



National University Rail Center - NURail
US DOT OST-R Tier 1 University Transportation Center

What is the extent of harm in rail-pedestrian crashes?

By

Asad J. Khattak

Beaman Professor & Transportation Program Coordinator
Department of Civil and Environmental Engineering
The University of Tennessee, Knoxville
Email: akhattak@utk.edu

David Clarke

Director, Center for Transportation Research
The University of Tennessee, Knoxville
Email: dclarke@utk.edu

Jun Liu, Ph.D.

Research Associate
Department of Civil & Environmental Engineering
The University of Tennessee
Email: jliu34@utk.edu

Xin Wang, Ph.D.

Senior Modeling Systems Analyst
Virginia Department of Transportation
Email: xwang757@gmail.com

&

Meng Zhang

Research Assistant
Department of Civil and Environmental Engineering
University of Tennessee, Knoxville
Email: mzhang27@vols.utk.edu

Grant Number: DTRT12-G-UTC18

DISCLAIMER

Funding for this research was provided by the NURail Center, University of Illinois at Urbana-Champaign under Grant No. DTRT12-G-UTC18 of the U.S. Department of Transportation, Office of the Assistant Secretary for Research & Technology (OST-R), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

ACKNOWLEDGMENT

The data used in this study was provided by the Federal Railway Administration of the US Department of Transportation. We are very thankful for the support received from the National University Rail (NURail) Center (a US DOT OST-R Tier 1 University Transportation Center), the Center for Transportation Research, and the Transportation Engineering & Science Program at the University of Tennessee. Support from the Office of the Secretary of Transportation is greatly appreciated. The views presented in this paper are those of the authors, who are responsible for the facts and accuracy of the data. The statistical software GWR 4 and STATA were used for modeling and ArcGIS for visualization of results.



TECHNICAL SUMMARY

This project provides a comprehensive understanding of rail-trespassing crashes. Specifically, the project explores injury severity in rail-trespassing crashes that occurred at highway-rail grade crossings and non-crossings. The authors examined the role of pre-crash behaviors, socio-demographics, time of crash, location and traffic controls where crash occurred on injury severity. Relying on the accident and incident data as well as inventory and highway-rail crossing data maintained by the Office of Safety Analysis of the Federal Railroad Administration, the project is split into three efforts:

The first effort examines spatial patterns of rail-trespassing crashes at non-crossings, using 10 years of crash data. Given the geocode information in rail-trespassing crash data, this part of the study visualizes the distribution of rail-trespassing crashes and analyzes the spatial correlates of injury severity with related factors (e.g., crash time, individual attributes, darkness, crash location and pre-crash behaviors) across the United States. The results indicate that: 1) higher rail-trespassing fatality chances, given a crash, are associated significantly with pre-crash behaviors, especially lying and sleeping on rail tracks. Seniors are more likely to be killed in rail-trespassing crashes while youths are less likely, compared with adults; 2) a critical finding is uncovering spatial variations in the correlates, given the variations in spatial locations of crashes. The study finds that relationships between crash injury severity and certain variables from one location cannot be generalized to other locations. By visualizing the associations to show where certain pre-crash actions are associated with higher probability of trespassing crash injury, this project provides information that can help stakeholders focus attention on critical factors in geographies that show higher risks of fatalities in crashes. This project provides insights into countermeasures that can be targeted regionally to reduce higher risk trespassing behaviors.

The second effort analyzes rail-pedestrian and bicyclist trespassing crashes at highway-rail grade crossings to non-crossings by comparing them. The research effort focuses on the effects of pre-crash behaviors on crash injury severity. The analysis was done separately for highway-rail grade trespassing crashes and non-crossing trespassing crashes. The project explored differences in injury and correlates of trespasser injury severity between crossings and non-crossings. Results show that: 1) lying or sleeping on or near tracks is associated with higher chance of fatal injury at both crossings and non-crossings, but more so at rail grade crossings; 2) sitting/standing/bending/stooping are more injurious at non-crossings, while crossing/crawling are more injurious at highway-rail grade crossings; 3) crashes occurring during darkness and summer are more injurious at grade crossings; 4) the trends show more fluctuations for crossing crashes, but relatively more stability for non-crossing crashes over a 10 year period (2005-2014). The study points to using different types of countermeasures in order to reduce injuries at rail grade crossings and non-crossings.

The third effort focuses on injury outcomes at highway-rail grade crossings, expanding the analysis from non-motorized trespassers to include motorized trespasser crashes. Path analysis quantifies the direct and indirect associations of passive control (crossbucks and stop signs) and active controls (gates, flashing lights, audible warnings and highway signals) with

pre-crash behaviors and injury severity. The study reveals that: 1) some crossing controls (e.g., presence of gates at crash sites) do not have a significant direct association with injury severity, but are indirectly associated with injury severity through pre-crash behaviors—the presence of gates is indirectly associated with lower injury severity; 2) more broadly, given a crash, active controls were associated with lower driver injuries compared with passive controls. This project indicates that understanding key correlates of injury severity can come through understanding of pre-crash behaviors. The results emphasize the need to develop a deeper understanding of motorists' gate-violation behaviors at highway-rail grade crossings. Indeed, pre-crash behaviors are important predictors of rail-trespassing safety. This study also showed that safety improvement strategies can be made more effective by customizing them to specific regions in the United States.

NOTE:

The project has resulted in the following publications and presentations:

1. Wang, X., Liu, J., Khattak, A. J., & Clarke, D. (2016). Non-crossing rail-trespassing crashes in the past decade: A spatial approach to analyzing injury severity. *Safety Science*, 82, 44-55. (TRB paper # 15-0955, Presented at the Transportation Research Board, National Academies, Washington, D.C., 2015.)
2. Zhang, M., A. Khattak, J. Liu, & D. Clarke, A comparative study of rail-pedestrian/bicyclist trespassing crash injury severity at highway-rail grade crossings and non-crossings. Presented at 2015 Road Safety & Simulation International Conference, Orlando, FL, 2015.
3. Liu, J., Khattak, A. J., Richards, S. H., & Nambisan, S. (2015). What are the differences in driver injury outcomes at highway-rail grade crossings? Untangling the role of pre-crash behaviors. *Accident Analysis & Prevention*, 85, 157-169. (TRB paper # 15-0959, Presented at the Transportation Research Board, National Academies, Washington, D.C., 2015.)
4. Liu J., A. Khattak & S. Richards, Examining motorist gate-violation behaviors at highway-rail grade crossings using geo-spatial modeling integrated with path analysis: a screening tool for safety improvement zones. Presented at 2015 Road Safety & Simulation International Conference, Orlando, FL, 2015.

NON-CROSSING RAIL-TRESPASSING CRASHES IN THE PAST DECADE: A SPATIAL APPROACH TO ANALYZING INJURY SEVERITY

Xin Wang, Jun Liu, Asad Khattak, and David Clarke
The University of Tennessee, Knoxville

Abstract: Transportation professionals have long recognized the harm of trespassing along railway rights-of-way. However, the non-crossing rail trespassing issue has received less attention than highway-rail grade crossing crashes, despite the fact that nearly 8,800 rail-trespassing crashes occurred on non-crossing rail tracks during the past decade, with a large number of them resulting in fatality. Also, geographic and socio-demographic diversity within the US imply that trespassing crash severity and its correlates may vary across geographic entities or regions. The purpose of this paper is to investigate these issues using rail-trespasser crash data maintained by Federal Railroad Administration (N=8,794 over 2004 to 2013). The unique aspects of the study are the development of a framework that explores correlates of injury severity in crashes, and applies appropriate analysis methods. Specifically, using rigorous spatial modeling methods (Geographically Weighted Regression), the study uncovered spatial variations in correlates between rail-trespassing injury and revealed contributing factors. The factors include personal attributes of individuals, environmental and location attributes, time of crash, and pre-crash behaviors. The results show that non-crossing trespass crashes are generally severe with 52.1% involving a fatality. Pre-crash behaviors were found to be key factors showing significant associations with the probability of rail-trespassing injury, especially lying or sleeping (on or near tracks). Fundamentally, the basic assumption of spatial stationarity in traditional regression models does not fully hold in the situation explored. Correlates of injury severity are found to be non-stationary across space. Therefore, regional considerations in specific situations should guide the implementation of treatments and policies.

Keywords: rail-trespassing, crash, geographically weighted regression, injury severity, spatial pattern

SECTION 1: INTRODUCTION

While crashes at highway-rail grade crossings with vehicles, pedestrians, bicyclists and other users are a continuing societal concern, a growing concern is train crashes with trespassers on railroad rights-of-way other than designated grade crossings. Railway trespassers are individuals who commit the act of trespassing on the railway property without the permission of the property owner, costing billions of dollars annually in injuries and fatalities. USDOT/FHWA guidance recommends that non-motorist-crossing safety should be considered at all rail grade crossings to minimize pedestrian crossing time and avoid trapping pedestrians between sets of tracks. But there is no such safety guidance at non-crossing trails. Safety at non-crossings is lightly researched and needs to be investigated further. Trespassers not only endanger themselves but also expose railway staff and passengers to unnecessary delays and strain public services. Rail-trespasser crashes are particularly problematic, as they are associated with more fatalities than

any other form of railway-related crashes (Lobb, 2006). In addition, individuals are more likely to be killed or irreparably injured in crashes with trains, compared to the crashes with highway vehicles (Evans, 2003). With rail-trespasser crashes a key concern in the future (due to a resurgence in goods movement through trains in the US), the problem may get worse if appropriate actions are not taken, e.g., adding surveillance and enhancing public education.

Little is known about people who violate or trespass by crossing, walking or taking other actions along tracks at places other than a designated level crossing. Most trespassers are pedestrians, but some can be people who driving or riding a bike, ATV, dirt bike, snowmobile, vehicles, etc. Investigation into the characteristics of trespassers as well as their pre-crash actions could help in the development of countermeasures to reduce the number of crashes.

To provide valuable information to decision-makers and take advantage of the expansion in computation power, availability of geo-referenced data, and geographic information systems, this study investigates injuries in rail-trespassing crashes that occur along railway tracks. It explores 1) how rail-trespassing crashes are distributed spatially in the United States; 2) the correlates of injury severity to trespassers that include personal attributes, environmental and location attributes (e.g., railroad yard), time of crash and trespasser actions, i.e., pre-crash behaviors; 3) how such associations are distributed across the country, given the geographical and social diversity.

SECTION 2: LITERATURE REVIEW

Much of the previous research has focused on rail-pedestrian and bicyclist crashes, especially on crashes at highway–rail grade crossings (Khattak and Luo, 2011; Metaxatos and Sriraj, 2013). The motivations for trespassers to cross railway tracks at improper locations or their presence in other railroad right-of-way areas vary substantially. Taking the shortest or most convenient route by crossing tracks is one of the most common reasons for pedestrian trespassing; people in one of the study reported that the safe, legal route via an overbridge took more time and effort and hence they decided to trespass (Lobb et al., 2001).

A macabre motivation may be the desire to commit suicide. In European countries, most of the rail-pedestrian crashes turned out to be suicides (Van Houweligen and Beersma, 2001; Silla and Luoma, 2012a, b). In the United States, there is also strong speculation that a substantial amount of rail-pedestrian crashes may be suicides (Savage, 2007). But perhaps suicides in rail crashes are lower in the US than in some of the other countries that restrict access to firearms.

Socio-demographics are usually used to draw the picture of train-pedestrian crashes, based on the possibility that people belonging to certain socio-economic groups may be more likely to be involved in trespassing crashes. Children and senior trespassers are vulnerable, though relatively few such crashes involved children under the age of 10 or seniors above the age of 60 (Pelletier, 1997; Silla and Luoma, 2012a). Summarizing a decade of train-pedestrian crashes in Charleston, South Carolina, Cina et al. found that young males accounted for a majority of rail-pedestrian crashes in their data (Cina et al., 1994). They further found that 80% of such crashes involved blood alcohol levels greater than 99 mg/dL. Pelletier reported that trespasser fatalities typically involved unmarried males with less than a high school education (Pelletier, 1997). He also pointed out the problem of alcohol intoxication in such trespassers. Lobb et al. conducted a self-reported survey to investigate the behaviors of individuals crossing

the railway (Lobb et al., 2001). Their findings suggested that teenagers and males have more dangerous attitudes and are more likely to walk across railroad tracks.

The time of rail-pedestrian crash occurrence is also a concern in the literature. Silla et al., analyzed documented rail-pedestrian crashes on the Finnish railway network and reported a large portion of crashes occurred in the afternoon and evening and a great number of crashes occurred on weekends. Also summer and winter time had more such crashes than spring and fall months (Silla and Luoma, 2012a). Pelletier found that fatality-involved crashes typically occurred at night at the end of a week (Pelletier, 1997). Lerer et al., report that rail-pedestrian crashes occurred at peak commute times in Cape Town, South Africa (Lerer and Matzopoulos, 1996). Investigations of crash locations have revealed that rail-pedestrian crashes typically occurred in areas with dense population and train activity. Such places included the vicinity of residential communities and train stations and rail yards (Silla and Luoma, 2012a).

Researchers also have found that trespasser pre-crash behaviors have a strong connection with the severity of injury. Pre-crash behaviors include walking, running, standing, sitting, lying, etc. in railroad rights-of-way. A study based on three-year rail-pedestrian crash records revealed that walking and lying were the two major pre-crash behaviors associated with trespassers (Savage, 2007). Another study found that pre-crash behaviors are helpful in revealing the reasons of the crash occurrence—a large portion of sitting or lying behaviors were strongly suspected to be suicides or intoxicated individuals (Savage, 2007; Silla and Luoma, 2012a).

