrate the gravity of large truck involved crashes, Figure 1 shows the number of passenger vehicles and large trucks involved in fatal crashes over the period from 1999 to 2008. As seen from Figure 1, large truck involved crashes on average lead to more fatalities compared to passenger vehicles per 100 million VMT. Although the trend slopes downwards (possibly due to advancements in safety technologies and some combination of increased fuel prices and economic factors), the numbers are still concerning especially given the percentage of trucks on the nation's highways.

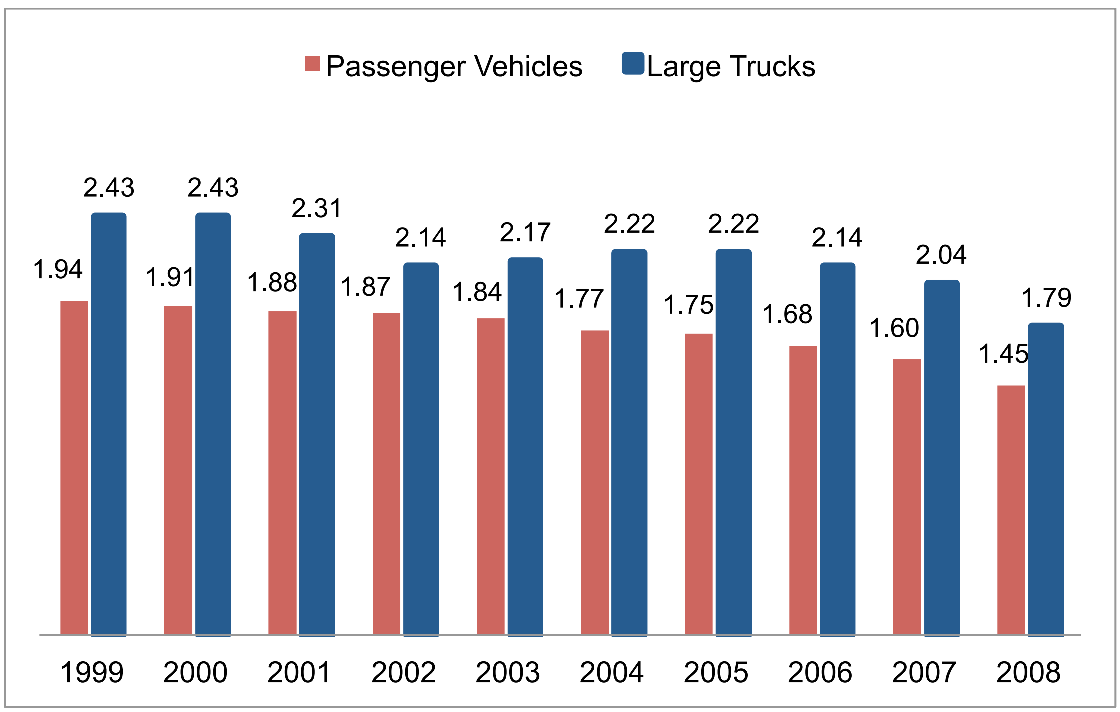

Figure 1: Vehicles involved in Fatal Crashes per 100 million VMT (FHWA, 2010)

While fatalities are a major aftermath of large truck involved crashes, the societal effects and cost associated with the resulting crashes are remarkably high—for example, expenses related to loss of life, medical attention, and insurance, and short term and long term physical and emotional effects (Miller, 1993). Moreover, large truck involved crashes greatly influence the level of injury severity experienced by those involved (Chang and Mannering, 1999). As such, these types of crashes are gnarring increased public and media attention as well increased interest from academia, transportation safety professionals, and the trucking industry. Consequently, large trucks drive the national economy through daily freight movements and would not be going away anytime soon.

To better understand the safety impacts related to increased large truck traffic on the nation's highway system, tools need to be developed that can aid transportation safety professionals as well as trucking industry operations managers in the avoidance and mitigation (i.e., aid them in the development of countermeasures) of large truck involved crashes. With this in mind, our study aims to add to the current literature by proposing a methodological approach that takes into account fatalities per million truck-miles traveled and fatalities per ton-miles of freight for large truck involved crashes. This is done through the application of a random parameters tobit modeling (censored at zero) framework. Through this, we seek to shed light on possible contributing factors to large truck involved crashes.

Over the last two decades, crash frequency modeling approaches have been widely used in traffic safety analysis. The most frequently applied models in this regard have been the Negative Binomial and Poisson models (Shankar et al., 1995; Poch and Mannering, 1996; Abdel-Aty and Radwan, 2000; Savolainen and Tarko, 2005) and their variants the zero-inflated Poisson and zero-inflated Negative Binomial models (Shankar et al., 1997; Carson and Mannering, 2001; Lee and Mannering, 2002), random parameter Negative Binomial models (Shankar et al., 1998; Chin and Quddus, 2003; Anastasopoulos and Mannering, 2009), Markov switching of two different state of crash occurrence (Malyshkina and Mannering, 2009) and Bayesian statistics on Negative Binomial models (Park et al., 2010). Although literature in crash frequency modeling is rich, severe crash rates in terms of number of crashes per VMT has not been widely studied. Specifically, literature pertaining to the modeling of fatalities per million truck-miles traveled or fatalities for ton-miles with respect to freight movements is relatively sparse. Using Exposure-based crashes such as crashes per 100 million VMT instead of traditional crash frequency as the dependent variable carries more practical significance since crash rates are widely used in crash reporting (Anastasopoulos et al., 2008).

