ARTICLE IN PRESS

Accident Analysis and Prevention xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

# Injury severity analysis of commercially-licensed drivers in single-vehicle crashes: Accounting for unobserved heterogeneity and age group differences

Mohamed Osman[a,*], Sabyasachee Mishra[a], Rajesh Paleti[b]

[a] Department of Civil Engineering, University of Memphis, 3815 Central Avenue, Memphis, TN 38152, United States
[b] Department of Civil & Environmental Engineering, Old Dominion University, 135 Kaufman Hall, Norfolk, VA 23529, United States

ARTICLE INFO

*Keywords:*
Commercial driver license
Injury severity
Mixed generalized ordered response probit
Heterogeneity
Transportation safety

ABSTRACT

This study analyzes the injury severity of commercially-licensed drivers involved in single-vehicle crashes. Considering the discrete ordinal nature of injury severity data, the ordered response modeling framework was adopted. The moderating effect of driver's age on all other factors was examined by segmenting the parameters by driver's age group. Additional effects of the different drivers' age groups are taken into consideration through interaction terms. Unobserved heterogeneity of the different covariates was investigated using the Mixed Generalized Ordered Response Probit (MGORP) model. The empirical analysis was conducted using four years of the Highway Safety Information System (HSIS) data that included 6247 commercially-licensed drivers involved in single-vehicle crashes in the state of Minnesota. The MGORP model elasticity effects indicate that key factors that increase the likelihood of severe crashes for commercially-licensed drivers across all age groups include: lack of seatbelt usage, collision with a fixed object, speeding, vehicle age of 11 years or more, wind, night time, weekday, and female drivers. Also, the effects of several covariates were found to vary across different age groups.

## 1. Introduction

The U.S. Congress passed the Commercial Motor Vehicle Safety Act of 1986 to establish uniform standards for testing and licensing of operators of commercial motor vehicles (Commercial Motor Vehicle Safety Act, 1986). This Act prohibits any person from operating a Commercial Motor Vehicle (CMV) without a valid a commercial driver's license (CDL). This study aims to analyze factors contributing to the level of injury severity sustained by commercially-licensed drivers involved in single-vehicle crashes. A MGORP model was adopted to investigate potential heterogeneous effects associated with the set of explanatory variables being investigated (Eluru et al., 2008). It was essential to consider the differences among various drivers' age groups for modeling injury severity outcomes of single-vehicle crashes involving CDL holders. For example, older drivers tend to have longer reaction times and likely to be more vulnerable in a crash occurrence. In contrast, younger drivers may have less driving experience and likely to drive aggressively compared to other age groups (Lee and Mannering, 2002). In this paper, we specifically analyzed potential heterogeneous effects due to "age" on the injury severity outcomes through segmenting the variable effects by drivers' age groups.

The Federal Motor Carrier Safety Administration (FMCSA) states that drivers are required to have a CDL in order to operate certain commercial vehicles since April of 1992 (FMCSA, 2014a). FMCSA has developed standards to be adopted by the different States when issuing commercial licenses. A CDL is issued when the potential driver passes a set of knowledge and skills tests administered by the State, which directly corresponds to the specific type of vehicle a driver is seeking to operate. Three types of CDLs are classified by FMCSA (Class A, B, and C) depending on the vehicle's gross weight and the different combinations of units or trailers. According to the U.S. Department of Transportation (USDOT), in 2013, there were approximately 3.9 million registered commercially-licensed drivers operating in the U.S. (FMCSA, 2014a). Between 2009 and 2013, there were approximately 650,000 commercially-licensed drivers involved in roadway crashes, although this statistic only accounts for large trucks and buses only (FMCSA, 2014b). In 2016 alone, there were approximately 165,000 commercial crashes that involved nearly 4700 fatalities and more than 91,000 injuries (U.S. Department of Transportation, 2017). The economic impact is substantial; the FMCSA states that in 2011, commercial crashes costs equated to $87 Billion (adjusted to 2012 dollars) (Zaloshnja and Miller, 2002). However, the economic impacts of crashes involving commercially- licensed drivers are outside the scope of this study.