Prevention of the rail-pedestrian crashes through treatments has been investigated in the literature (Pelletier, 1997; CDCP 1999; Lobb, 2006; Savage, 2007; Liu et al., 2011; Silla and Luoma, 2011, 2012b, a). Mohanty et al. suggested surveillance and public education as useful ways to decreasing the frequency of rail-pedestrian crashes (Mohanty et al., 2007). Surveillance in particular places that include typical trespassing crash locations, rail yards and highway-rail grade crossings can help. However, it is impractical and cost-prohibitive to monitor all such places. Public education may help people realize the danger of crossing tracks illegally (Lobb et al., 2001). Studies by Mok and Savage indicated that increasing the amount of public educational activities on railway safety can be effective in reducing the number of train-motor vehicle crashes at highway-rail crossings, through investigating crash reductions after a campaign called Operation Lifesaver initiated in the 1970s (Mok and Savage, 2005; Savage, 2006). However, a follow-up study by Savage did not show a relationship between trespassing crashes and the implementation of Operation Lifesaver, owing to two potential reasons: 1) there was simultaneous growth in Operation Lifesaver programs and railroad abandonments (which would decrease train activity and crashes), and 2) Operation Lifesaver mainly focused on risks at grade crossings rather than non-crossings (Savage, 2007). Savage further discussed the associations of educational activities with trespassing behaviors and suggested Operation Lifesaver to redirect some of their activities to places (school and civic groups) located close to the tracks (Savage, 2007). Studies have investigated the effects of three countermeasures (landscaping, building a fence and prohibitive signs) on the frequency of trespassing, and found that fencing can reduce frequency of trespassing by 94.6% (Silla and Luoma, 2011).

Nearly all the above mentioned studies have focused on trespassing crash frequency instead of the crash severity (harm suffered by the trespasser) given a crash, except a study conducted by Pelletier (1997). While a few previous studies have investigated non-crossing trespassing crashes, even fewer have taken advantage of the available computation power that allows more data-intensive spatial analysis of rail-pedestrian crashes. While widely used to assess highway crashes (Levine et al., 1995; Loo, 2009; Plug et al., 2011), spatial analysis has

not received a large application in rail trespassing studies. This is partially due to the limited availability of geo-coded rail-trespassing data. However, this situation has improved considerably because the Federal Railroad Administration (FRA) now includes geocode information in its rail-trespassing data, making it possible to apply location-aware modeling methodology to help demonstrate the spatial patterns of rail-trespassing and understand the factors associated with rail-trespassing crash outcomes. Given that rail-trespassing is a national issue, it is of particular interest to use robust spatial visualization and state-of-art modeling methods to analyze relevant data.

SECTION 3: APPROACH AND METHODOLOGY

After obtaining the relevant data, the study first conducted univariate analyses for exploring the distribution of variables and descriptive statistics. They provided information about outliers in the data. Data were visualized using spatial statistics methods. Next, bivariate analyses helped understand first-order correlations in the data and understand simple hypothesized relationships. In addition, multivariate statistical models, in particular global and local regression models, were estimated to explore key relationships. The role of pre-crash behaviors and actions of trespassers was of interest and how these lead to certain types of collisions, and hence injuries. The severity measure used was whether the injuries were fatal or non-fatal. The methods are described in greater detail below.

3.1 Visualizing Non-crossing Rail-Trespassing Crashes

Visualization and analysis of kernel density can improve our understanding of spatial distribution of railway track trespassing at non-crossings. A distribution density analysis is used to calculate the density of trespassing crashes in a search radius. Therefore it can show where trespassing crashes are concentrated. Kernel density is used in this paper.

The kernel function K determines the shape of the bumps while the radius h determines their width (Silverman 1986). The incident density is used to create a surface showing the spatial distribution of trespassing crashes (fatal and/or non-fatal) throughout the country. The kernel density function $\hat{f}(x)$ is described in Equations 1 and 2.

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K \left\{ \frac{1}{h} (x - X_i) \right\} \quad (\text{Equation 1})$$

Where: n = sample size

h = bandwidth parameter (kernel radius)

x = Location of crash around which kernel is centered

X_i = Observed trespass crashes in the kernel

The function $K(x)$ will be a symmetric probability density function—the normal density, for instance, or Gaussian function (Shown in Equation 2) with mean zero and variance one.

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad (\text{Equation 2})$$

3.2 Model structure

A-priori, given a rail-trespassing crash, fatality (vs. non-fatal injury) will likely be associated with lower visibility (measured by darkness-a binary variable), certain dangerous pre-crash behaviors (such as lying on or near tracks, walking on or near tracks, sitting or standing, driving, riding a bicycle on railway tracks, etc.), whether the track is located in a yard, people at higher risk (such as young or senior, derelicts, drunk or intoxicated individuals), and the crash location type is rural (as opposed to urban or suburban). Figure 1 shows the different types of risky trespassing actions that are exhibited along railway tracks.

A logistic model, referred to as the global model, was estimated for testing these hypotheses:

$$Y = \beta_0 + \beta_1 (\text{dark}) + \beta_2 (\text{pre-crash action}) + \beta_3 (\text{location}) + \beta_4 (\text{personal attributes}) + \beta_5 (\text{season}) + \varepsilon \quad (\text{Equation 3})$$

$$Y = \text{Ln} \left(\frac{\text{Prob}(\text{Fatal Injury})}{1 - \text{Prob}(\text{Fatal Injury})} \right)$$

Prob (Fatal Injury) = Probability of trespassing fatality (0=non-fatal injury, 1=fatal injury)

β = Coefficients for variables

ε = Error term in Model 1.

Note that, unlike studies on the crash frequency that use count data modeling techniques, i.e., Poisson or negative binomial models (Famoye et al., 2004; Savage, 2007; Millegan et al., 2009; Ye et al., 2009; Russo and Savolainen, 2013; Ye et al., 2013), this study attempts to untangle the correlates of the trespassing injury severity, given a crash. Injury severity is coded as a binary variable, justifying the use of binary logistic model.

The traditional logistic model (Model 1) estimates the average associations between trespassing fatality and explanatory variables, which is a “one size fits all” model and does not consider the possibility of spatial variations (i.e., “spatial heterogeneity”) in associations of variables with injury. This has two drawbacks-specifically the traditional logistic model does not consider that 1) the magnitude of model coefficients can change across the country; 2) the statistical significance of coefficients can also vary in space. This means that statistically significant associations in one part of the country may not necessarily hold in other parts of the country. For a geographically diverse country like the United States, these two assumptions may not hold, or at least should be tested empirically. Small local variations can justify the use of a global model. If local variations are relatively large, then the relationships are not stationary across space and local models (GWR) should be applied. Overall, the spatial model applied in this study captures spatial heterogeneity in associations between the rail-trespassing crash injury and explanatory variables.

3.2 Spatial Modeling

To overcome drawbacks of the traditional logistic model, a spatial model known as the Geographically Weighted Logistic Regression (GWLR, shown in Model 2) was estimated in this study. It explores geographical drifts in regression parameters. Specifically, GWLR estimates spatial variations of associations between dependent and explanatory variables by relaxing the

assumption that estimated parameters hold globally in a traditional (global) model. The coefficients in the spatial model are no longer fixed but vary according to their locations; therefore, the local GWR model takes the following form with coefficients varying for each location:

$$Y_i = \beta_{0i} + \beta_{1i} (\text{dark}) + \beta_{2i} (\text{pre-crash action}) + \beta_{3i} (\text{location}) + \beta_{4i} (\text{personal attributes}) + \beta_{5i} (\text{season}) + \varepsilon_i$$

This can be written in its basic form as follows:

$$Y_i = \beta_{i0} + \sum_{k=1}^p \beta_{ik} x_{ik} + \varepsilon_i \quad (\text{Equation 4})$$

$$Y_i = \text{Ln} \left(\frac{\text{Prob}(\text{Fatal Injury})_i}{1 - \text{Prob}(\text{Fatal Injury})_i} \right)$$

$\text{Prob}(\text{Fatal Injury})_i$ = probability of trespassing injury (0=non-fatal injury, 1=fatal injury) for each trespassing crash location i ($i = 1, 2, \dots, n$, where n is the sample size: 8,794);

β_{i0} = the constant at trespassing crash location i ;

β_{1i} = the parameter at trespassing crash location i for explanatory variable x_{ik} ;

x_{ik} = explanatory variables of the k^{th} parameter for trespassing crash location i ,

ε_i = error term at trespassing crash location i ,

To obtain locally changing parameter estimates, instead of using all 8,794 samples for a global model regression, GWLR performs a regression for each trespassing crash location i using only a subset of the trespassing crashes that are close to i in space (this neighborhood area is called the “kernel”). This means a specific local model will be estimated for each trespassing crash location by using those samples within its kernel neighborhood. This is based on the assumption that crashes in close proximity share more similarity than crashes that are far away. The size of the kernel (defined as the distance between location i and the edge of kernel in space) is termed as “bandwidth.” In this sense, if a kernel with fixed bandwidth is used for every trespassing crash location, local sample size (the number of subsamples) for each regression location would be different since trespassing events are not distributed evenly in space. Therefore kernels with fixed bandwidth will result in large local sample sizes in those areas with high trespassing density, but small local sample sizes or even near zero local sample size in those areas with sparse trespassing events. This will cause problems in the regression model. To solve this problem, an adaptive kernel was used, which ensures the bandwidth was selected so that each regression location had the same local sample size. That means larger kernels are used for locations with sparse trespassing crashes in its neighborhood while small kernels are used for those locations with denser trespassing crashes nearby.

Moreover, the subsamples in the kernel are not treated equally but each of them is given a weight which is inversely proportional to its distance from the trespassing regression location. In this sense, each local model is actually a weighted logistic model using the sub samples (crashes) in its kernel. Given that each trespassing crash is surrounded by a unique set of other trespassing crashes, the results of local calibration are unique to the particular location. In this regard, by plotting the results of these local calibrations on a map, continuous surfaces of parameter estimates can be generated while discontinuities and sudden changes in magnitude are minimized. This makes the technique attractive for various aspects of spatial analysis, a benefit demonstrated by Paez and Scott (2005). GWR has been also used in the transportation field.

Wang and Khattak (2013) used Geographically Weighted Logistic Regression (GWLR) to explore spatial patterns of traveler information and travel decisions and they have found improvements in goodness of fit and forecasts over other model forms. Also, in transportation safety, Hadayeghi et al. (2009) utilized geographically weighted Poisson regression (GWPR) to model zonal collision counts and concluded that the local model estimation technique of GWPR can improve analysis of transportation networks. Park et al. (2010) used GWR to identify hazardous locations based on severity scores of highway crashes. As a relatively new and novel modeling approach, GWR is used to analyze rail-trespassing crashes.

Instead of using GWR, the mixed-effects model can also be used for conducting similar analysis (correlates of injury and variations across the country), yet allowing random effects for different observations (McCulloch and Neuhaus, 2001; Liu et al., 2015a). Unlike GWR, which examines spatial variations of correlates across the space, the mixed-effects logit regression reveals the variations of correlates between observation groups, e.g., state-wide or county-wide. GWR consistently uses the same number of observations for each regression (sub-sample), while the mixed-effects model often uses different number of observations for each entity, e.g., state or county. A model may not converge for an entity that has only a few observations. A key advantage of GWR is that it allows regression across state borders and estimates the correlates within a kernel containing geographically connected observations. Both GWR and Mixed-effects models allow variations of correlates between observations, while GWR provides more flexibility in terms of variation in estimates and does not consider artificial boundaries, given spatially distributed observations. Thus GWR models were selected for this study. However, the results of mixed-effects models are also available from the authors.

SECTION 4: DATA DESCRIPTION

The Office of Safety Analysis of the Federal Railroad Administration (FRA) maintains a detailed national database for railroad safety information including crashes, inventory, and highway-railroad crossing data. All railway-related crashes are documented in a safety data website (<http://safetydata.fra.dot.gov/>) and are available to the public. Ten years (2004-2013) of railway trespassing data were extracted for this study. Since this study concentrates on non-crossing rail-trespassing only, crashes occurring at highway-rail grade crossings were excluded.

The data describing each railway trespasser crash includes crash features (date, time etc.), crash location, limited personal attributes of the individuals (age), pre-crash behaviors (walking, sitting, lying, driving, etc. in railroad right-of-way), injury nature (amputation, fracture, killed, etc.) and location (latitude and longitude, county, state). Note that limited geocode information was available for trespassing crash data before 2010; therefore crash data without a geocode were coded into the centroids of the county where the trespassing crash occurred (railroad mileposts that provide more accurate location information were also not available). GIS tools were applied to merge the railway track trespassing data, state boundary and county boundary (downloaded from Census.gov) and to assist with identifying darkness based on sunrise and sunset times by each month. The final data were verified and error checked for reasonableness.

4.1 Descriptive Statistics

The data used in this paper include 8,797 non-crossing rail-trespassing crashes that occurred during the past decade (2004-2013). Table 1 shows the descriptive statistics for the variables used in models. Over one-half of the trespassers (52.91%, 4,590 out of 8,794 individuals) were killed in trespassing crashes. Most trespassers involved in crashes were adults ranging between 17-64 years old. One-half of the trespassing crashes happened during the night and one third happened during the weekend. However, t-tests for differences between means of two groups do not show significantly different effects (5% level) of nighttime and daytime on trespassing injury severity (getting hit by a train seems to be uniformly bad irrespective of day or night). Also, there is no significant difference between weekdays and weekends regarding trespassing injury severity—a result that is inconsistent with a previous study (Silla and Luoma, 2012a).