Trucking is important to the national economy, but it also presents a significant safety concern (Zhu and Srinivasan, 2011). In the 2007 Commodity Flow Survey trucks accounted for 70.7 percent of all freight movement, 68.8 percent by weight, and 39.8 percent by ton-miles of freight (USDOT/BTS, 2008). Zhu and Srinivasan (2011) illustrate that the unique operating characteristics, driving behavior and skills, design-weight related issues for trucks as a mode for freight movements significantly impacted the frequency of crashes, and severity of injuries sustained. This is further illustrated from the fact that 413,000 large trucks were involved in traffic crashes resulting in 4,808 fatalities, accounting for 12 percent of the total fatality of all crashes in 2007 (NHTSA, 2008).

In summary, the objective of this study is then to seek those factors related to human (i.e., drivers and passengers), vehicle and road-environment and weather that influence fatalities rates as the highest level of injury severity for large truck involved crashes using a random parameters tobit modeling framework to account for heterogeneity (Tobit model applications to transportation problems have primarily assumed fixed parameter estimates see Weiss, 1992; Talley, 1995; Nolan, 2002; Anastasopoulos et al., 2008). The number of fatalities per million truck-miles traveled and number of fatalities per ton-miles for large truck freight movements is considered as a continuous variable instead of discrete integer (non-negative count) over a period of time. Since there is a likelihood of zero fatalities per million truck-miles traveled or zero fatalities per ton-miles of freight, this research is focused on fatalities higher than zero as a rate of safety indicator over a time period on US interstates, where the random parameters tobit modeling framework provides the flexibility of censoring the irrelevant count process in the regression estimation and at the same time account for unobserved factors that may vary across observations. To best of the authors' knowledge, these are the first attempts to model fatalities per million truck-miles traveled and number of fatalities per ton-miles for large truck freight movements utilizing a random parameters tobit modeling framework.

METHODOLOGY

To achieve a better understanding of the causal factors associated to larger tuck involved crashes, we seek to develop a statistical model that can be used to determine those influencing factors that affect the fatalities per million truck-miles traveled and fatalities per ton-miles for large truck freight movements using a tobit modeling framework first introduced by James Tobin (1958).
The standard tobit model (i.e., fixed parameters) is a statistical model in which the range of the response variable is constrained in some way (i.e., censored). Censoring occurs when data on the response variable are limited (or lost) and can result in data clustering at either upper or lower thresholds. In contrast to truncated data, censored data provides information on non-limited values not considered in the former—that is, in censored data all the observations are included in the dataset.

For this work, the standard tobit model is then expressed (for large truck involved in crash *i*) using a lower limit of zero (i.e., censored at zero) which is regarded the condition in our analysis for zero fatalities per million truck-miles traveled and zero fatalities per ton-miles of freight as (Washington et al, 2011):

$$
\begin{aligned}
&Y_i^* = \boldsymbol{\beta} \boldsymbol{X}_i + \varepsilon_i, \quad i = 1,2,\dots,N \\
&Y_i = Y_i^* \quad \text{if} \quad Y_i^* > 0 \\
&Y_i = 0 \quad \text{if} \quad Y_i^* \le 0
\end{aligned}
\tag{1}
$$

where:

Y_i : is the dependent variable (fatalities per million truck-miles traveled or fatalities per ton-miles of freight),

$\boldsymbol{X}_i$: is a vector of independent variables (e.g., human, roadway segment, vehicle, and crash mechanism characteristics),

$\boldsymbol{\beta}$: is a vector of estimable parameters,

N : is the number of observations in the sample used in the model, and
ε_i : is normally and independently distributed error term with zero mean and constant variance σ^2.

However, to account for heterogeneity (unobserved factors that may vary across observations), Greene (2007) has developed estimation procedures (simulation based maximum likelihood estimation) for incorporating random parameters in tobit (censored regression) models (see Moeltner and Layton, 2002 for power outage costs application). To allow for such random parameters in tobit models, estimable parameters can be written as

$$\beta_i = \beta + \gamma_i \tag{2}$$

where:
γ_i : is randomly distributed term (for example a normally distributed term with mean 0 and variance σ^2)

With this equation, the tobit model for large truck involved in crash *i* becomes $Y_i^*|\gamma_i = \boldsymbol{\beta X}_i + \varepsilon_i$. The corresponding log-likelihood can be written as

$$LL = \sum_{\forall i} ln \int_{\gamma_i} \mathrm{g}(\gamma_i) P(Y_i^*|\gamma_i) d\gamma_i \tag{3}$$

where:
$\mathrm{g}(\cdot)$: is the probability density function of the γ_i, and
$P(\cdot)$: is the probability for the tobit model.

Maximum likelihood estimation of the tobit model shown in Eq. (3) is undertaken with simulation approaches due to the difficulty in computing the probabilities. The most widely accepted simulation approach uses Halton draws which is a technique developed by Halton (1960) to generate a systematic non-random sequence of numbers. Halton draws have been shown to provide a more efficient distribution of the draws for numerical integration than purely random draws (Bhat, 2003; Train, 1999).

For estimation procedures of the standard tobit model and marginal effects derivations the reader is referred to Amemiya (1973, 1985), McDonald and Moffitt (1980), Roncek, (1992), and Anastasopoulos et al. (2008).

EMPIRICAL SETTING

To illustrate the application of the fixed- and random-parameters tobit models, crash data were collected from the Fatality Analysis Reporting Systems (FARS) from 2005 to 2008. FARS is a nation-wide crash census system where a set of files have been built documenting all qualifying fatal crashes that occurred within all the states in the U.S. The observation in the model is a fatal crash (A variable – *Fatals* includes the total number of fatalities in a fatal collision reported in the FARS database system) involving a motor vehicle where at least a large truck is involved in the fatal collision traveling on U.S. interstate system resulting in a fatal (or fatalities) within 30 days for the collision. Annual average daily traffic (AADT) is not considered in this study.