A review of the dataset utilized in this study reveals that CDL

* Corresponding author.
*E-mail addresses:* mosman1818@outlook.com (M. Osman), smishra3@memphis.edu (S. Mishra), rpaleti@odu.edu (R. Paleti).

https://doi.org/10.1016/j.aap.2018.05.004
Received 4 December 2017; Received in revised form 8 April 2018; Accepted 4 May 2018
0001-4575/ 

Please cite this article as: Osman, M., Accident Analysis and Prevention (2018), https://doi.org/10.1016/j.aap.2018.05.004

holders are involved in more than just crashes in commercial vehicles; to include privately owned passenger-vehicles; possibly outside of work hours. Regardless of the type of vehicle being operated, CDL holders are likely operating on the road for longer periods of times and for greater distances, which may lead to higher risks of fatigue; raising the possibility of crash occurrence compared to non-CDL holders (Park et al., 2017). CDL holders are individuals who possess a higher level of knowledge, experience, skills, and physical abilities compared to standard driver's license holders. Serious traffic violations committed by CDL holders can affect their ability to maintain their certification (FMCSA, 2014a). Due to the possible differences in behavior between both CDL and non-CDL holders and the different nature of crashes both categories may be involved in, this study aims to target all commercial and non-commercial crashes involving drivers who, at the time of the crash, held a valid commercial driver's license.

The primary reasons to focus on commercially-licensed drivers involved in single-vehicle crashes are threefold: (1) the characteristics of multi-vehicle crashes are potentially different, (2) a multi-vehicle crash involves interactions between a CDL holder and likely a non-CDL holder with possible behavioral differences, and (3) a separate study that accounts for the role of the driver's age in the interaction between multiple vehicles in a crash (CDL to CDL, or CDL to non-CDL) is needed. This study aims more to specify how commercial drivers involved in single-vehicle crashes interact with the roadway, vehicle, temporal, and environmental factors, while accounting for age group differences and possible heterogeneous effects of the risk factors. To our knowledge, this study would be the first to analyze injury severity of commercially-licensed drivers involved in single-vehicle crashes. This study attempts to contribute to the literature of CDL driver's safety by adopting econometric models to investigate possible contributing factors to the severity of drivers involved, while investigating potential unobserved heterogeneity in the covariates as well as the differences across age groups.

The remainder of this paper is structured as follows. A literature review is presented in the next section followed by the methodology section presenting an overview of the econometric approach adopted and its statistical interpretation. The data section presents the dataset utilized and the final estimation sample assembly process. The results section presents an overview of the estimation results, statistical measures-of-fit, elasticity effects, and implications of variables' effects and recommendations. Finally, the conclusion section provides an overall summary of this research along with major findings, limitations, and future scope of research.

## 2. Literature review

Literature shows a number of past studies analyzed severity of single-vehicle crashes in different settings, while other studies analyzed single-vehicle versus multi-vehicle crashes (Geedipally and Lord, 2010, p.; Martensen and Dupont, 2013; Yu and Abdel-Aty, 2013). Various studies identified the different types of roadways where single-vehicle crashes have occurred (Gong and Fan, 2017; Rusli et al., 2017; Wu et al., 2016b; Xie et al., 2012). Other studies identified the effects of specific factors (for example: age, gender, time, curb, etc.) on the severity of a single-vehicle crash (Anderson and Searson, 2015; Gong and Fan, 2017; Jiang et al., 2013; Kim et al., 2013; Martensen and Dupont, 2013; Wu et al., 2016a). Several studies identified the type-of-crash as a rollover crash for single-vehicles (Anarkooli et al., 2017; Bambach et al., 2013; Fréchède et al., 2011).