This study considers the interaction between trespassing behaviors and the land use. A variable termed “Land Use Mix Index,” was included in the analysis. The index was developed by Environmental Protection Agency (EPA) to capture the land use mix entropy level, which varies from 0 (homogeneous land use, such as in rural areas or suburban subdivisions) to 1 (most mixed, such as diverse city centers) land use (GeoDa Center, 2015). The land-mix index is calculated based on how different land development with various densities are mixed, e.g., developed open space, developed low intensity, developed medium intensity, developed high intensity, etc. Only one county in Texas has zero mix use which contains developed open space only.

Non-crossing trespassing crashes happened most frequently in situations when the individual was lying or sleeping, sitting or standing, walking or running, driving or riding on railway tracks. Crashes that involve lying or sleeping on tracks are particularly noteworthy as they account for about 23.35% of all trespassers. While some of them may involve suicide, such crashes cannot be directly linked to suicides. Notably, suicide cases are judged by the Coroners or medical examiners. FRA regulations did not require railways to report suicides so there is no system to reconcile the FRA trespasser database with local records from Coroners or medical examiners (Savage, 2007). An early FRA report (by Cadle Creek Consulting, 2008) analyzing data from 2002-2004 with records from local coroners and medical examiners suggests that about 20% of the deaths in the database were suicides. As a result of the Rail Safety Improvement Act of 2008 (RSIA), in June 2011 FRA began to collect incident reports of suicides as specified by the revised Code of Federal Regulations (CFR) and reflected in the FRA Guide for Preparing Accident/Incident Reports (Federal Railroad Administration, 2011). A recent FRA report by Gabree et al., (2014) states that 242 fatalities were determined to be suicides on railroad rights-of-way in 2012, compared with 429 trespasser fatalities (non-suicides) and 232 grade crossing fatalities. This gives 26.7% ($242 / (429 + 232 + 242)$) of all railroad right-of-way fatalities were suicides; the study further reports that the percentage of suicides can be as high as 35%. The study also mentioned that suicides are treated differently from the trespasser fatalities for reporting purposes and the data on suicides are stored in a separate database, which is not publically available. The authors of this study did not have access to suicide data.

4.2 Spatial distribution of non-crossing rail-trespassing crashes

Figure 2 shows the kernel density for all non-crossing rail-trespassing crashes and fatal trespassing crashes in the past decade. Each point in the graph represents a single trespassing crash. Green coloring indicates areas of low crash density, yellow moderate density, and red high density. The kernel density revealed the spatial patterns of railroad-trespassing crashes in the

United States from 2004 to 2013. On the whole, the highest crash densities are in mega-regions such as the Northeastern New York-Boston Area, Great Lakes, Southern and Northern California, Piedmont Atlantic Area, Cascadia Area, and the Texas Triangle plus Gulf coast areas. These densities are likely related to the surrounding population, train frequencies and trespassing exposure. However, the focus of this study is on correlates of injury severity given a crash. Therefore crash frequency and its relationship with exposure/geometry are not modeled. Also, information about train exposure variables is not commonly available in FRA safety databases. The variation in spatial distributions underscores the need for exploring variations in correlates of injury severity in space.

SECTION 5: MODEL RESULTS

After checking the correlation between variables, the global (traditional) logistics regression models are estimated and compared with corresponding GWLR (Table 2). Variables with positive coefficient values contribute to higher probability of a fatality, given a crash. Both models are statistically significant overall (0.05% level). However, the goodness-of-fit for global and local models are relatively low, indicating limited explanatory power for the crash injury data. This may be partly because some of the key explanatory variables that are associated with crash injury severity, e.g., train speed, are missing.

The global model shows that:

- Most of the pre-crash actions were found to be significantly associated with trespassing injury severity. Among them, lying and sleeping shows statistically significant positive associations with probability of fatality compared with the base (“other pre-crash actions” than those listed in Table 1). The odds of death for lying on track are 60%, $[\exp(\beta)-1]*100\%$, higher than the “base.”
- Pre-crash actions such as climbing, jumping, stepping into or off the track, riding (bicycles), crossing or crawling (over tracks) are less likely to be associated with fatal crashes relative to the base. They account for almost one-half of trespassers. Besides these actions that do not involve a motor vehicle, 5.5% of pre-crash actions involved driving vehicles on or along tracks (in close proximity to the tracks so it was categorized as trespassing).
- Driving, riding, climbing, jumping on tracks or even crossing over tracks have higher mobility compared with more static pre-crash actions (lying down) so it would be easier for these trespassers to take evasive action in a collision. Therefore, they are less likely to be involved in fatal crashes. The odds of driving a motorized vehicle are 73% lower for fatal crashes compared with the base. One possible reason for this driving action is that drivers may get out of their vehicle if the vehicle gets stuck on the tracks (Liu et al., 2015b).
- Sitting or standing, bending, stooping or running/ walking on tracks do not show statistically significant associations with injury in the global logistic model. However, it is still possible for them to be associated with injuries in local GWR models.
- For the higher risk population groups, individuals who are older than 65 years (3.6% of all involved in trespassing crashes) show a higher chance of fatality compared with the base group (30-39 year old). This is in line with expectations, considering seniors maybe more fragile and more easily injured. Unexpectedly, individuals who are

younger than 16 years (6.5% of all involved) show a lower probability of fatality compared with adults. Notably, trespassers younger than 16 years were more likely to be involved in pre-crash actions of climbing, jumping and stepping than the base age group (Pearson $\chi^2=89.47$, p-value = 0.000). Accordingly younger individuals had lower injury severity, on average.

- Yard trespassing (3.7% of all trespassing) is associated with a lower chances of fatality compared with those occurring on regular tracks. Using similar logic, i.e., examining correlations with pre-crash actions, this study found that trespassing crashes in yard areas were more likely to involve climbing, jumping, stepping, riding, operation, crossing, crawling, and driving than in other places (χ^2 test is statistically significant at 95% level). These actions are associated with a lower injury severity.
- One-half of trespassing cases occurred at night, but on average they are less likely to be related with a fatal outcome. This result is also related to the pre-crash actions.
- No statistically significant evidence was found relating injury severity with surrounding land use mix.

Although the conventional (global) Logit regression model explains the average correlation between trespassing injury severity and pre-crash action types, involved persons' attributes, and location and temporal factors, the relationships uncovered by the conventional model described above is a "global" point of view, which means the model assumes the correlates do not vary in space. However, geographical and social contexts vary substantially across the United States. Therefore, differences in associations of the coefficients across space are expected. GWR can uncover local spatial variation since it allows the regression parameters to vary across space.

Comparing the AIC between the global and local models, the local model outperforms its counterpart, as the AIC values for the local model are substantially lower than the global model. As a general rule, improvements in the AIC that are less than three (3) in value could easily arise as a result of sampling error (Fotheringham et al., 2002), while here the difference between the global and local models is substantially greater than three, indicating that the local models are statistically better than the global model. As shown in Table 2, the local GWLR models provide ranges for coefficients. One criterion to tell the significance of spatial variations compares the difference between the upper quartile and lower quartile of the parameter in the local model along with twice the standard error of that parameter in the global model. Accordingly all the independent variables show that their associations with trespassing fatality vary significantly over the United States (indicated by "TRUE" in the table).

Based on the local parameter estimates, a set of parameter surfaces were generated to map the spatial variations of major independent variables. An Inverse Distance Weighted (Tidwell Jr and Humphreys) interpolation algorithm assigned values to unknown points based on the 8,794 known trespassing crash locations, creating a continuous coefficient surface covering the whole country. IDW assumes that each measured point has a local influence that diminishes with distance; higher weights are given for locations closer to the prediction location than those farther away. Also, the parameter contours are generated based on the IDW interpolation results which can better visualize the parameter values in space. T-statistic surfaces are created and shown together with the parameter surface.

Figure 3 demonstrates the spatial distributions of the local parameter estimates for the higher risk population groups including (a) youths (age less than or equal to 16 years) and (b)

seniors (age equal to or greater than 65 years), as well as (c) trespassing fatalities in railway yard. The global models show that individuals who are younger than 16 years have a lower probability of fatality compared with the base trespass group (30-39 years old). However, the local GWLR models show that trespassing crashes in Northern California and Nevada are more likely to be fatal for trespassers younger than 16 years old, as well as crashes in south New Mexico, South West Virginia, west part of Virginia and North Carolina. The fatality odds for youths in rail-trespassing crashes in Northern California and Nevada are 49% to 120% higher than the base trespass group. Based on empirical evidence from the spatial model, the trespassing issues in these areas should be given more attention when it comes to youths. Note that, in the global model, the parameter estimation of youths is -0.294 (odds ratio is 25% lower than the base), but this parameter changes between -1.189 (-70%) and 0.795 (120%) in the local model, representing substantial variation across the country. Moreover, the coefficient map of seniors shows that consideration should be also given to the West and East Coast areas, as well as Eastern Missouri and Southern Illinois. In these areas, senior trespassers are even more likely to be killed than other areas, given a crash. Also, the odds of fatality for seniors in rail-trespassing crashes in Northern California and Nevada are 230% to 570% higher than the base trespasser group.

In terms of their locations, trespassing crashes in railway yards are also associated with a spatially varying fatality chance. The coefficient map in Figure 3 (c) shows that only the positive estimates are statistically significant in the local GWLR models. Yard trespassing crashes in the States of Washington, Kansas, Georgia, Florida, South Carolina and western New York, are less likely to be fatal than in other states.

Figure 5 shows the parameter maps for each pre-crash action type. Note that the dark grey color in the maps represents those areas showing no locally statistically significant association with trespassing injury severity (at 95% level). This highlights that, although a parameter can be significant or insignificant in the global model, the associations are not necessarily the same throughout the study area. Therefore, using a global logistic model may generate information in terms of statistical significance that may not be true in all areas, and it can potentially hide detailed information on spatial distribution of the associations.

An important finding is that the pre-crash action types show substantial spatial variations in their associations with the probability of a trespassing fatality. For instance, the coefficients of lying remain almost positive across the United States, while the coefficients have higher values in the Great Lakes area (Illinois and Indiana), and in New Mexico. That means that lying on tracks has a higher chance to be fatal in these areas, relative to the “base.” Specifically, the odds of being involved in fatal crash are substantially higher, by 5.7 times (or 570 %), for the Great Lakes area, compared with the base action type (listed as “other action types” in the database, and includes actions such as exercising on or near tracks at the time of collision). Although, the database used in this study does not link the behavior of lying on tracks to suicides, it is still suspected that a significant number of the trespassers in these regions were likely attempting suicide, based on the literature showing that about 20% of such actions are suicides (Cadle Creek Consulting, 2008) and given the mixed and dense population in these areas and the economic decline in the last decade. The Great Lakes area has a dense railway network with substantial train activity, which seems to provide an option to people who wish to end their life by lying on train tracks. In addition, some trespassers who are lying on tracks may be drunk and passed out. Firm reasons for such differences in outcomes need further investigation, with consideration given to various factors that include sociodemographic, economic, cultural, and environmental, and intentional vs. unintentional harm variables.

For running or walking on (or near) tracks, the global model does not provide support for its significant association with a trespassing fatality. However, local models provide more nuanced explanations, e.g., in the Northwest and Southeast, as well as West Coast areas, running or walking on (near) tracks are less likely to involve fatal crashes while for the Midwest to Great Plains, plus the New Mexico, this type of pre-crash actions are associated with higher likelihood of fatality, given a crash. The difference might be because of the train operation speeds. The train speed in flat areas may be higher than in rolling and mountain areas. Higher speeds are often associated with an increased injury severity. Running or walking shows both positive and negative associations with a trespassing fatality. This highlights the fact that the traditional model can mask important spatial variance existing in both the signs and magnitudes of coefficients, while spatially disaggregated regression methods (e.g., GWR) provides a good remedy.

For other pre-crash actions such as climbing, riding, crossing or driving, the signs remain negative in almost all over the country with different significance levels. Table 3 summarizes the change of odds of trespassing fatality for different pre-crash action types in various mega-regions across the US. All the odds shown in Table 3 represent chances of fatality compared with particular pre-crash action type relative to “base” pre-crash action (coded as ‘Other action’ in table 1 that include all pre-crash actions which cannot fit in the action categories, such as exercise on or near tracks). Figure 4 illustrates where these mega-regions are located in the US. The table provides a complete big picture view of how different pre-crash actions are related to deadly railway trespassing crashes. As different metropolitan areas show substantially different impacts of different pre-crash actions on non-crossing rail-trespassing fatalities, strategy and policy inferences that suit local settings should be considered. Notably, some of the regional differences may reflect the different geography of the regions. For example, Eastern portions of the US have relatively older, larger and more extensive rail networks (including frequent commuter service) that are located in more densely-populated areas.

SECTION 6: LIMITATIONS

The GWR family of models (including GWLR) is “data-hungry” requiring intensive computations. Models in this paper were calibrated using the GWR 4 package, which took more than 4 hours with a workstation level computer. This is much longer than estimation of a conventional logistic model. The model itself largely depends on the accuracy of location information since the spatial weights in the model are based on physical locations in space. The data of years prior to 2011 were geo-imputed into the centroid of the county where the trespassing occurred while the data after 2011 were coded into their actual location (note that there are 3,144 counties and county equivalents in the US). This geo-imputation will impact the accuracy of the models, and it is recognized as a limitation. However, considering that the model covers a relatively large space (the entire country) and this study uses a relative large kernel size (500 crashes) for GWLR estimation, the extent of errors in terms of stationarity is likely to be small.