The ton-miles of freight data from 2005 to 2007 were collected from the Bureau of Transportation Statistics special tabulation (BTS/RITA, 2010), whereas, truck-miles traveled data from 2005 to 2008 were collected from FHWA travel reports (FHWA, 2009) and secondary estimation procedures includes use of State supplied data. Since the crash data were limited to the U.S. interstate system, data for the truck-miles traveled and ton-miles of freight models are limited to the U.S. interstate system.

For model estimation, the truck-miles traveled and ton-miles of freight were aggregated for the range of years of 2005 to 2008 and 2005 to 2007, respectively. Then, fatalities per million truck-miles traveled and fatalities per ton-miles of freight were calculated as follows:

$$Fatality\ Rate = \left[\frac{Number\ of\ fatalities}{truck-miles\ traveled}\right] * 1{,}000{,}000 \tag{4}$$

$$Fatality\ Rate = \left[\frac{Number\ of\ fatalities}{ton-miles\ of\ freight * 1{,}000{,}000}\right] * 1{,}000{,}000 \tag{5}$$

The total number of observations for fatalities per million truck-miles and fatality per ton-miles of freight are 3498 and 2714, respectively. The crash data were processed using the statistical software SAS. The LIMDEP software was utilized to estimate the fixed- and random-parameter tobit models. Table 1 illustrates descriptive statistics for key variables.

Table 1 Descriptive statistics of key variables

Variables	Fatalities per million truck-miles traveled		Fatalities per ton-miles of freight	
	Mean	Std. Dev.	Mean	Std. Dev.
Fatalities per million truck-miles traveled	5.676	3.261	-	-
Fatalities per ton-miles of freight	-	-	0.985	0.58
Manner of collision (1 if rear-end, 0 otherwise)	0.367	0.482	-	-
Manner of collision (1 if angle, 0 otherwise)	0.065	0.246	0.069	0.253
Ambient light condition (1 if dawn time, 0 otherwise)	0.027	0.163	0.027	0.163
Surface condition (1 if wet, 0 otherwise)	0.134	0.341	-	-
Weather condition (1 if foggy, 0 otherwise)	-	-	0.017	0.129
Weather condition (1 if rainy, 0 otherwise)	0.093	0.289	-	-
Traffic median barrier (1 if divided highway with traffic barrier, 0 otherwise)	0.291	0.454	0.28	0.449
Time of the day (1 if 5 pm in the evening, 0 otherwise)	0.041	0.198	0.043	0.203
Time of the day (1 if 6 pm in the evening, 0 otherwise)	0.034	0.183	0.034	0.181
Trailing unit (1 if two trailing unit, 0 otherwise)	0.051	0.219	0.052	0.222
State specific crash information (1 if Texas, 0 otherwise)	0.091	0.287	0.093	0.29
Month of the year (1 if month is August, 0 otherwise)	0.083	0.276	-	-
Month of the year (1 if month is December, 0 otherwise)	-	-	0.076	0.265
Day of the weekend (1 if Friday, 0 otherwise)	-	-	0.16	0.367
Crash related human factors (1 if driving too fast, 0 otherwise)	0.068	0.252	0.079	0.271
Driver's license type (1 if license is valid, 0 otherwise)	0.863	0.343	-	-
Involved vehicles in crash	1.843	1.369	1.81	1.195
Number of person not fatally insured	2.802	4.206	2.739	4.152

EMPIRICAL RESULTS

Table 2 and Table 3 present estimation results for the tobit fixed- and random-parameters models for fatalities per million truck-miles traveled and fatalities per ton-miles of freight, respectively. The random parameters tobit models were estimated using simulation-based maximum likelihood with 200 Halton draws. This number of draws has been empirically shown to produce accurate parameter estimates (Bhat, 2003; Milton et al., 2008; Gkritza and Mannering, 2008). With regard to the distribution of the tobit random parameters, consideration was given to the normal, lognormal (which restricts the impact of the parameters to be either negative or positive), triangular, and uniform distributions. However, only the normal distribution was found to be significant. The estimation results in Tables 2 and 3 show the estimated parameters with their respective statistical significance (*t-stat* and *P-value*) and plausible sign based on the sample sizes of 3498 (fatalities per million truck-miles traveled) and 2714 (fatalities per ton-miles of freight) of crash observations that had complete information of all variables used.

The Madalla pseudo R^2 was estimated for both the fixed- and random-parameter tobit models (see Tables 2 and 3) (Madalla, 1983). Veall and Zimmermann (1996) show that the Madalla pseudo R^2 is good indicator of overall goodness of fit and is computed as (also see Anastasopoulos et al., 2008)

$$Madalla\ pseudo\ R^2 = 1 - e^{[-2(LL(\beta)-LL(0))/N]} \tag{6}$$

where:
$LL(\beta)$: is log-likelihood at convergence,
$LL(0)$: is log-likelihood at zero, and
N : is the number of observations.

For the fatalities per million truck-miles traveled model, the pseudo R^2 were found to be 0.227 and 0.355 for the fixed- and random-parameter tobit models, respectively. Similarly, for the fatalities per ton-miles of freight model, the pseudo R^2 were found to be 0.222 and 0.360 for the fixed and random parameter tobit models, respectively. The pseudo R^2 for the tobit models indicate that the random parameter tobit models are more robust in explaining unobserved heterogeneity than fixed parameter tobit models. Furthermore, a likelihood ratio test comparing the fixed- and random-parameters models for the fatalities per million truck-miles traveled (χ^2 = 629.51) and fatalities per ton-miles of freight (χ^2 = 529.54) indicates that we are more than 99.99% (a *p-value* near zero) for both models (see Washington et al., 2011). Therefore, the interpretation of the estimation of results will be confined to both the fatalities per million truck-miles traveled and fatalities per ton-miles of freight random parameter tobit models.