On the contrary, literature exclusively examining CDL holders irrespective of what type of vehicle being operated within the context of injury severity is scarce. Most of large truck or bus crash injury severity studies account for crashes only involving those types of vehicles, yet limited to reflect the remaining of all possible combinations of vehicle types in the commercial fleet (Al-Bdairi and Hernandez, 2017; Chang and Chien, 2013; Chang and Mannering, 1999; Chen and Chen, 2011; Dong et al., 2015; Duncan et al., 1998; Islam and Hernandez, 2013; Khattak and Targa, 2004; Khorashadi et al., 2005; Lemp et al., 2011; Pahukula et al., 2015; Wang and Shi, 2013; Zhu and Srinivasan, 2011a, 2011b). In terms of studies specifically addressing injury severity of CMV, one study used a cross-classified multilevel model to investigate the severity of CMVs while addressing heterogeneity among firms and regions (Park et al., 2017). Another study analyzed the medical condition and the severity of CMV drivers but not specifically in a single-vehicle crash setting (Laberge-Nadeau et al., 1996). Other studies addressed seatbelt usage among CMV drivers in Utah (Cook et al., 2008; Eby et al., 2002; Kim and Yamashita, 2007). Few studies addressed sleeping quality, duration, and patterns (Bunn et al., 2005; Chen et al., 2016; Hanowski et al., 2007; Lemke et al., 2016; Sparrow et al., 2016). Based on the review of the literature that focused on commercially-licensed drivers, a gap in the literature certainly exists with respect to injury severity analysis. So, additional research is needed to understand the factors that influence the injury severity of CDL holders in the event of a crash.

## 3. Methodology

Several different modeling methods have been employed to analyze crash severity data. Typically these methods can be grouped into two categories – unordered (Chang and Mannering, 1999; Holdridge et al., 2005; Savolainen and Mannering, 2007; Shankar et al., 1996; Ulfarsson and Mannering, 2004) and ordered (Eluru et al., 2008; Wang et al., 2010; Zhu and Srinivasan, 2011a; Osman et al. 2016; Osman et al. 2018). With respect to the unordered frameworks, the multinomial logit model has been widely used in injury severity literature. The multinomial logit model brings constraints such as the "independence of irrelevant alternatives (IIA)" which is, in the literature, known as the red bus/blue bus problem (McFadden, 1973). The multinomial logit model also ignores the natural ordering of injury severity outcomes which can account for misleading or inaccurate results.

In the ordered response framework (such as the ordered probit model), a single latent propensity function is assumed to be translated into the observed severity outcome depending on the value of the propensity function relative to threshold parameters (number of thresholds = number of possible severity outcomes – 1). The latent propensity function is specified as a function of different factors along with a stochastic component to account for all unobserved factors that influence injury severity. The parameters in the single propensity equation and the thresholds constitute the set of parameters that are estimated using methods such as the maximum likelihood (ML). An ordered probit model is constrained to find only one coefficient on each variable that is also in one direction, towards either higher or lower injury severity levels; a constraint that is relaxed by the MNL model. Eluru et al. (2008) extended the standard ordered response framework to develop Generalized Ordered Response (GOR) models that allow parameterization of the threshold parameters providing additional flexibility to the ordinal models (Eluru et al., 2008). So, it is not surprising that a recent comparison analysis of unordered and ordered frameworks that considers generalized version of ordered models found minor differences between the two models (Anowar et al., 2014).

Injury severity conditional on crash occurrence can depend on numerous factors all of which are most certainly not observed in crash databases. These unobserved factors can moderate the influence of other observed covariates in the model leading to variation in the parameter effects across different observations. These unobserved variations are referred to as "unobserved heterogeneity", which is of considerable importance in injury severity analysis. One important feature of the MGORP model is that it addresses possible heterogeneity in covariates. Mannering et al. (2016) describes this issue in greater detail and present alternate modeling methods available in the literature for handling the problem (Mannering et al., 2016). Among these methods, the random parameters approaches are the most prominent.