Another limitation of this study may be the issue of under-reporting of less severe trespassing crashes. FRA Guide for Preparing Accident/Incident Reports indicates that the reported incident/accidents should satisfy the reporting threshold in stipulated dollars, ranging from \$6,700 to \$8,200 (FRA, 2011). If less severe crashes are under-reported, then the models

estimated in this study may predict a slightly higher probability of fatal trespassing crashes than it actually is.

The estimated models may have limited explanatory power due to missing variables. Unfortunately, some key explanatory variables which are usually reported in crossing crash reports were not available for track trespassing crashes. This is partially because the railway tracks are usually open facilities with no specially designated management equipment which makes it difficult to collect appropriate data, e.g., tracking train speeds.

This study explores rail-trespassing injury by emphasizing fatal and injury crashes, further expansion of the study can consider other measures to assess the harm of the rail-trespassing crashes, e.g., the monetary costs of crashes. Also, correlation among explanatory variables is a concern. Variance inflation factors were calculated for the explanatory variables and the ones used in the study were less than 10, which is the threshold at which multi-collinearity among explanatory variables becomes problematic.

SECTION 7: CONCLUSIONS AND RECOMMENDATIONS

Rail trespassing crashes at non-crossings has received less attention compared with highway-rail grade crossing crashes, which has received wider attention, e.g., the Rail Safety Improvement Act of 2008 (US Congress 2008) has one chapter discussing highway-rail grade crossing safety issues. Such crashes are an increasing concern, and it is certainly not an easy task to prevent trespassing on nearly 140,000 route-miles of rail tracks. Nevertheless, trespassing on tracks in non-crossing locations is not a trivial issue. On average, one person is killed for every 300 rail route-miles per year. The proportion of fatalities (at more than 50%) is much higher than crashes that occur at highway-rail grade crossings (where the comparable proportion of fatalities is 25% for all incidents based on FRA 2004-2013 data, after excluding property-damage only crashes). This research is timely because it highlights trespassing crashes that occur on non-crossing rail track locations and it should be of interest to researchers, practitioners, and the general public.

While researchers have focused on understanding factors associated with rail-trespassing crashes, most of them concentrate on crashes that occurred at crossings and few studies have explored crashes caused by trespassing along railway tracks/right-of-way. This paper fills a gap by using a fairly comprehensive database to investigate an under-researched area— railway track trespassing crashes at non-crossings—and applying a novel location-sensitive modeling methodology to uncover the relationships between trespassing crash injuries and their associated factors. The study takes advantage of the recent developments in geo-referenced data and higher computational power of the computers.

The study has the following key findings: 1) higher rail-trespassing fatality chances, given a crash, are associated significantly with pre-crash behaviors, especially lying and sleeping on tracks; seniors are more likely to be killed in rail-trespassing crashes while youths are less likely to be killed compared with adults; rail-trespass crash injury severity was not statistically significantly (5% level) related with land use and weekend or weekday; 2) a critical finding is uncovering spatial variations in the correlates, given the variations in spatial locations of crashes. That is to say, the coefficients are not identical (stationary) over space but vary across the country. Therefore, relationships between injuries and certain variables from one location cannot be generalized to other locations. Fundamentally, rail-trespass crashes are complex and nuanced

and the basic assumption of independence of spatial proximity in traditional regression models does not fully hold in the situation explored.

Using GWLR, local models provide maps of coefficients, which visualize the associations by showing where certain pre-crash actions are associated with higher probability of trespassing crash injury. The study provides valuable information that can help rail safety stakeholders focus their attention on critical factors in geographies that show higher risks of fatalities in crashes. Due to the limitations of the crash data, it is difficult to draw firm conclusions about how to design treatments to prevent railway track trespassing crashes or reduce their intensity. However, the national level spatial models show an intuitive big picture of railway track trespassing crash facts in the United States. Some areas become prominent based on spatial model results, e.g., lying (some of which are potential suicides) and sitting or standing on the tracks has the highest probability of fatality in the Great Lakes area. Senior trespassers have a higher probability of fatality in Northern California. Such findings should motivate studies concentrating on these particular areas to further explore the issue. Far more detailed information including social, economic, land use, community crime rate, demographics can be incorporated in such case studies to identify reasons for such outcomes. The case studies would complement a national study by providing more detailed suggestions on effective design of treatments in order to avoid rail-trespasser crashes at non-crossings. The study will investigate effectiveness of various countermeasures that include adding fences in critical areas or safety alert signs, surveillance, and public education. Notably, countermeasures sometimes come with certain issues, e.g., fencing can be subject to vandalism or may relocate a person to a non-fenced area such as a station platform. Overall, there is a need to study effective countermeasures that can be targeted regionally to specific trespassing behaviors that have the highest risk.

SECTION 8: REFERENCES

- Cadle Creek Consulting. 2008. Rail Trespasser Fatalities, Developing Demographic Profiles. <https://www.fra.dot.gov/eLib/Details/L02669>. Accessed 11-01 2014.
- Centers for Disease Control Prevention. 1999. Injuries among railroad trespassers--Georgia, 1990-1996. *MMWR: Morbidity and Mortality Weekly Report* 48, 537-541.
- Cina, S., J. Koelpin, C. Nichols, and S. Conradi. 1994. A decade of train-pedestrian fatalities: the Charleston experience. *Journal of forensic sciences* 39, 668-673.
- Evans, A. W. 2003. Accidental fatalities in transport. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 166, 253-260.
- Famoye, F., J. Wulu, and K. P. Singh. 2004. On the generalized Poisson regression model with an application to accident data. *Journal of Data Science* 2, 287-295.
- Fotheringham, A., C. Brunsdon, and M. Charlton. 2002. Geographically Weighted Regression, the analysis of spatially varying relationships. John Wiley & Sons, Ltd.
- Federal Railroad Administration. 2011. FRA Guide for Preparing Accident/Incident Reports, Federal Railroad Administration, U.S. Department of Transportation, Washington, D.C.
- Gabree, S., S. Chase, A. Doucette and M. Martino, Countermeasures to Mitigate Intentional Deaths on Railroad Rights-of-Way: Lessons Learned and Next Steps. 2014. <http://www.fra.dot.gov/Elib/Document/14240>. Accessed 6-22 2015.
- GeoDa. 2015. Land Use Mix, 2001, in counties, PCSAs, and ZCTAs. https://geodacenter.asu.edu/drupal_files/Land%20use%20mix%20calculations.pdf.

Accessed 6-22 2015.

- Hadayeghi, A., A. Shalaby, and B. Persaud. 2009. Development of Planning Level Transportation Safety Tools Using Geographically Weighted Poisson Regression. In: *TRB Annual Meeting CD-ROM*, Washington, D.C.
- Khattak, A., and Z. Luo. 2011. Pedestrian and Bicyclist Violations at Highway-Rail Grade Crossings. *Transportation Research Record: Journal of the Transportation Research Board* 2250, 76-82. doi:10.3141/2250-10.
- Lerer, L., and R. Matzopoulos. 1996. Meeting the challenge of railway injury in a South African city. *The Lancet* 348, 664-666.
- Levine, N., K. Kim, and L. Nitz. 1995. Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns. *Accident Analysis & Prevention* 27, 663-674. doi :10.1016/0001-4575(95)00017-T.
- Liu, J., B. Bartnik, S. H. Richards, and A. J. Khattak. 2015a. Driver Behavior at Railway-Highway Grade Crossings with Passive Traffic Controls: A Driving Simulator Study. *Journal of Transportation Safety and Security*. doi:10.1080/19439962.2015.1043478.
- Liu, J., A. Khattak, S. Richards, and S. Nambisan. 2015b. What are the Differences in Driver Injury Outcomes at Highway-Rail Grade Crossings? Untangling the Role of Pre-Crash Behaviors. In: *94th Annual Meeting of the Transportation Research Board* Washington D.C.
- Liu, R., Z. Huang, and W. Hao. 2011. Potential Issues and Challenges in Implementing the Railroad Safety Improvement Act of 2008 Based on a National Survey. *Journal of Transportation Safety & Security* 3, 252-271.
- Lobb, B. 2006. Trespassing on the tracks: a review of railway pedestrian safety research. *Journal of safety research* 37, 359-365.
- Lobb, B., N. Harre, and T. Suddendorf. 2001. An evaluation of a suburban railway pedestrian crossing safety programme. *Accident Analysis & Prevention* 33, 157-165.
- Loo, B. 2009. The Identification of Hazardous Road Locations: A Comparison of the Blacksite and Hot Zone Methodologies in Hong Kong. *International Journal of Sustainable Transportation* 3, 187-202. doi:10.1080/15568310801915583.
- McCulloch, C. E., and J. M. Neuhaus. 2001. Generalized linear mixed models. Wiley Online Library.
- Metaxatos, P., and P. Sriraj. 2013. Pedestrian/Bicyclist Warning Devices and Signs at Highway-Rail and Pathway-Rail Grade Crossings Urban Transportation Center, University of Illinois at Chicago.
- Millegan, H., X. Yan, S. Richards, and L. Han. 2009. Evaluation of Effectiveness of Stop-Sign Treatment at Highway–Railroad Grade Crossings. *Journal of Transportation Safety & Security* 1, 46-60.
- Mohanty, M., M. Panigrahi, S. Mohanty, and K. Patnaik. 2007. Death due to traumatic railway injury. *Medicine, science and the law* 47, 156-160.
- Mok, S., and I. Savage. 2005. Why Has Safety Improved at Rail - Highway Grade Crossings? *Risk Analysis* 25, 867-881.
- Páez, A., and D. Scott. 2005. Spatial Statistics for Urban Analysis: A Review of Techniques with Examples. *GeoJournal* 61, 53-67.
- Park, S., D. Kim, S. Kho, and S. Rhee. 2010. Identifying Hazardous Locations Based on Severity Scores of Highway Crashes. In: *12th World Conference on Transport Research*, Lisbon

- Pelletier, A. 1997. Deaths among railroad trespassers: the role of alcohol in fatal injuries. *JAMA* 277, 1064-1066.
- Plug, C., J. Xia, and C. Caulfield. 2011. Spatial and temporal visualisation techniques for crash analysis. *Accident Analysis & Prevention* 43, 1937-1946.
doi:<http://dx.doi.org/10.1016/j.aap.2011.05.007>.
- Russo, B., and P. Savolainen. 2013. An Examination of Factors Affecting the Frequency and Severity of Crashes at Rail-Grade Crossings. In: *Transportation Research Board 92nd Annual Meeting*
- Savage, I. 2006. Does public education improve rail-highway crossing safety? *Accident Analysis & Prevention* 38, 310-316.
- Savage, I. 2007. Trespassing on the Railroad. *Research in Transportation Economics* 20, 199-224.
- Silla, A., and J. Luoma. 2011. Effect of three countermeasures against the illegal crossing of railway tracks. *Accident Analysis & Prevention* 43, 1089-1094.
- Silla, A., and J. Luoma. 2012a. Main characteristics of train-pedestrian fatalities on Finnish railroads. *Accident Analysis & Prevention* 45, 61-66.
- Silla, A., and J. Luoma. 2012b. Opinions on railway trespassing of people living close to a railway line. *Safety science* 50, 62-67.
- Silverman, B. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, New York.
- US Congress. 2008. Public Law 110-432 - Rail Safety Improvement Act of 2008.
<https://www.fra.dot.gov/Page/P0395>. Accessed 8-22 2014.
- Van Houwelingen, C., and D. Beersma. 2001. Seasonal changes in 24-h patterns of suicide rates: a study on train suicides in The Netherlands. *Journal of affective disorders* 66, 215-223.
- Wang, X., and A. Khattak. 2013. Role of travel information in supporting travel decision adaption: exploring spatial patterns. *Transportmetrica*, 1-19.
doi:10.1080/18128602.2011.577041.
- Ye, X., R. M. Pendyala, V. Shankar, and K. C. Konduri. 2013. A simultaneous equations model of crash frequency by severity level for freeway sections. *Accident Analysis & Prevention* 57, 140-149.
- Ye, X., R. M. Pendyala, S. P. Washington, K. Konduri, and J. Oh. 2009. A simultaneous equations model of crash frequency by collision type for rural intersections. *Safety Science* 47, 443-452.