To assess the degree of influence of specific variables, Table 4 illustrates the computed marginal effects for the fatalities per million truck-miles traveled and fatalities per ton-miles of freight for the random parameter tobit models, respectively. Finding the marginal effect of an independent variable on the expected value of a dependent variable for all cases, $E[Y]$, was calculated using the McDonald and Moffitt (1980) formula:

$$\partial E[Y]/(\partial X_i) = F(z)\times(\partial E[Y^*]/(\partial X_i)) + E[Y^*]\times(\partial F(z)/(\partial X_i)) \tag{7}$$

where:
$F(z)$: is the cumulative normal distribution function, associated with the proportion of cases above the limit (in this case zero),
$E[Y^*]$: denotes observations above zero which indicates fatalities per million VMT and fatalities per ton-miles of freight (not censored),
$\frac{\partial E[Y^*]}{\partial X_i}$: denotes observations above zero which indicates fatalities per million VMT and fatalities per ton-miles of freight (not censored),
$\frac{\partial F(z)}{\partial X_i}$: is the change in the cumulative probability of being above zero associated with an independent variable.

Table 2 Tobit regression estimation for fatalities per million truck-miles traveled

Variables	Fixed Parameter Tobit			Random Parameter Tobit		
	Coeff.	*t-stat*	*P-value*	*Coeff.*	*t-stat*	*P-value*
Constant	3.430	19.263	0.000	3.916	19.435	0.000
Crash Mechanism						
Manner of collision (1 if rear-end, 0 otherwise)	-0.186	-1.744	0.081	-0.229	-2.122	0.034
Manner of collision (1 if angle, 0 otherwise)	0.591	2.901	0.004	0.725	4.961	0.000
Temporal Characteristics						
Ambient light condition (1 if dawn time, 0 otherwise)	0.598	1.995	0.046	0.920	4.175	0.000
Time of the day (1 if 5 pm in the evening, 0 otherwise)	0.756	3.063	0.002	0.898	5.046	0.000
Time of the day (1 if 6 pm in the evening, 0 otherwise)	0.809	3.040	0.002	0.806	4.574	0.000
Month of the year (1 if month is August, 0 otherwise)	0.416	2.362	0.018	0.432	3.131	0.002
Location Characteristics						
State specific crash information (1 if Texas, 0 otherwise)	0.501	2.950	0.003	0.361	2.504	0.012
Environment - Weather						
Weather condition (1 if rainy, 0 otherwise)	-0.834	-3.266	0.001	-0.741	-3.275	0.001
Road Surface condition (1 if wet, 0 otherwise)	0.565	2.606	0.009	0.478	2.609	0.009
Road - Geometry						
Traffic median barrier (1 if divided highway with traffic barrier, 0 otherwise)	-0.234	-2.182	0.029	-0.233	-2.007	0.045
Vehicle Configuration						
Trailing unit (1 if two trailing unit, 0 otherwise)	-0.517	-2.334	0.019	-0.488	-1.881	0.060
Human Factor						
Vehicle maneuver (1 if going straight, 0 otherwise)	0.398	3.527	0.000	0.174	1.508	0.132*
Crash related human factors (1 if driving too fast, 0 otherwise)	1.167	5.970	0.000	0.741	4.897	0.000
Driver's license type (1 if license is valid, 0 otherwise)	0.399	2.813	0.005	0.332	2.003	0.045
Exposure to Injury Severity						
Number of vehicles involved in the crash	0.438	10.211	0.000	0.293	8.887	0.000
Std. dev. of parameter distribution				*0.409*	*36.118*	*0.000*
Number of persons not fatally injured in the crash	0.246	17.328	0.000	0.236	19.659	0.000
Std. dev. of parameter distribution				*0.247*	*32.781*	*0.000*
Number of variables	17			17		
Log-likelihood at zero, $LL(\mathbf{0})$	-9097.403			-9097.403		
Log-likelihood at convergence, $LL(\boldsymbol{\beta})$	-8646.047			-8331.289		
$X^2 = -2[LL(\mathbf{0}) - LL(\boldsymbol{\beta})]$	902.71			1532.228		
Number of observations	3498			3498		
Madalla pseudo-R^2	0.227			0.355		

**the p-value is considered upto 0.15 indicating that we are 85% confident that coefficient estimates are significantly different from zero.*

