Accident Analysis and Prevention 107 (2017) 20–30

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Using fixed-parameter and random-parameter ordered regression models to identify significant factors that affect the severity of drivers' injuries in vehicle-train collisions

Essam Dabbour[a,*], Said Easa[b], Murtaza Haider[c]

[a] Center of Transportation & Traffic Safety Studies at Abu Dhabi University, Abu Dhabi University, P.O. Box 59911, Abu Dhabi, United Arab Emirates
[b] Department of Civil Engineering, Ryerson University, Toronto, Ontario M5B 2K3, Canada
[c] Ted Rogers School of Management, Ryerson University, Toronto, Ontario M5B 2K3, Canada

ARTICLE INFO

Keywords:
Injury severity
Ordinal regression
Random-parameter ordered probit models
Mixed ordered models
Temporal stability
Negative log–log function
Railroad-grade crossings

ABSTRACT

This study attempts to identify significant factors that affect the severity of drivers' injuries when colliding with trains at railroad-grade crossings by analyzing the individual-specific heterogeneity related to those factors over a period of 15 years. Both fixed-parameter and random-parameter ordered regression models were used to analyze records of all vehicle-train collisions that occurred in the United States from January 1, 2001 to December 31, 2015. For fixed-parameter ordered models, both probit and negative log–log link functions were used. The latter function accounts for the fact that lower injury severity levels are more probable than higher ones. Separate models were developed for heavy and light-duty vehicles. Higher train and vehicle speeds, female, and young drivers (below the age of 21 years) were found to be consistently associated with higher severity of drivers' injuries for both heavy and light-duty vehicles. Furthermore, favorable weather, light-duty trucks (including pickup trucks, panel trucks, mini-vans, vans, and sports-utility vehicles), and senior drivers (above the age of 65 years) were found be consistently associated with higher severity of drivers' injuries for light-duty vehicles only. All other factors (e.g. air temperature, the type of warning devices, darkness conditions, and highway pavement type) were found to be temporally unstable, which may explain the conflicting findings of previous studies related to those factors.

1. Introduction

Collisions between trains and motor vehicles usually cause severe injuries to vehicle drivers due to the large weight of a typical train as compared to that of a typical motor vehicle. That weight ratio is estimated to be approximately 4000 to 1 (Hao and Daniel 2016; Yan et al., 2010). According to statistical data obtained from the Federal Railway Administration (FRA, 2017), there were 35,751 reported collisions between motor vehicles and trains in the United States during the period from January 1, 2001 to December 31, 2015. Those collisions include 2849 fatal collisions and 9003 injury collisions.

Several studies attempted to identify risk factors that may increase the frequency of collisions between trains and motor vehicles so that safety treatments may be implemented to reduce the frequency of those collisions (Saccomanno et al., 2007; Oh et al., 2006; Saccomanno et al., 2004; Austin and Carson 2002). Other studies focused on analyzing risk factors affecting the severity of collisions between trains and motor vehicles. Hao et al. (2016) used ordered probit modeling to analyze all collisions that occurred at railroad-grade crossing in the United States during the period from 2002 to 2011, and they found that injury severity increased during AM and PM peak hours as well as during PM off-peak period. Hao and Daniel (2013) used ordered probit modeling to determine the significant factors influencing the severity of drivers' injuries resulting from collisions with trains at railroad-grade crossings in the United States. They found that female and older drivers were more likely to be involved in severe collisions compared to male and younger drivers, respectively. Eluru et al. (2012) developed a latent class model to identify risk factors that may increase the severity of drivers' injuries resulting from collisions at railroad-grade crossings by using collision data from 14,532 crossings in the United States from 1997 to 2006. The factors that were found to significantly influence injury severity included driver's age, time of the collision, and weather conditions. Hu et al. (2010) used logit modeling to identify risk factors that increase the severity of collisions at railroad-grade crossings using data from Taiwan for the period from 1995 to 1997. Variables such as daily number of trains, daily number of trucks, and the use of obstacle-

* Corresponding author.
E-mail addresses: essam.dabbour@adu.ac.ae (E. Dabbour), seasa@ryerson.ca (S. Easa), murtaza.haider@ryerson.ca (M. Haider).

http://dx.doi.org/10.1016/j.aap.2017.07.017
Received 3 December 2016; Received in revised form 16 June 2017; Accepted 12 July 2017

detection devices were found to have an impact on the severity of collisions. McCollister and Pflaum (2007) used logistic regression to develop models that predict the probability of collisions, injuries, and fatalities at railroad-grade crossings. The risk factors that were found to increase injury severity included train speed, number of trains, traffic volume, and collision history. The proportion of heavy vehicles in the traffic volume and the proximity of traffic control devices were also found to impact collision severity.

It should be noted that all the aforementioned studies that investigated the factors related to drivers' injury severity, resulting from collisions with trains at railroad-grade crossings, were based on compiling data from different years into one dataset that was used for analysis. However, the factors identified in those studies might be temporally unstable so that they might have been significant during the period being analyzed and then their significance might have changed due to the ongoing changes in vehicle and train technologies, drivers' attitudes, traffic volumes, road conditions, railroad-crossing conditions, law-enforcement technologies, and implemented policies. Understanding the temporal stability trends of the factors affecting injury severity would be helpful in assessing the effectiveness of implementing different safety treatments so that researchers would be able to identify whether any safety improvements, obtained after implementing a specific treatment, are attributed to that treatment or simply attributed to the temporal instability of the factors being addressed.