LIST OF TABLES AND FIGURES

Table 1. Descriptive Statistics for Non-Crossing Rail-Trespassing Events.....	22
Table 2. Global and Local Models for Non-Crossing Rail-Trespassing Crash Injury Severity	23
Table 3. Odds of Fatality in Non-Crossing Rail-Trespassing Crashes for Different Pre-Crash Behaviors (Relative to Base Pre-Crash Action) in Major Mega-Regions.....	24
Figure 1. Different Risky Actions of Non-Crossing Railway-Trespassing. Source: Internet.....	25
Figure 2. Kernel Density Distribution for Non-Crossing Rail-Trespassing Crashes	26
Figure 3. Local Parameter Estimates for Non-Crossing Rail-Trespassing Crashes in Higher Risk Population Groups	27
Figure 4. Mega-Regions for This Study.....	28
Figure 5. Local Parameter Estimates for Non-Crossing Rail-Trespassing Crashes Associated with Pre- Crash Action Types.....	29

Table 1. Descriptive Statistics for Non-Crossing Rail-Trespassing Events (N=8,794 Observations)

Variable		N	Mean/Percent	Std.	Min	Max
Injury Severity (0-other, 1-fatal)		459	52.19%	0.500	0	1
Age	<=16 years old	572	6.50%	0.247	0	1
	17-29 years old	245	27.91%	0.449	0	1
	30-39 years old	249	28.35%	0.451	0	1
	40-54 years old	234	26.62%	0.442	0	1
	55-64 years old	619	7.04%	0.256	0	1
	>=65 years old	315	3.58%	0.186	0	1
Env.	Darkness (0-no, 1-yes)	449	51.16%	0.500	0	1
Time	Weekend (0-no, 1-yes)	295	33.64%	0.472	0	1
	Summer (0-no, 1-yes)	259	29.55%	0.456	0	1
	Winter (0-no, 1-yes)	168	19.10%	0.393	0	1
	Spring or Autumn (0-no, 1-yes)	451	51.34%	0.500	0	1
Location attributes	Land Use Mix Index #		0.419	0.280	0.00	0.98
	Railway Yard (0-no, 1-yes)	326	3.71%	0.189	0	1
Pre-crash trespasser actions	Climbing, jumping, stepping	753	8.56%	0.280	0	1
	Riding, operation	468	5.32%	0.224	0	1
	Lying, sleeping (on or near	205	23.35%	0.423	0	1
	Running, walking	308	35.07%	0.477	0	1
	Crossing, crawling (over tracks)	215	2.44%	0.154	0	1
	Sitting, standing, bending,	146	16.60%	0.372	0	1
	Driving	486	5.53%	0.229	0	1
	Other action (0-no, 1-yes) *	275	3.13%	0.174	0	1

Notes:

- # Land Use Mix Index was developed by Environmental Protection Agency (EPA) to capture the land use mix entropy level, which varies from 0 (homogeneous land use, such as in rural areas or suburban subdivisions) to 1 (most mixed, such as diverse city centers) land use (GeoDa, 2015).
- * Other pre-crash actions include those action types which cannot fit in the categories listed in the table, e.g., exercising.

Table 2. Global and Local Models for Non-Crossing Rail-Trespassing Crash Injury Severity

Models →		Global Model (binary logit)		Global Model (dropped insignificant variables)			Local GWLR Model					
Variables		β	P> t	β	SE	P> t	Min	Max	Lwr Quartile	Upr Quartile	(Upr-Lwr)>2SE	
Constant		0.251	0.069	0.261	0.131	0.045	-1.832	1.486	-0.369	0.669	TRUE*	
Age (Base: 30-39 years old)	<=16 years old	-0.294	0.003	-0.294	0.100	0.003	-1.189	0.795	-0.491	0.107	TRUE	
	17-29 years old	-0.007	0.909	-0.007	0.060	0.906	-0.546	0.670	-0.107	0.130	TRUE	
	40-54 years old	0.185	0.002	0.186	0.061	0.002	-0.149	0.841	0.027	0.363	TRUE	
	55-64 years old	0.250	0.008	0.250	0.095	0.008	-0.517	1.294	0.075	0.529	TRUE	
	>=65 years old	0.641	0.000	0.641	0.131	0.000	-0.175	2.094	0.412	0.949	TRUE	
Env. attributes	Darkness (0-no, 1-yes)	-0.110	0.019	-0.110	0.047	0.018	-0.543	0.288	-0.185	-0.024	TRUE	
	Weekend (0-no, 1-yes)	0.005	0.917	Dropped								
Time	Summer (Base: Spring and Autumn)	-0.142	0.006	-0.142	0.052	0.006	-0.812	0.352	-0.290	-0.038	TRUE	
	Winter (Base: Spring and Autumn)	0.053	0.383	0.053	0.060	0.383	-0.325	0.370	-0.102	0.149	TRUE	
Location attributes	Land Use Mix Index	0.017	0.833	Dropped								
	Railway Yard (0-no, 1-yes)	-1.037	0.000	-1.036	0.142	0.000	-4.765	0.359	-1.258	-0.652	TRUE	
Pre-crash trespasser actions (Base: Other actions***)	Climbing, jumping, stepping	-1.357	0.000	-1.359	0.151	0.000	-4.213	0.537	-1.811	-0.850	TRUE	
	Riding, operation	-1.032	0.000	-1.035	0.160	0.000	-2.596	0.605	-1.541	-0.564	TRUE	
	Lying, sleeping (on or near tracks)	0.469	0.000	0.467	0.132	0.000	-1.113	2.468	-0.107	1.194	TRUE	
	Running, walking	-0.023	0.857	-0.024	0.128	0.850	-1.435	1.907	-0.523	0.493	TRUE	
	Crossing, crawling (over tracks)	-0.766	0.000	-0.767	0.191	0.000	-4.605	2.227	-1.412	-0.461	TRUE	
	Sitting, standing, bending, stooping	0.171	0.203	0.170	0.134	0.206	-1.364	1.968	-0.292	0.740	TRUE	
	Driving	-1.297	0.000	-1.299	0.162	0.000	-3.371	0.548	-1.828	-0.497	TRUE	
Summary Statistics		Sample Size: 8794		Sample Size: 8794			Local Sample Size**: 500					
		Log likelihood = -5642.09		Log likelihood = -5642.12			Pseudo-R ² = N/A, % deviance explained: 0.115					
		Prob > χ^2 = 0.00		Prob > χ^2 = 0.00			AIC = 11308.19					
		Pseudo-R ² = 0.073		Pseudo-R ² = 0.073								
		AIC = 11322.19		AIC = 11318.24								

Notes: 1. * True means the significance of spatial variance of coefficient.

2. ** Local sample size is the number of subsamples used in each kernel; 500 local closest surrounding trespassing cases were used as the subsample for these regressions. This kernel size was chosen based on balancing model performance and conceptual considerations. On average, one mega-region in the US contains between 400-700 trespassing cases over the past decade.

3. ***Other actions may include activities like exercise on or near tracks.

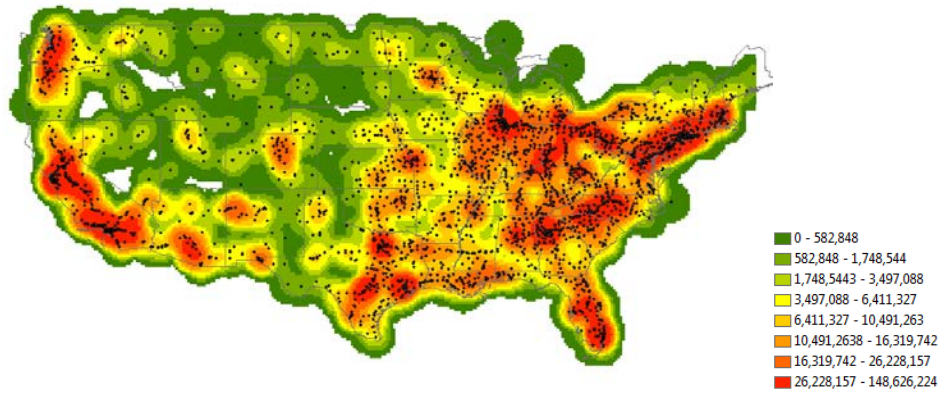
Table 3. Odds of Fatality in Non-Crossing Rail-Trespassing Crashes for Different Pre-Crash Behaviors (Relative to Base Pre-Crash Action) in Major Mega-Regions

Regions	Climbing, jumping, stepping	Riding Operation	Lying on or near tracks	Running Walking	Crossing	Sitting or Standing	Driving	Other pre-crash actions (base)
Northeastern	-75%	-90%	-	-14%	-55%	-	-77%	0
Great Lakes	-45%	-	570%	172%	-55%	350%	80%	0
Northern CA and NV	-90%	-90%	-	-63%	-83%	-	-86%	0
Southern CA	-75%	55%	-	-14%	-45%	-	-	0
Piedmont Atlantic	-98%	80%	-	-63%	-90%	-	-95%	0
Cascadia Area	-63%	-52%	82%	-	-63%	-	-	0
New Mexico	-	-	1000%	420%	-	500%	-	0
Gulf Coast	-80%	-74%	-	-53%	-70%	-	-92%	0
Florida	-94%	-74%	-	-65%	-70%	-	-80%	0

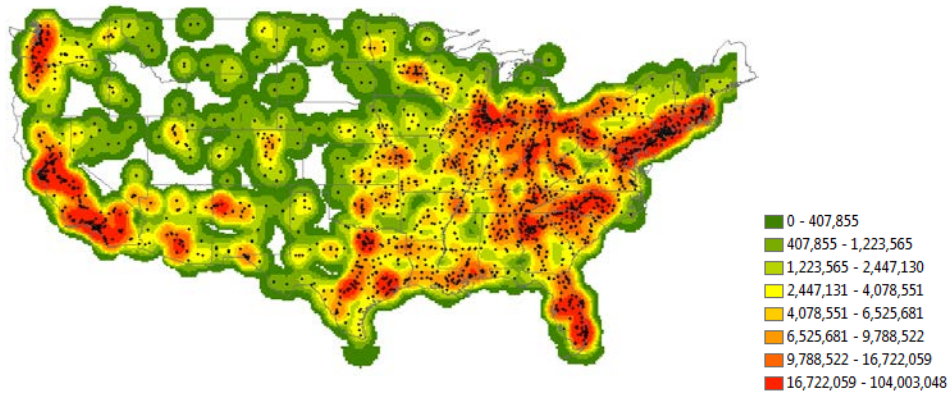
Notes: “-“means no statistical significant associations (95% level) found in such area. These regions are only based on the results of GWR modeling and do not represent (Control and Prevention)defined mega-regions.



Figure 1. Different Risky Actions of Non-Crossing Railway-Trespassing. Source: Internet



(a) Total Crashes-Fatal and Non-Fatal ($N= 8,794$)



(b) Fatal Crashes ($N=4,204$)

Figure 2. Kernel Density Distribution for Non-Crossing Rail-Trespassing Crashes

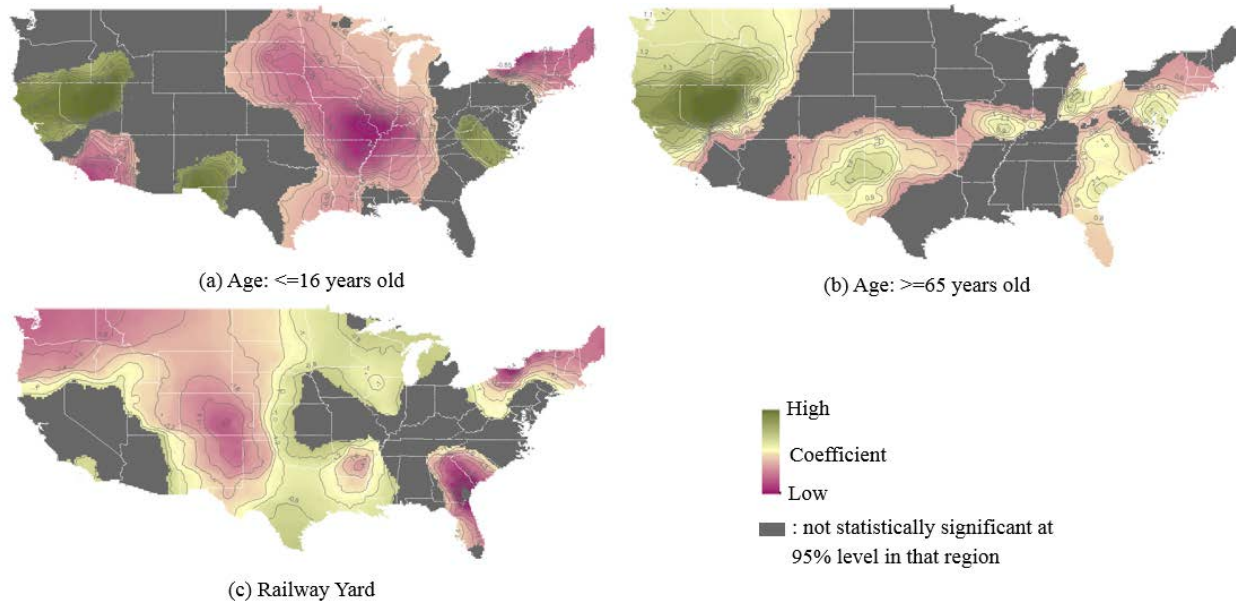


Figure 3. Local Parameter Estimates for Non-Crossing Rail-Trespassing Crashes in Higher Risk Population Groups