Table 3 Tobit regression estimation for fatalities per ton-miles of freight

	Fixed Parameter Tobit			Random Parameter Tobit		
Variables	*Coeff.*	*t-stat*	*P-value*	*Coeff.*	*t-stat*	*P-value*
Constant	0.679	32.485	0.000	0.696	33.487	0.000
Std. dev. of parameter distribution	-	-	-	*0.025*	*3.096*	*0.002*
Crash Mechanism Manner of collision (1 if angle, 0 otherwise)	0.136	3.469	0.000	0.135	5.038	0.000
Temporal Characteristics						
Ambient light condition (1 if dawn time, 0 otherwise)	0.110	1.916	0.055	0.108	3.146	0.002
Time of the day (1 if 5 pm in the evening, 0 otherwise)	0.147	3.025	0.003	0.137	4.205	0.000
Time of the day (1 if 6 pm in the evening, 0 otherwise)	0.185	3.391	0.001	0.164	5.232	0.000
Day of the week (1 if Friday, 0 otherwise)	0.062	2.312	0.021	0.077	3.117	0.002
Month of the year (1 if month is December, 0 otherwise)	-0.100	-2.686	0.007	-0.062	-1.775	0.076
Location Characteristics						
State specific crash information (1 if Texas, 0 otherwise)	0.109	3.194	0.001	0.092	3.226	0.001
Environment - Weather Weather condition (1 if foggy, 0 otherwise)	-0.267	-3.471	0.000	-0.214	-2.488	0.013
Road - Geometry						
Traffic median barrier (1 if divided highway with traffic barrier, 0 otherwise)	-0.052	-2.350	0.019	-0.036	-1.571	0.116*
Vehicle Configuration						
Trailing unit (1 if two trailing unit, 0 otherwise)	-0.102	-2.309	0.021	-0.082	-1.627	0.104*
Human Factor						
Crash related human factors (1 if driving too fast, 0 otherwise)	0.223	6.041	0.000	0.173	6.366	0.000
Exposure to Injury Severity Number of vehicles involved in the crash	0.087	8.923	0.000	0.075	11.489	0.000
Std. dev. of parameter distribution	-	-	-	*0.068*	*27.787*	*0.000*
Number of persons not fatally injured in the crash	0.043	15.117	0.000	0.039	17.377	0.000
Std. dev. of parameter distribution	-	-	-	*0.055*	*35.470*	*0.000*
Number of variables	14			14		
Log-likelihood at zero, $LL(\mathbf{0})$	-2373.439			-2373.439		
Log-likelihood at convergence, $LL(\boldsymbol{\beta})$	-2032.271			-1767.500		
$X^2 = -2[LL(\mathbf{0}) - LL(\boldsymbol{\beta})]$	682.336			1211.878		
Number of observations	2714			2714		
Madalla pseudo-R^2	0.222			0.360		

****the p-value is considered upto 0.15 indicating that we are 85% confident that coefficient estimates are significantly different from zero.***

Table 4 Marginal effects comparison for fixed- and random-parameter tobit models for fatalities per million truck-miles traveled and fatalities per ton-miles of freight

Variables	Fatalities per million truck-miles traveled		Fatalities per ton-miles of freight	
	Random	Fixed	Random	Fixed
Constant	3.869	3.349	0.688	0.661
Manner of collision (1 if rear-end, 0 otherwise)	-0.226	-0.182	-	-
Manner of collision (1 if angle, 0 otherwise)	0.716	0.576	0.133	0.132
Ambient light condition (1 if dawn time, 0 otherwise)	0.909	0.583	0.107	0.108
Surface condition (1 if wet, 0 otherwise)	0.472	0.551	-	-
Weather condition (1 if foggy, 0 otherwise)	-	-	-0.211	-0.26
Weather condition (1 if rainy, 0 otherwise)	-0.732	-0.815	-	-
Traffic median barrier (1 if divided highway with traffic barrier, 0 otherwise)	-0.23	-0.229	-0.036	-0.05
Time of the day (1 if 5 pm in the evening, 0 otherwise)	0.888	0.738	0.136	0.143
Time of the day (1 if 6 pm in the evening, 0 otherwise)	0.797	0.789	0.162	0.18
Trailing unit (1 if two trailing unit, 0 otherwise)	-0.482	-0.505	-0.081	-0.099
Vehicle maneuver (1 if going straight, 0 otherwise)	0.172	0.388	-	-
State specific crash information (1 if Texas, 0 otherwise)	0.356	0.489	0.091	0.106
Month of the year (1 if month is August, 0 otherwise)	0.426	0.406	-	-
Month of the year (1 if month is December, 0 otherwise)	-	-	-0.062	-0.097
Day of the weekend (1 if Friday, 0 otherwise)	-	-	0.076	0.061
Crash related human factors (1 if driving too fast, 0 otherwise)	0.732	1.139	0.171	0.217
Driver's license type (1 if license is valid, 0 otherwise)	0.328	0.39	-	-
Number of vehicles involved in the crash	0.29	0.428	0.074	0.006
Number of persons not fatally injured in the crash	0.233	0.24	0.039	0.084

Fatalities per Million Truck-miles Traveled Model

Two parameters were found to be random with statistically significant standard deviations for their assumed distributions. Also, for the parameters whose standard deviations were not statistically different from zero, the parameters were fixed to be constant across the observations. The estimation results shown in Table 2 indicate that the number of vehicles involved in the crash, and the number of persons not fatally injured in the crash were found to produce statistically significant random parameters.

With regard to the parameters found to be random, the exposure to injury severity variable the more vehicles involved in a crash resulted in a random parameter that is normally distributed, with mean of 0.293 and standard deviation of 0.409. The positive sign indicates that an increase in number of vehicles involved in a crash per million trucks-mile traveled increases the likelihood of fatalities (less than 23.7 percent of the distribution would have a negative value). On possible explanation for this finding is that crashes with many cars (e.g., pile ups) varies in severity (may not always lead to fatalities) due to some unforeseen pile up dynamics and preventive technologies present in vehicles (Chakravarthy et al., 2009). With respect to marginal effects, Table 4 shows that a unit increase in the number of vehicles involved in the crash results in an average 0.29 increase in the number of fatalities per million truck-miles traveled. This variable was also found to be significant by Chen and Chen (2011) for multi-vehicle collisions.