A few studies have attempted to investigate temporal stability of risk factors related to crash severity. Dabbour (2017a) used logistic regression to investigate temporal stability of the factors related to the severity of drivers' injuries in single-vehicle collisions by analyzing all single-vehicle collisions that occurred in North Carolina for seven years. It was found that most of the factors were temporally stable. Those factors included rural and undivided highways, darkness conditions, favorable environmental conditions, vehicle age, driver's gender, influence by alcohol or other illicit drugs, and driver's failure to properly use seatbelt. Behnood and Mannering (2015) used mixed-logit modeling to explore temporal stability of the factors affecting the severity of drivers' injuries in single-vehicle collisions using nine-year crash records from Chicago, Illinois. They found that the effects of most factors were temporally unstable. They attributed the temporal instability to several factors, including ongoing changes in vehicle safety features, drivers' response to those improved safety features, and drivers' response to changes in microeconomic conditions. Another study (Xiong et al., 2014) investigated temporal stability of the factors affecting the severity of single-vehicle collisions by analyzing police-reported single-vehicle crash data in Indiana over five years. It was found that the factors related to drivers' injury severity in single-vehicle collisions were generally temporally unstable. The authors attributed the temporal instability to several factors, including ongoing changes in average annual daily traffic, international roughness index, and rut depth.

The purpose of this study is to investigate temporal stability of the factors affecting the severity of drivers' injuries when colliding with trains at railroad-grade crossings by analyzing individual-specific heterogeneity related to those factors over a period of 15 years. Both fixed-parameter and random-parameter ordered regression models are used to analyze records of all vehicle-train collisions that occurred in the United States from January 1, 2001 to December 31, 2015. For fixed-parameter ordered models, both probit and negative log–log link functions were used. The latter function accounts for the fact that lower injury severity levels are more probable than higher ones. Due to the substantial differences in vehicle and driver characteristics between heavy and light-duty vehicles, separate models were developed for each of the two categories.

In the next two sections, the dataset and the methodology used for the research are both presented. The other sections provide details and discussion on the developed models with validating the identified factors and highlighting the novel contributions of the study. The last section provides concluding remarks and recommendations for future research.

2. Data collection and preparation

The dataset used for this research is obtained from the Federal Railway Administration (FRA, 2017). The dataset covers all collisions that occurred in the United States between trains and motor vehicles, at railroad-grade crossings, for 15 years starting from January 1, 2001 to December 31, 2015 after excluding suicides. Using collision records that extend for fifteen years provides the benefits of capturing year-to-year temporal changes over that period. During that period, there were 37,846 reported collisions at railroad-grade crossings, from which 35,751 collisions were between trains and motor vehicles. The remaining 2095 collisions were between trains and either pedestrians, motorcycles, or non-motorized vehicles.

For the purpose of this research, the outcome variable, y_i, is the level of severity of driver's injury, which takes one of the following injury severity levels:

- Level 1: the driver was uninjured;
- Level 2: the driver was injured; or
- Level 3: the driver was killed.

Note that the severity levels reported in the original dataset (FRA, 2017) were actually in descending order (with level 1 corresponding to the driver being killed and level 3 corresponding to the driver being uninjured). However, for the purpose of this research, the order has been reversed into the shown ascending order. The use of the ascending order makes it easier to interpret the signs of the estimated coefficients so that a coefficient with a positive sign indicates that the factor associated with it increases injury severity, and vice versa.

Due to the substantial differences in vehicle and driver characteristics between heavy and light-duty vehicles, separate models were developed for each of the two categories. For the purpose of developing the models, the following dichotomous explanatory variables (factors) were investigated:

- Darkness (*DA*): takes a value of "1" if the crossing was in darkness at time of the collision (with no daylight or illumination), and a value of "0" otherwise;
- Favorable weather (*FW*): takes a value of "1" if the weather at the time of the collision was favorable (i.e. clear or cloudy), and a value of "0" otherwise (i.e. rain, snow, sleet, or fog conditions);
- Freezing temperature (*FR*): takes a value of "1" if the air temperature at the time of the collision was freezing (i.e. below 32 degrees Fahrenheit), and a value of "0" otherwise;
- Obstructed view (*OB*): takes a value of "1" if the view of the vehicle driver was obstructed (by a natural object, a built object, or another train on an adjacent track), and a value of "0" otherwise;
- Paved highway (*PV*): takes a value of "1" if the highway was paved, and a value of "0" otherwise;
- Smallest crossing angle (*AN*): takes a value of "1" if the smallest crossing angle was between 60° and 90°, and a value of "0" otherwise;
- Private crossing (*PC*): takes a value of "1" if the crossing was private, and a value of "0" if it was public;
- Passive warning devices (*PW*): takes a value of "1" if the crossing was equipped with only passive warning devices (e.g. signs or pavement markings with no bells, gates or flashing lights), and a value of "0" if the crossing was equipped with active warning devices (i.e. bells, gates, and flashing lights);
- Young driver (*YD*): takes a value of "1" if the age of the vehicle driver was less than 21 years, and a value of "0" otherwise;
- Senior driver (*SD*): takes a value of "1" if the age of the vehicle