Figure 4. Mega-Regions for This Study

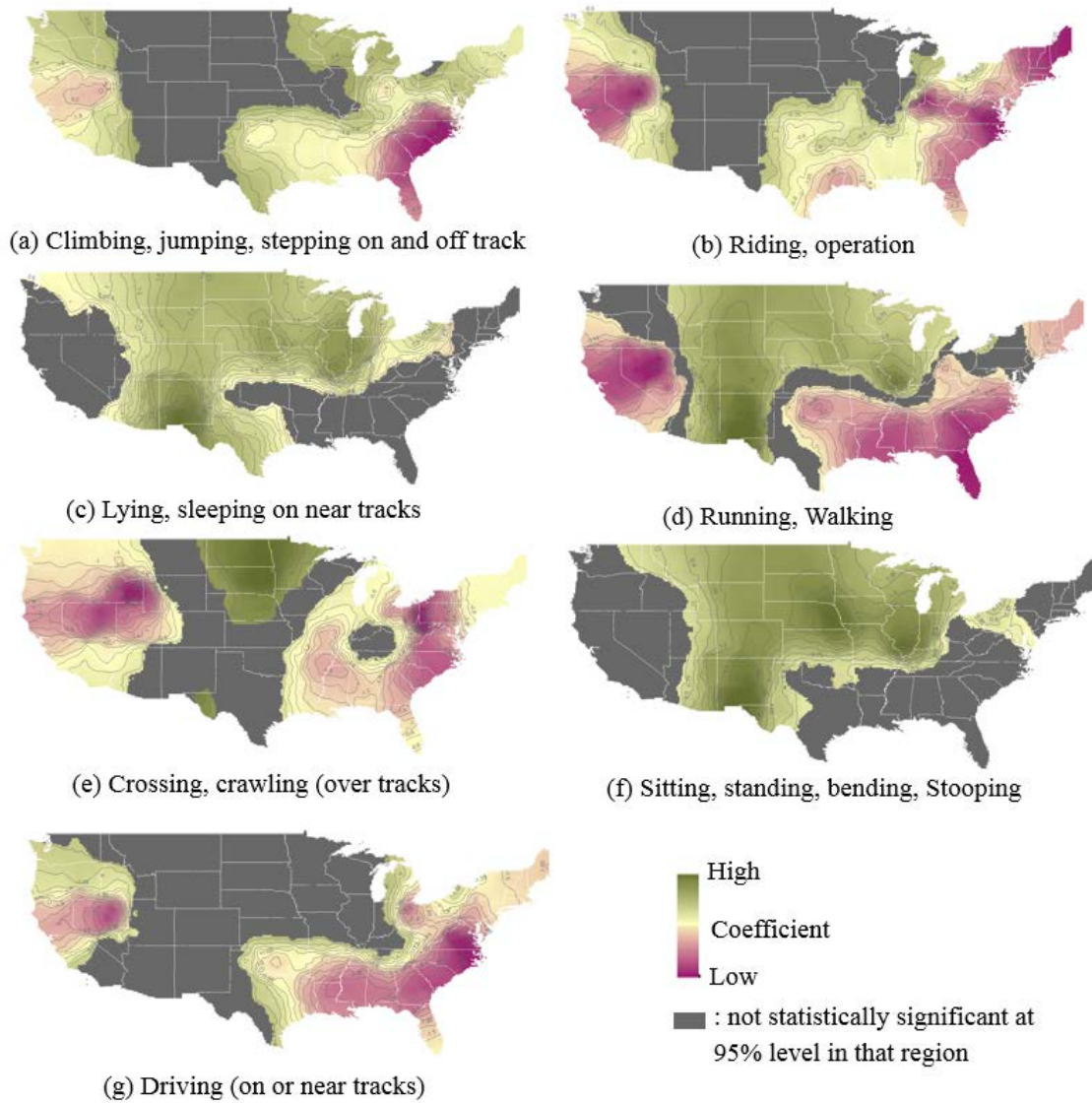


Figure 5. Local Parameter Estimates for Non-Crossing Rail-Trespassing Crashes Associated with Pre-Crash Action Types

A COMPARATIVE STUDY OF RAIL-PEDESTRIAN AND BICYCLIST TRESPASSING CRASH INJURY SEVERITY AT HIGHWAY-RAIL GRADE CROSSINGS AND NON-CROSSINGS

Meng Zhang, Asad Khattak, Jun Liu, and David Clarke
The University of Tennessee, Knoxville

Abstract - This study examined rail-pedestrian and cyclist trespassing crash injury severity in two different settings, highway-rail grade crossings and non-crossings, by using railroad injury and illness incidents data. It explores how pre-crash trespassing behaviors and other factors (e.g., crash time, location and social demographic) differ between highway-rail grade crossings and non-crossings. The analysis used ten years (2005-2014) of rail-pedestrian and bicyclist trespassing crash records extracted from a Federal Railroad Administration (FRA) safety database. The total number of rail-pedestrian and bicyclist trespassing crashes used for analysis were 10,146; of this number, 8,258 (81% of the total) are non-crossing trespassing crashes and 1,888 (19%) are highway-rail grade crossing trespassing crashes. At highway-rail grade crossings, 43.4% of the crashes were fatal, whereas at non-crossings, 53.4% of the crashes involved a fatality. The most prevalent pre-crash trespassing behavior is riding (a bicycle or other non-motorized transport) or operating equipment (51.5%) at highway-rail grade crossings. Lying or sleeping and running or walking account for about 61% of non-crossing crashes. To explore relationships, this study applied separate and pooled models for injury severity in highway-rail grade crossing and non-crossing trespassing crashes. To account for the ordinal nature of severity data, ordered logistical regression models (proportional odds models) were estimated. Next, the proportional odds assumption was tested, and the results of partial proportional odds models were provided. Modeling results show that the correlates of injury severity differ across highway-rail grade crossings and non-crossings. For example, lying or sleeping accounted for a higher probability of fatal injury in rail-pedestrian and cyclist trespassing crashes in both settings-notably such crashes were even more injurious at grade crossings than at non-crossings. The role of pre-crash behaviors was explored and the implications are discussed.

Keywords: rail-trespassing crash; injury severity; crossing; non-crossing; pre-crash action; comparative analysis

SECTION 1: INTRODUCTION

For railway safety, rail-pedestrian crashes are a major concern, as rail-pedestrian crashes are a leading cause of fatalities in rail-related crashes (Pelletier 1997, Lobb, Harre et al. 2001, Savage 2007). However, an increasing societal concern is rail-trespasser crashes, especially non-motorized trespassers (pedestrian and bicyclist). In 2014, pedestrian and bicyclist trespassing on rail properties was a fact in more than 500 fatalities and more than 450 injuries (Administration 2015). These incidents annually produce billions of dollars in personal and societal costs (Trottenberg and Rivkin 2013). According to Federal Railroad Administration (FRA) statistics,

rail trespassing fatalities and injuries are rising. With increasing exposure of trespassers and more train activities, rail crashes with pedestrian and bicyclist trespassers will remain a critical concern in the future.

FRA defines trespassers as persons who should not be present on the railway right-of-way. Normally, a person at highway-rail grade crossing would not be recognized as a trespasser unless the person goes around or through crossings with physical barriers (e.g., gates are down). A person deliberately ignoring the barrier in an attempt to cross will also be coded as a trespasser (Safety 2011). Most trespassers are pedestrians, but some are individuals who are driving or riding a bicycle, an all-terrain vehicle (ATV), snowmobile, etc. (George 2008). As there is little in common between motorized trespassers and non-motorized trespassers, this study will focus only on vulnerable, non-motorized trespassers (pedestrian and bicyclist).

The key question to be answered in this research is how pre-crash behaviors and other correlates (e.g., crash time, location and social demographic) of rail-pedestrian and bicyclist trespassing injury severity vary between highway-rail grade crossings and non-crossings. To the best of our knowledge, studies have not compared variations in injury risks at highway-rail grade crossings with non-crossings. A variety of factors, including the trespassers' pre-crash actions, may be associated with injury severity. How these associations vary from highway-rail grade crossings to non-crossing tracks has not been fully explored. Therefore, the objective of this study is to explore and compare differences in injury severity and correlations of pre-crash behaviors and other factors with injury severity. The analysis involved estimation of injury severity models together and separately for highway-rail grade crossings and non-crossings.

SECTION 2: LITERATURE REVIEW

Previous studies investigate rail-related crashes with pedestrians, bicyclists, vehicles and other users (Oh, Washington et al. 2006, Horton, Carroll et al. 2009, Khattak and Luo 2011, Metaxatos and Sriraj 2013, Gabree, Chase et al. 2014, Liu, Khattak et al. 2014, Liu, Bartnik et al. 2015). Examining trespassing motivations can help us understand trespassers' intentions. Studies have shown various types of trespassing motivations (Pelletier 1997, Lobb, Harre et al. 2001, Savage 2007). Choosing a shorter or more convenient path to a destination is a common motivation for trespassers (Lobb, Harre et al. 2001, Safety 2007).

Some studies suggest suicide as a motivation to trespass. According to a European report, more than 3,000 people were killed in train crashes yearly owing to suicides or trespassing (Burkhardt, Radbo et al. 2014). However, determining whether a fatality is suicide or accident is always challenged by inadequate information (Mishara 2007). In addition to inadequate information, the social, legal, financial and ethical implications also complicate identifying rail suicide as a cause of fatality (Lobb 2006). Nonetheless, about 20% to 27% of deaths were recorded as suicides (Gabree, Chase et al. 2014, George 2008). Although the Rail Safety Improvement Act of 2008 (RSIA) requires information about suicides to be collected, such information is not publicly available.

Literature on rail-related crashes also highlights other associated factors, including personal or environmental characteristics, timing and location attributes, trespassers' behaviors (pre-crash behaviors) and countermeasures implemented to prevent trespassing events (Cina, Koelpin et al. 1994, Pelletier 1997, Silla and Luoma 2012).

2.1 Rail-pedestrian crash analysis

Trespasser socio-demographic characteristics are commonly investigated. Individuals belonging to a specific social group (e.g., young males, intoxicated with alcohol) may be frequently involved in a rail-pedestrian fatal crash. Youths and seniors would seem to be more vulnerable as trespassers, though few rail crashes included youths under 10 years old and seniors above 60 years old (Pelletier 1997, Silla and Luoma 2012). Nixon et al. illustrated that young people involved in rail crashes were associated with risk-taking and daring behaviors (Nixon, Corcoran et al. 1985). Most fatalities in rail-pedestrian crashes were young males (Cina, Koelpin et al. 1994, Pelletier 1997, Silla and Luoma 2012, George 2008). Compared to females and seniors, young males tended to lack awareness of dangers for a specific traffic situation (Lobb, Harre et al. 2001). Pelletier reported that trespassing fatalities typically involved individuals who were unmarried males without high school education (Pelletier 1997). Individuals who were intoxicated with problems of alcohol, medicines or drugs were easier contributed to being struck by trains (Silla and Luoma 2009, George 2008).

Previous studies have explored the timing and location of rail-pedestrian crashes. For example, fatal crashes occurred frequently from March to August (Pelletier 1997). Rail-pedestrian crashes have shown temporal clustering. They occur frequently at the end of a week (from Friday to Sunday) and during rush hours (Silla and Luoma 2012). Lerer and Matzopoulos noticed that rail injuries commonly occurred during peak commuting times in the city of Cape Town, South Africa (Lerer and Matzopoulos 1996). As for geographic clustering, most trespasser fatalities tended to be specific to locations like high-density populated areas and rail yards (Matzopoulos and Lerer 1998, Silla and Luoma 2009, Silla and Luoma 2012). Geographically Weighted Logistic Regression (GWRL) model were estimated to investigate the spatial patterns of rail non-crossing trespassing crashes across the United States (Wang, Khattak et al. 2014). However, little information is available about how rail-trespasser crashes patterns differ between grade crossings and non-crossings. This study investigates injury severity levels of rail-trespasser crashes given a crash, focusing on the role of pre-crash behaviors at grade crossings vs. non-crossings.

Trespassers' pre-crash behaviors were found to be highly associated with the consequences of train crashes. Lying or walking on the railroad track were common precursor behaviors (Patterson and Authority 2004, Savage 2007). Several studies showed that most fatal train crashes happened when individuals were walking, sitting or lying on or near to the railway tracks (Cina, Koelpin et al. 1994, Lerer and Matzopoulos 1996, Pelletier 1997). Information on pre-crash behavior will help address reasons of trespassing crashes, such as committing suicide (Savage 2007, Silla and Luoma 2012). A retrospective analysis on suicidal behavior (jumping, lying and wandering) revealed that fatality rate was highest when the victim was lying and lowest when jumping.

2.2 Rail-related injury severity

Injury severity is another concern in rail-trespasser crashes. Published research on rail crash injury severity largely relates to drivers or pedestrians at grade crossing (Fan and Haile 2014, Liu, Khattak et al. 2014, Zhao and Khattak 2015). Fan et al. (2015) estimated the multinomial logit model to explore correlates of injury severity of rail-pedestrian crashes at grade crossing,

showing that fatal crashes are more likely to occur under cloudy weather, on paved highways, with low temperature (<50° F) and in cities. Khattak (2013) applied an ordered logit model to investigate the severity outcome of rail-pedestrian crashes at grade crossings. The results show that high train speed and female pedestrians were associated with higher chances of injury severity, given a crash (Khattak 2013).

Other studies on injury severity analysis have estimated various models that include the ordered probit logit model, random parameter logit model, mixed generalized ordered logit model and generalized logit model. These studies have identified the relationship between safety outcome and related factors (Eluru, Bhat et al. 2008, Hu, Li et al. 2010, Khattak and Luo 2011, Liu, Khattak et al. 2014, Zhao and Khattak 2015).

2.3 Trespasser safety interventions

Some studies have focused on countermeasures that prevent rail-trespasser crashes (Pelletier 1997, Lobb 2006, Savage 2007, Silla and Luoma 2011). Silla et al. illustrated that the effectiveness of countermeasures varies with diversity of trespasser characteristics (Silla and Luoma 2011). They mainly recommended physical barriers (landscaping, fencing) for reducing trespassing, but prohibitive signs was recommended when the sites are not suitable for physical barriers. Punishment, education and communication were found to be effective in preventing rail trespassing behaviors (Lobb, Harre et al. 2003). Lobb et al. found that public education can be more effective than communications but less than punishment (Lobb, Harre et al. 2003). Lobb et al. also applied a self-reported survey to evaluate the effectiveness of educational and environmental countermeasures in New Zealand. The authors further pointed out that the access preventions was an effective strategy (Lobb, Harre et al. 2001).