Similarly, the exposure to injury severity variable for the number of persons not fatally injured in the crash was also found to be random and normally distributed, with mean of 0.236 and standard deviation of 0.247. Given the distributional patterns, an increase in the number of persons not fatally injured in a crash increases fatalities per million truck-miles traveled but with varying magnitude—that is, less than 16.9 percent of the distribution (less than zero) would have a negative value (would increase fatalities). A possible reason for this finding may be due to under reporting by police because persons dying sometime later due to injuries sustained during the crash may not be updated later on the police reports themselves. Marginal effects show that a unit increase in persons not injured in the crash results in an average 0.23 increase in the number of fatalities per million truck-miles traveled. More broadly, Islam and Mannering (2006) also indicate that the likelihood of fatality increases when one or more occupants travel with the driver.

The indicator variable representing rear-end collisions decreases fatalities per million truck-miles traveled. This may be due to most occupants being in the front seats of their vehicle (trucks) and are afforded more full body protection from the rear seats (trailers) and head restraints (airbags) upon collision. In addition, the direction of the impact and the resulting relative movement of the occupants minimizes' the chance of more serious injuries of striking more lethal objects in the vehicle (Duncan et al., 1998). The average marginal effect for this variable is -0.226 (and only a decrease of -0.182 for the fixed parameter model)

The angle collision indicator variable increases fatalities per million truck-miles traveled. In contrast to rear-end collisions, angled collisions lead to more severe injury outcomes (e.g., fatalities) especially when large trucks are involved. This may be due to the structural dynamic makeup of vehicles especially when struck in an angle—not as energy absorbing as the front or rear of vehicles (Abdel-Aty and Abdelwahad, 2004). Marginal effect for this variable is 0.716 compared to 0.576 for the fixed parameter model.

With regards to the temporal variables, all the indicator variables increase fatalities per million truck-miles traveled. First, the dawn variable (before the sunrise) increases the likelihood of fatalities. This may be a result of driver experiencing drowsiness and maybe capturing, among other factors, the effects related to long hours of driving. Next, the times from 5 to 6 pm increases fatalities per million truck-miles traveled. This maybe also capturing some driver related factors (as in the dawn variable) with regards to the level of alertness and fatigue. During the summer periods, in particular, August increases fatalities per million truck-miles traveled. This may be reflecting vehicular interactions on highways due to preferable weather condition for outdoor activities especially during this time of year. A marginal effect of 0.426 for the August indicator variable is observed for the random parameters tobit model compared to 0.402 for the fixed parameter tobit model.

Crashes occurring in the state of Texas indicator variable increases fatalities per million truck-miles traveled. It is interesting that this variable was found to increase the fatalities per million truck-miles traveled, a possible explanation may be the number truck related freight movements in the State of Texas due to it sharing a border with North American Free Trade Agreement (NAFTA) member Mexico. This variable may be capturing the driving complexities related to the diverse geographical nature of the State of Texas.

With respect to weather, the indicator variable for rain was found to be significant and decreased the fatalities per million truck-miles traveled. A possible explanation is that truck drivers are more cautious while driving through rain. This result is supported by Zhu and Srinivasan (2011) and Chen and Chen (2011) based on the risk-averse behavior of the drivers in the adverse weather conditions. On one hand, the indicator variable for surface condition being wet increases fatalities per truck-miles traveled. This is possibly capturing, among other factors, vehicular conditions (e.g., tire wear leading to hydroplaning). Chen and Chen (2011) also show for wet surface conditions due to snow/slush, increases the likelihood of collisions.

The presences of median barriers (or not) separating the opposing traffic flow decreases fatalities per million truck-miles traveled. As shown in Anastasopoulos et al. (2008) median barrier potentially reduces head-on collisions and may lower injury severity, which significantly reduces the likelihood of fatalities.

The indicator variable for a truck hauling two trailers decreases the likelihood of fatalities per million truck-miles traveled. A possible reason is that these large trucks are primarily driven by professional truck drivers with practical safety training especially for hauling more than one trailer unit.

Driving the truck in the straight in a traffic lane as a crash avoiding maneuver (or not) increases fatalities per million truck-miles traveled. This may be due to, among other factors, the kinematics revolving around large truck involved crashes. Akin, driving too fast was identified as increasing the fatalities per million truck-miles traveled. Speed (being the top factor identified in the FARS data) has been shown to increased fatalities rates due to due higher energy transfer between colliding bodies (Craft, 2010).

The indicator variable for a truck driver who poses a valid license (or not) increases fatalities per million truck-miles traveled. This variable may be capturing factors related to the level of experience or years of driving.

Fatalities per Ton-miles of Freight Model

To avoid repetition in the explanation of the specified estimates found in the two models, only variables specific to the fatalities per ton-miles of freight will be explained in this section. Turning to the model specification, three parameters were found to be random with statistically significant standard deviations for their assumed distributions. Also, for the parameters whose standard deviations were not statistically different from zero, the parameters were fixed to be constant across the observations. The estimation results shown in Table 3 indicate that the constant, the number of vehicles involved in the crash, and the number of persons not fatally injured in the crash were found to produce statistically significant random parameters.

The constant for fatalities per ton-miles of freight is found to be random and normally distributed with mean of 0.696 and standard deviation of 0.025. With these distributional patterns, the constant term is less than zero for 0% and more than zero for 100% of the large truck involved fatalities per ton-miles of freight. This variability is likely capturing the unobserved heterogeneity in the severity outcomes that could include factors such as traffic condition, among other factors, which was not directly measured in the dataset for this model.

With regards to the significant temporal variables, the indicator variable for December was found to be significant and decreased fatalities per ton-miles of freight. The significance of this variable may stem from the lower activity of freight movements due to winter (the possibility of adverse weather conditions such as snow), and seasonal effects (e.g., Christmas holidays). Typically, freight movements are at their highest in the early fall for the winter holiday season. In addition, the day of the week the Friday indicator variable increases fatalities per-ton miles. Although freight movements are made pretty uniformly from Monday thru Friday, this variable may be capturing, among other factors, some week-end effects.

Consistent to Zhu and Srinivasan (2011) we find that the presence of foggy weather conditions has a negative effect on fatalities per ton-miles of freight. As was the finding with the rain indicator variable earlier, truck drivers are more cautious while driving through foggy conditions. Additionally, this variable may be capturing some risk-averse behavior of drivers.

SUMMARY AND CONCLUSIONS

This study provides a demonstration of the random parameters tobit regression as a viable methodological approach to gain new insights into factors that significantly influence fatalities per million truck-miles traveled and fatalities per ton-miles of freight. The random-parameters tobit regression modeling framework is an important approach because it allows us to account and correct for heterogeneity that can arise from factors such as human (i.e., drivers and passengers), vehicle, road-environment, weather, variations in police reporting, temporal and other unobserved factors not captured.

Using four years of data for fatalities per million truck-miles traveled and three years of data for fatalities per ton-miles of freight our estimation results provide some interesting findings, respectively. For example, factors related to the type of collision were found to be significant including rear-end and angled crashes as was driving too fast. Temporal factors were also found to be significant such as the effects of dawn, evening times between 5 and 6 pm, and the months of August and December. In terms of locational variables the State of Texas was found to be a contributing factor for both models. Also, factors related to weather which included rain, foggy, and wet surfaces were significant. With regards to road geometry, the presences of traffic medians impacted both models. The hauling of two trailers by a truck was also found to be significant for both models. And, exposure variables number of vehicles involved in a crash and the number of persons not fatally injured were significant. Although traffic data such as AADT has not been incorporated in the dataset for the developed models, there are variables in both models representing the time of the day (dawn time, between 5 pm to 6 pm), day of the week (Friday) and month of the year (August, December) serve as a proxy for traffic conditions on the highway system.

Although this study is exploratory in nature, the modeling approach presented in this paper offers a flexible methodology that has considerable potential to analyze fatalities per million truck-miles traveled and fatalities per ton-miles of freight. Applying this approach to state specific datasets with available AADT (average annual daily traffic) data and for more years, would potentially provide more information on the effects of contributing factors present and new on fatalities per million truck-miles traveled and fatalities per ton-miles of freight.

REFERENCES

Abdel-Aty, M.A., and Abdelwahad, H., (2004). "Analysis and Prediction of Traffic Fatalities Resulting from Angle Collisions Including the Effect of Vehicles' Configuration and Compatibility", *Accident Analysis and Prevention* 36(3), 457–469.

Abdel-Aty, M. A., and Radwan, A. E., (2000). "Modeling traffic accident occurrence and involvement", *Accident Analysis and Prevention* 32(5), 633–642.

Amemiya, T., (1973). "Regression-analysis when Dependent Variable is Truncated Normal", *Econometrica* 41(6), 997–1016.

Amemiya, T., (1985). "*Advanced Econometrics*". Harvard University Press, Cambridge, MA.

Anastasopoulos, P. Ch., T, Andrew P., and Mannering, F. L., (2008). "Tobit Analysis of Vehicle Accident Rates on Interstate Highways", *Accident Analysis and Prevention* 40(2), 768–775

Anastasopoulos, P. and Mannering, F.L., (2009). "A Note on Modeling Vehicle Accident Frequencies with Random-Parameters Count Models", *Accident Analysis and Prevention* 41(1), 153–159.

Bhat, C., (2003). "Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences", *Transportion Research Part B*, 37(1), 837–855.

Bureau of Transportation Statistics (BTS), Research and Innovative Technology Administration (RITA), Special Tabulation. U.S. Department of Transportation. (http://www.bts.gov/publications/national_transportation_statistics/html/table_01_46b.html).

Carson, J., Mannering, F., (2001). "The Effect of Ice Warning Signs on Accident Frequencies and Severities", *Accidient Analysis and Prevention* 33(1), 99–109.

Chakravarthy A., Song, K., and Feron, E., (2009). "Preventing Automotive Pileup Crashes in Mixed-Communications Environments", IEEE Transactions on Intelligent Transportation Systems 10(2), 211–225.

Chang, Li-Yen, Mannering, F., (1999). "Analysis of Injury Severity and Vehicle Occupancy in Truck- and Non-truck involved Accidents", *Accident Analysis and Prevention* 31(5), 579–592.

Chen, F., and Chen, S., (2011). "Injury severities of truck drivers in single- and multi-vehicle accidents on rural highways", *Accident Analysis and Prevention* 43(5), 1677-1688.

Chin, H. C., and Quddus, M. A., (2003). "Applying the Random Effect Negative Binomial Model to Examine Traffic Accident Occurrence at Signalized Intersections", *Accident Analysis and Prevention* 35(2), 253–259.

Craft, R., (2010). “2009: Historic Truck Crash Declines”, webinar, Federal Motor Carrier Safety Administration, United States Department of Transportation, Washington, D.C.

Douglas, J.G., (2003). “Strategies for Managing Increasing Truck Traffic”, *NCHRP Report 314*, TRB, National Research Council, Washington, D.C.

Duncan, C.S., Khattak, A.J., and Council, F.M., (1998). “Applying the Ordered Probit Model to Injury Severity in Truck-Passenger Car Rear-End Collisions”, *Transportation Research Record: Journal of the Transportation Research Board* , No.1635, Transportation Research Board of the National Academies, Washington, DC,63–71.