The relationship between rail-related trespassing crashes and relevant factors is of interest. Especially, no studies have compared injury severity in crossing and non-crossing crashes. In fact, there are only a few studies of injury severity in rail trespassing crashes. Noticing the geographical diversity of rail-trespasser crashes, Wang et al. investigated spatial patterns of trespasser injury severity at non-crossings (Wang, Khattak et al. 2014). However, differences in injury severity correlates from highway-rail grade crossing to non-crossing pedestrian and bicyclist trespassing crashes have not been fully explored. With increasing concern about non-motorized trespassers and given the diversity of rail-pedestrian and bicyclist trespassing crashes across highway-rail grade crossings and non-crossings, it is important to uncover the potentially different correlates of injury severity.

SECTION 3: APPROACH AND METHODOLOGY

3.1 Data Source

The data used in this study were railroad safety information obtained from FRA Office of Safety Analysis Web Site (<https://safetydata.fra.dot.gov/OfficeofSafety/Default.aspx>). Railroad safety information, including accidents and incidents, inventory and highway-rail crossing data, were available to public in this site. The data structure of dataset related to this study was shown in figure 1. A total of 419,164 highway-rail crossing inventories (6180.71) were filed across United

States. Between 2005 and 2014 year, 23,638 highway-rail grade crossing accidents (6180.57) have occurred and 96,285 railroad injury/illness records (6150.55a) were reported. Any reportable individual fatality, injury, or illness that meets one or more of the general reporting criteria listed in FRA report guide (Safety 2011) should be filed in 6180.55a, such as injury to any person that results in medical treatment.

For this study, Railroad Injury/Illness Summary dataset (6180.55a) was used for analysis. Ten years (2005-2014) of rail-trespasser crashes data representing 12,254 incidents (personal level) were extracted from this dataset based on the rule that the type of person whose injury/illness was reported as “Type E = trespassers”. Notably, the trespasser defined in this study will not contain a person at highway-rail grade crossing unless the grade crossing protected by gates or other similar interventions that can be regarded as physical barriers. Thus, a pedestrian or vehicle that accesses grade crossing without a physical barrier will not be defined as a trespasser (Safety 2011). Then based on occurring location (variable “LOCC” in the dataset), these rail-trespassing crashes were classified into three categories: non-crossing trespassing crashes, highway-rail grade crossing trespassing crashes and other-rail crossing trespassing crashes. Note that trespassing crashes occurred at other-rail crossings were not involved in this study due to the low sample size (21 out of 12254 as shown in figure 1). As this study only focus on rail-pedestrian and bicyclist trespassing crashes, rail-motorists trespassing crashes were also exclusive. Finally, this study obtained 10,146 rail-pedestrian and bicyclist incidents after data cleaning and error check. Of these, 8,258 (81%) occurred at non-crossing locations and 1,888 (19 %) at highway-rail grade crossings.

Rail-pedestrian and bicyclist trespassing crash data contains personal attributes (e.g., age of individuals), geographical and temporal characteristics (e.g., location and time), pre-crash behaviors (e.g., lying, sitting, running, climbing) and the nature of injuries (fatal injury and non-fatal injury). The data are of reasonably good quality, with FRA’s guidance on standard procedures for coding variables. As population density was recognized as contribute factor to rail-trespassing crashes (Silla and Luoma 2009, Silla and Luoma 2012), census information including population density, education and income level obtained from United States Census Bureau (<http://www.census.gov/en.html>) were involved in the analysis. For each row in rail-pedestrian and bicyclist trespassing crash dataset, the corresponding census information was coded into the census information of county where the trespassing crash occurred. The data were verified and error checked using descriptive statistics.

3.2 Model Structure

Unlike studies focusing on crash rates or frequency at specific grade crossing locations, this study investigates the correlates of non-motorized trespasser injury severity given a rail-pedestrian and bicyclist trespassing crash. A scale of three categories (Y= 3 is fatal injury – reported as fatality; Y=2 is severe injury – reported as bruise, fracture, amputation and laceration; Y=1 is minor injury – otherwise, i.e. nervous shock) is defined in this study, justifying the estimation of ordered logistical regression model to explore relationships between injury severity and a set of explanatory variables. The predicted probability for each category Y in a simple ordered logit model (proportional odds model) can be calculated as follows (Long and Freese 2006):

$$\Pr(Y = 1|X) = \frac{\exp(\alpha_1 - X\beta)}{1 + \exp(\alpha_1 - X\beta)} \quad (1)$$

$$\Pr(Y = j|X) = \frac{\exp(\alpha_j - X\beta)}{1 + \exp(\alpha_j - X\beta)} - \frac{\exp(\alpha_{j-1} - X\beta)}{1 + \exp(\alpha_{j-1} - X\beta)} \text{ for } j = 2 \text{ to } J - 1 \quad (2)$$

$$\Pr(Y = J|X) = 1 - \frac{\exp(\alpha_{J-1} - X\beta)}{1 + \exp(\alpha_{J-1} - X\beta)} \quad (3)$$

Where,

Y = The injury severity category level from $j = 1$ to J ;

X = A vector of explanatory variable in the model, e.g., pre-crash behavior;

α = Cut points (thresholds) for injury severity level from $j = 1$ to $J - 1$;

β = A vector of estimated coefficients for explanatory variables;

Note that the parallel regression assumption should be satisfied in proportional odds model, i.e., the coefficient will be the same for all values of j in the proportional odds model. It is necessary to test this parallel regression assumption as the coefficient may vary across different level of injury categories. If the proportional odds assumption is violated, then the partial proportional odds model is more appropriate. This model overcomes a key limitation of proportional odds model. For some variables, the coefficients will be the same for all categories of j while for other variables coefficients can differ across injury categories. The predicted probability can be formulated as follows (Long and Freese 2006, Williams 2006):

$$\Pr(Y = 1|X) = \frac{\exp(\alpha_1 - X\beta_1)}{1 + \exp(\alpha_1 - X\beta_1)} \quad (4)$$

$$\Pr(Y = j|X) = \frac{\exp(\alpha_j - X\beta_j)}{1 + \exp(\alpha_j - X\beta_j)} - \frac{\exp(\alpha_{j-1} - X\beta_{j-1})}{1 + \exp(\alpha_{j-1} - X\beta_{j-1})} \text{ for } j = 2 \text{ to } J - 1 \quad (5)$$

$$\Pr(Y = J|X) = 1 - \frac{\exp(\alpha_{J-1} - X\beta_{J-1})}{1 + \exp(\alpha_{J-1} - X\beta_{J-1})} \quad (6)$$

Where,

β_j = A vector of estimated coefficients that vary based on category level.

α = Cut points (thresholds) for injury severity level from $j = 1$ to $J - 1$.

SECTION 4: FINDINGS

4.1 Descriptive analysis

Table 1 shows the descriptive statistics of rail-pedestrian and bicyclist trespassing crashes occurring at highway-rail grade crossings and non-crossings during the study period. Most of these trespassers had direct collisions with trains (at grade crossings) or were struck by on-track equipment (at non-crossings). Total yearly crashes ranged from 902 to 1,117 with a small peak in 2006 and a slight increase from 2011. Annually, grade crossing crashes were more uniform (ranging from 159 to 238) than non-crossing crashes (ranging from 738 to 956). Non-crossing trespassing crashes are important because they account for nearly 81% of the total.

On the whole, over 51% (5,231) of the crashes resulted in a rail-pedestrian and bicyclist trespassing fatality. Thus, these crashes annually killed about 510 individuals. Most fatalities involved pedestrian and bicyclist trespassers who were running or walking, sitting or standing, lying, crossing and climbing at grade crossing or rail tracks. Individuals involved in rail-pedestrian and bicyclist trespassing crashes tended to be adults between 17 and 64 years old, as expected (87% overall).

Figure 2 presents the percentage of rail-pedestrian and bicyclist trespassing crashes (crossing vs. non-crossing) by pre-crash behaviors during the study period. Riding or operation of equipment accounts for about more than 51% of annual trespasser crashes at grade crossings versus 5.8% at non-crossings. Trespassers who were running/walking accounted for 37.7% of non-crossing crashes and 29.5% of grade crossing crashes. It is also notable that about 24% of non-crossing crashes involved lying/sleeping, whereas these showed a much lower frequency at 1.6% at crossings. Generally, rail-pedestrian and bicyclist trespassers involved in crashes exhibit substantially different pre-crash behaviors at crossings versus non-crossings.

Age shows substantial variations between grade crossings and non-crossings. Senior trespassers (aged 65 or greater) are involved in 10.7% of grade crossing crashes, but only 3.1% of non-crossing crashes. About 48.7% of the trespasser crashes occurred at night. Of crossing crashes, 42% were nighttime compared with 50% of non-crossing crashes.

Rail yards are associated with 3.6% of non-crossing trespasser crashes and 0.6% of crossing crashes. The number of tracks and operational needs of the railroads tend to minimize the placement of at-grade crossings at yard locations. The average population density is 1214 per square mile for non-crossing crashes, but 997 per square mile for grade crossing crashes. There is no significant difference between non-crossings and grade crossings regarding to education and income level.

Correlations between variables were estimated and analyzed, especially focussing on correlations among potential explanatory variables. The estimated matrix indicated that almost all correlations among potential explanatory variables were lower than 0.3. A few variables, such as youth and adult categories in the age variable, show higher correlations due to the structure of the data. That is, they are indicator variables representing high-frequency age groups and as such are correlated. Overall, the correlation matrix implies that multicollinearity problems may be minimal.

Figure 3 presents the kernel density of rail crossing and non-crossing trespassing crashes across the country. Blue and red colors indicate low density area and high density areas, respectively. The map is provided for visualization and verification/error-checking purposes only. This paper does not analyze the spatial locations of crashes. The distribution of non-crossing crashes are more dispersed (at 5% level). A large proportion of crashes occurred in the northeastern U.S. This region has a large population and high train traffic volumes.

4.2 Model results

The outputs of pooled and separate models quantifying the correlates of rail-pedestrian and bicyclist trespassing injury severity are shown in Table 2 and Table 3. The first model shown in Table 2 is a simple ordered logit model (proportional odds model) with the restrictive parallel regression assumption while the model shown in Table 3 is a partial proportional odds model which relaxes the restriction of parallel regression assumption. The backward stepwise selection procedure autofit was used, showing that the parallel regression assumption is violated (at 5% level). All models shown in Table 2 and Table 3 have goodness-of-fit that is on the low side. However, most variables show statistically significant correlations with the response variable (at 5% level) and the model is statistically significant overall.

The use of separate models is justified by the Likelihood Ratio test, at 5% level (Neyman and Pearson 1992), which indicates that separate models are preferred. Based on the results, the correlates of rail-pedestrian and bicyclist trespassing fatality differ between grade crossings and

non-crossings. Also, note that in Table 3, most correlates vary across the ordinal categories at non-crossings while they are consistent at grade crossings. The results highlight the fact that pooled model can hide some of the variations in both the signs and magnitudes of coefficients between grade crossings and non-crossings. Marginal effects are provided to give a cleaner sense of correlates. Marginal effects in ordered regression model are the changes of probability from a lower injury category to a higher injury category associated with a unit change in the explanatory variable. Variables with a positive sign indicate an increasing probability of a non-motorized trespasser being severely injured or killed with an increase in the explanatory variable.

4.2.1 Discussion of Key Factors

4.2.1.1 Pre-crash behaviors

Key variables of interest in this study are pre-crash behaviors. The pooled model shows that the chances of rail-pedestrian and bicyclist trespassing crashes fatality are statistically significantly higher when the person is lying/sleeping, running/walking, sitting/standing/bending/stooping. Lying has been suspected to be suicides or intoxication, indicating that these individuals potentially expose themselves to trains and would not escape from the train (Savage 2007, Silla and Luoma 2012). Riding/operation and climbing/jumping are associated with lower injury severity compared with the base of “other actions” which include exercising, for example.

The separate models show interesting results. Notable are the differences in magnitudes and signs of pre-crash behaviors. Lying/sleeping has the largest association with fatalities both at grade crossings and non-crossings, but more injurious at grade crossings. Additionally, running/walking are more injurious at grade crossings, than at non-crossings. Climbing/jumping and riding/operation are statistically significantly associated with lower probability of fatality at non-crossings while they do not associate significantly with injury severity at crossings. Crossing/crawling is associated with more injurious at grade crossings, but not significant at non-crossings. Sitting/standing/bending/stooping is more injurious at non-crossings. Generally, pre-crash behaviors are more statistically significantly associated with rail-pedestrian and bicyclist trespassing crashes severity at crossings.

Notably, the difference between proportional odds model and partial proportional odds model is that partial proportional odds model quantifies the potentially different contributions of explanatory variables across injury categories. In this case, pre-crash behaviors have shown different signs and magnitudes of coefficients across injury categories at non-crossings, while the signs and magnitudes are more consistent across injury categories at grade crossings. For example, sitting/standing/bending/stooping only makes contribution to severe injury at grade crossings while the coefficients for sitting/standing/bending/stooping are consistently positive but increase from minor injury category across severe injury category. This means that people with such pre-crash behaviors are associated with higher change of fatal injury but they are less likely to put themselves in severe injury at non-crossings, compared with base level (other behaviors). Conversely, trespassers who are climbing/jumping are more likely to be severely injured at non-crossings while they do not associate significantly with minor injury.

4.2.1.2 Other trespassing-related factors

The pooled model shows that the chances of rail-pedestrian and bicyclist trespassing crashes fatality are statistically significantly higher when the person involved is a senior, and the crash occurs in a non-railyard location. Results from the separate model show that trespassers less than 65 years old have lower probability of fatal crash involvement (given a crash) at grade crossings, than at non-crossings overall. Crashes occurring in summer or darkness shows statistically significantly higher probability of fatality at crossings, while less injurious at non-crossings. The county median household income contributes higher chance of fatal injury. The more persons under poverty level in a county, the higher probability of fatal crash involvement at grade crossings. No statistical evidence was found relating injury severity with location (urban vs. rural) or population density.