FARS Analytic Reference Guide1975 to 2009. (2010). National Highway Safety Administration, DOT HS 811 352. (http://www-nrd.nhtsa.dot.gov/Pubs/811352.pdf).

Federal Highway Administration, (2010). Highway statistics, 2008. Washington, DC: U.S. Department of Transportation.

Greene,W., (2007). *Limdep*, Version 9.0. Econometric Software Inc., Plainview, NY.

Greene, W., (1999). “Marginal Effects in the Censored Regression Model”, *Economic Letters* 64, 43–49.

Gkritza, K., Mannering, F.L., (2008). “Mixed logit analysis of safety-belt use in single- and multi-occupant vehicles”, *Accident Analysis and Prevention* 40(2), 443–451.

Halton, J. (1960). “On the efficiency of evaluating certain quasi-random sequences of points in evaluating multi-dimensional integrals”, *Numerische Mathematik* 2, 84–90.

Islam, S., and Mannering, F., (2006). “Driver Aging and its Effects on Male and Female Single-Vehicle Accident Injuries: Some Additional Evidence”, Journal of Safety Research 37, 267–276.

Khorashadi, A., Niemeier, D., Shankar, V., and Mannering, F., (2005). “Differences in Rural and Urban Driver-Injury Severities in Accidents Involving Large-Trucks: An Exploratory Analysis”, *Accident Analysis and Prevention* 37(5), 910–921.

Lee, J., Mannering, F., (2002). “Impact of Roadside Features on the Frequency and Severity of Run-off-Roadway Accidents: An Empirical Analysis”, *Accident Analysis and Prevention* 34(2), 149–161.

Maddala, G.S. (1983). “*Limited-dependent and Qualitative Variables in Econometrics”*, New York: Cambridge University Press.

Malyshkina, N. V., Mannering, F. L., and Tarko A. P., (2009) “Markov Switching Negative Binomial Models: An application to Vehicle Accident Frequencies”, *Accident Analysis and Prevention* 34(2), 149–161.

McDonald, J.F., and Moffitt, R.A., (1980). “The Uses of Tobit Analysis”, *The Review and Economics Statistics* 62(2), 318–321.

Miller, T.R., (1993). “ Costs and Functional Consequences of U.S. Roadway Crashes”, *Accident Analysis and Prevention* 25(5), 593–607.

Milton, J.C., Shankar, V.N., Mannering, F.L., (2008). “Highway accident severities and th mixed logit model: An exploratory emperical analysis”, *Accident Analysis and Prevention* 40(1), 260–266.

Moeltner, K., and Layton, D.F., (2002). “A Censored Random Coefficients Model for Pooled Survey Data with Application to the Estimation of Power Outage Costs”, *The Review of Economics and Statistics* 84(3), 552–561.

NHTSA, 2008. Traffic Safety Facts 2007: Large Trucks, DOT HS 810 805. (http://www-nrd.nhtsa.dot.gov/Pubs/810989.PDF).

Park, B. J., Lord, D. and Hart, J. D., (2010) “Bias Properties of Bayesian Statistics in Finite Mixture of Negative Binomial Regression Models in Crash Data Analysis”. *Accident Analysis and Prevention* 42(2), 741–749

Poch, M., and Mannering, F., (1996). “Negative Binomial Analysis of Intersection Accident Frequencies”, *Journal of Transportation Engineering* 122(2), 105–113.

Roncek, D.W., (1992). “Learning More from Tobit Coefficients: Extending a Comparative Analysis of Political Protest”, *American Sociological Review* 57(4), 503–507.

Savolainen, P.T., and Tarko, A.P., (2005). “Safety Impacts at Intersections on Curved Segments”, *Transportation Research Record: Journal of the Transportation Research Board* , No.1908, Transportation Research Board of the National Academies, Washington, DC,130–140.

Shankar,V., Albin, R., Milton, J., and Mannering, F., (1998). “Evaluating Median Crossover Likelihoods with Clustered Accident Counts: An Empirical Inquiry using the Random Effects Negative Binomial Model”, *Transportation Research Record: Journal of the Transportation Research Board*, No. 1635, Transportation Research Board of the National Academies, Washington, DC, 44–48.

Shankar, V., Milton, J., Mannering, F., (1997). “Modeling Accident frequencies as Zero-Altered Probability Processes: An Empirical Inquiry”, *Accident Analysis and Prevention* 29(6), 829–837.

Shankar, V., Mannering, F., Barfield, W., (1995). “Effect of Roadway Geometrics and Environmental Factors on Rural Accident Frequencies”, *Accident Analysis and Prevention* 27 (3), 371–389.

Tobin, J. (1958). “Estimation of Relationships for Limited Dependent Variables”, *Econometrica* 26, 24–36.

Train, K., (1999). “Halton sequences for mixed logit”, Working Paper, University of California Berkley, Department of Economics.

USDOT/BTS, 2008. 2007 Commodity Flow Survey. (http://www.bts.gov/publications/ commodity flow survey/preliminary_tables_december 2008/index.html).

Veall, M. R., and Zimmermann, K. F. (1996). “Pseudo-R2 Measures for Some Common Limited Dependent Variable Models”, *Journal of Economic Surveys* 10(3), 241-259.

Washington, S.P., Karlaftis, M.G., and Mannering, F.L., (2011). “Statistical and Econometric methods for Transportation Data Analysis”, 2nd Ed. Chapman & Hall/CRC.

Zhu, X., and Srinivasan, S., (2011). “A comprehensive analysis of factors influencing the injury severity of large-truck crashes”, *Accident Analysis and Prevention* 43(1), 49–57.