Table 3 also shows the chance of rail-pedestrian and bicyclist trespassing fatality across highway-rail grade crossings and non-crossings. This provides information about how different attributes are related to rail-pedestrian and bicyclist trespassing fatal crashes. Take lying/sleeping for example (in partial proportional odds model), from the pooled model, the chance for lying/sleeping trespassers being killed is 23.9%, compared with trespassers who are choosing to engage in other behaviors (i.e., the base level of pre-crash actions in the model). For the non-crossing model, the chance of fatality is a little smaller which is 23.2%, while from the crossing model, the percentage is substantially larger, increasing by 35.2%.

To understand trends, Figure 4 shows the chance of fatality in rail-pedestrian and bicyclist trespassing crashes over time (relative to the base year of 2005) based on the results from partial proportional odds model. The trends of chance of fatality for rail crossing and non-crossing fatality vary according to year. During the years analyzed, the chances of fatality have remained stable in terms of non-crossing crashes (showing slight reductions since 2009 but slight increase after 2013), while the chance of crossing fatalities have greater fluctuations (small peaking in 2009 and showing a sharp reduction since 2009, but sharp increase after 2013), compared with the base level (of 2005). Clearly, the study has shown substantially different correlations in terms of time trends and various factors with rail-pedestrian and bicyclist trespassing fatality between highway-rail grade crossings and non-crossings.

SECTION 5: LIMITATIONS

While this study selected key variables for analysis, other factors (e.g., alcohol involvement) may contribute to pedestrian and bicyclist trespasser injury severity, but lack of data prevented their investigation. Available variables constrain the analysis, and consequently the estimated models have limited explanatory power. The data availability problem is more severe for non-crossings, since such data are difficult to obtain.

Another limitation of this study is the under-reporting issue. According to the FRA guide for preparing accident or incident reports, the reportable railroad accidents/incidents should meet the reporting criteria (Safety 2011). One criteria is that the reported railroad accidents/incidents should satisfy the monetary reporting threshold, such as \$7,700 for calendar year 2006. Besides that, accidents/incidents reported in 6180.55a should also meet with the general reporting criteria, like injury to any person that result in medical treatment. Thus, the database would not contain cases that do not meet the criteria. Then these under-reported or less severe injury cases would make the predicted probability of fatality in this study is higher than reality.

Variations in the magnitudes of dependent and independent variables may affect study findings. Using parametric statistical methods to analyze crash data from FRA, this study did not control variations in variables (as is the case with experimental studies where researchers have greater control). Further, ten years of data analyzed in the study could be affected by variation in the reporting methods or procedures (although the variables collected are consistent over these years). While the researchers checked the data using descriptive statistics, coding errors or inaccurate records may still exist.

SECTION 6: CONCLUSIONS

This study addresses a key question: how pre-crash behaviors and other correlates of rail-pedestrian and bicyclist trespassing injury severity differ between highway-rail grade crossings and non-crossings? To answer the question, ten years (2005-2014) of FRA maintained rail pedestrian and bicyclist trespassing crash records were analyzed by estimating rigorous models, i.e., partial proportional odds and proportional odds models. This comprehensive database contains crashes at both highway-rail grade crossings and non-crossings. Results from separate and pooled models revealed substantial differences in correlates of rail-pedestrian and bicyclist trespassing crash injury severity, especially in terms of pre-crash behaviors. Key findings include:

- Higher probability of rail-pedestrian and bicyclist trespassing fatality is associated with lying/sleeping; lying/sleeping and running/walking are injurious at both highway-rail grade crossings and non-crossings, but more so at highway-rail grade crossings.
- Sitting/standing/bending/stooping are more injurious at non-crossings, while crossing/crawling are more injurious at highway-rail grade crossings.
- Crashes occurring during darkness and summer are more injurious at grade crossings.
- Seniors (65 years or older) are more likely to be involved in fatal crashes than other age groups (less than 65 years old).
- The trends show more fluctuations for crossing crashes, but relatively more stability for non-crossings over a 10-year period.

Generally, rail-motorist collisions were used to evaluate railroad safety. However, such measurement is not appropriate without considering rail-pedestrian and bicyclists crashes. This study indicated that pedestrian and bicyclists trespassing behaviors are significant to railroad safety. Besides, rail-pedestrian and bicyclists trespassing crashes at non-crossings have received less attention in the literature, compared with crashes at grade crossings. For example, the Rail Safety Improvement Act has an entire chapter devoted to crashes at grade crossings (US Congress 2008), but non-crossing crashes are not covered. Furthermore, almost all available rail-trespasser related studies were dealing with crash frequency, but not crash injury severity. This study is timely because it highlights different risk factors associated with injury severity in rail-pedestrian and bicyclists trespassing crashes comparing grade crossings and non-crossings. The risk factors identified in the study can point transportation researchers, practitioners and policy makers toward safety interventions that can reduce risks, lowering injury severity in crashes.

This study reveals lying/sleeping is a critical pre-crash behavior associated with high probabilities of rail-pedestrian and bicyclists trespassing fatality for both highway-rail grade crossings and non-crossings. The behavior indicates suicidal intentions, involvement of alcohol/drugs, general negligence, or the person being transient. Studies on relevant

countermeasures are needed, e.g., posting information about suicide prevention hotlines. It is possible that countermeasures will have different outcomes at grade crossings than non-crossings. Past studies have pointed out that the effectiveness of some countermeasures, which are typically highly dependent on trespasser characteristics and location of crashes (Silla and Luoma 2011, Burkhardt, Radbo et al. 2014). Overall, to promote railroad safety at the national level, studies of countermeasures that reduce crashes and injuries at both crossings and non-crossings are needed.

SECTION 7: REFERENCES

- Administration, F. R. (2015). Casualty Summary Tables. *Federal Railroad Administration Office of Safety Analysis*. Available from: <http://safetydata.fra.dot.gov/OfficeofSafety/default.aspx> [Accessed 12-10 2015].
- Burkhardt, J.-M., Radbo, H., Silla, A., & Paran, F. (2014). *A Model of Suicide and Trespassing Processes to Support the Analysis and Decision Related to Preventing Railway Suicides and Trespassing Accidents at Railways*. Paper presented at the Transport Research Arena (TRA) 5th Conference: Transport Solutions from Research to Deployment.
- Cina, S. J., Koelpin, J., Nichols, C., & Conradi, S. (1994). A decade of train-pedestrian fatalities: the Charleston experience. *Journal of forensic sciences*, 39(3), 668-673.
- Eluru, N., Bhat, C. R., & Hensher, D. A. (2008). A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis & Prevention*, 40(3), 1033-1054.
- Fan, W., & Haile, E. W. (2014). *Analysis of Severity of Vehicle Crashes at Highway-Rail Grade Crossings: Multinomial Logit Modeling*. Paper presented at the Transportation Research Board 93rd Annual Meeting.
- Gabree, S. H., Chase, S., Doucette, A., & Martino, M. (2014). *Countermeasures to Mitigate Intentional Deaths on Railroad Rights-of-Way: Lessons Learned and Next Steps*. Retrieved from
- George, B. F. (2008). Rail Trespasser Fatalities Demographic and Behavioral Profiles [Online]. *Cadle Creek Consulting*. Available from: <https://www.fra.dot.gov/eLib/Details/L02669> [Accessed 07-12 2015].
- Horton, S., Carroll, A., Chaudhary, M., Ngamdung, T., Mozenter, J., & Skinner, D. (2009). *Success Factors in the Reduction of Highway-Rail Grade Crossing Incidents from 1994 to 2003*. Retrieved from
- Hu, S.-R., Li, C.-S., & Lee, C.-K. (2010). Investigation of key factors for accident severity at railroad grade crossings by using a logit model. *Safety science*, 48(2), 186-194.
- Khattak, A. (2013). *Severity of Pedestrian Crashes at Highway-Rail Grade Crossings*. Paper presented at the Transportation Research Board 92nd Annual Meeting.
- Khattak, A., & Luo, Z. (2011). Pedestrian and bicyclist violations at highway-rail grade crossings. *Transportation Research Record: Journal of the Transportation Research Board*, 2250(1), 76-82.
- Lerer, L. B., & Matzopoulos, R. (1996). Meeting the challenge of railway injury in a South African city. *The Lancet*, 348(9028), 664-666.

- Liu, J., Bartnik, B., Richards, S. H., & Khattak, A. J. (2015). Driver Behavior at Railway-Highway Grade Crossings with Passive Traffic Controls: A Driving Simulator Study. *Journal of Transportation Safety and Security*. doi:<http://dx.doi.org/10.1080/19439962.2015.1043478>
- Liu, J., Khattak, A. J., Richards, S. H., & Nambisan, S. (2014). What are the Differences in Driver Injury Outcomes at Highway-Rail Grade Crossings? Untangling the Role of Pre-Crash Behaviors. in *93rd Annual Meeting of the Transportation Research Board*. 2014. Washington D.C.
- Lobb, B. (2006). Trespassing on the tracks: a review of railway pedestrian safety research. *Journal of safety research*, 37(4), 359-365.
- Lobb, B., Harre, N., & Suddendorf, T. (2001). An evaluation of a suburban railway pedestrian crossing safety programme. *Accident Analysis & Prevention*, 33(2), 157-165.
- Lobb, B., Harre, N., & Terry, N. (2003). An evaluation of four types of railway pedestrian crossing safety intervention. *Accident Analysis & Prevention*, 35(4), 487-494.
- Long, J. S., & Freese, J. (2006). *Regression models for categorical dependent variables using Stata*: Stata press.
- Matzopoulos, R., & Lerer, L. B. (1998). Hours to hell and back: A social epidemiology of railway injury in a South African city, 1890–1995. *Social Science & Medicine*, 47(1), 75-83.
- Metaxatos, P., & Sriraj, P. (2013). Pedestrian/Bicyclist Warning Devices and Signs at Highway-Rail and Pathway-Rail Grade Crossings.
- Mishara, B. L. (2007). Railway and metro suicides: Understanding the problem and prevention potential. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, 28(S1), 36.
- Neyman, J., & Pearson, E. S. (1992). *On the problem of the most efficient tests of statistical hypotheses*: Springer.
- Nixon, J., Corcoran, A., Fielding, L., & Eastgate, J. (1985). Fatal and nonfatal accidents on the railways—a study of injuries to individuals. with particular reference to children and to nonfatal trauma. *Accident Analysis & Prevention*, 17(3), 217-222.
- Oh, J., Washington, S. P., & Nam, D. (2006). Accident prediction model for railway-highway interfaces. *Accident Analysis & Prevention*, 38(2), 346-356.
- Patterson, T., & Authority, L. T. S. (2004). Analysis of trespasser accidents. *Land Transport Safety Authority, New Zealand*.
- Pelletier, A. (1997). Deaths among railroad trespassers: the role of alcohol in fatal injuries. *JAMA*, 277(13), 1064-1066.
- Safety, O. o. R. (2011). FRA Guide for Preparing Accident/Incident Reports *Federal Railroad Administration*. Available from: <http://safetydata.fra.dot.gov/OfficeofSafety/default.aspx> [Accessed 07-12 2015].
- Safety, R. (2007). Standards Board (2007) Trespass and access via the platform end. Final Report. Halcrow Group Limited in partnership with Human Engineering.
- Savage, I. (2007). Trespassing on the Railroad. *Research in Transportation Economics*, 20, 199-224.
- Silla, A., & Luoma, J. (2009). Trespassing on Finnish railways: identification of problem sites and characteristics of trespassing behaviour. *European transport research review*, 1(1), 47-53.
- Silla, A., & Luoma, J. (2011). Effect of three countermeasures against the illegal crossing of railway tracks. *Accident Analysis & Prevention*, 43(3), 1089-1094.

- Silla, A., & Luoma, J. (2012). Main characteristics of train–pedestrian fatalities on Finnish railroads. *Accident Analysis & Prevention*, 45, 61-66.
- Trottenberg, P., & Rivkin, R. S. (2013). Guidance on treatment of the economic value of a statistical life in US Department of Transportation analyses. *Revised departmental guidance, US Department of Transportation*.
- US Congress. (2008). Public Law 110–432 - Rail Safety Improvement Act of 2008. Retrieved from <https://www.fra.dot.gov/Page/P0395>
- Wang, X., Khattak, A. J., Liu, J., & Clarke, D. (2014). Non-crossing Rail-Trespassing Crashes in the Past Decade: A Spatial Approach to Analyzing Injury Severity. *in 93rd Annual Meeting of the Transportation Research Board. 2014. Washington D.C.*
- Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata Journal*, 6(1), 58-82.
- Zhao, S., & Khattak, A. (2015). Motor vehicle drivers' injuries in train–motor vehicle crashes. *Accident Analysis & Prevention*, 74, 162-168.

LIST OF TABLES AND FIGURES

Table 1. Descriptive Statistics for Rail-Pedestrian/Bicyclist Trespassing Events Using FRA Data	44
Table 2. Ordered Logit Model for Injury Severity in Rail-Pedestrian/Bicyclist Trespassing Crashes at Crossings and Non-Crossings	45
Table 3. Partial Proportional Odds Models for Injury Severity in Rail-Pedestrian and bicyclist Crashes at Crossings and Non-Crossings	46
Figure 1. Data Structure of Railway Safety Dataset in FRA	47
Figure 2. Rail-Pedestrian and bicyclist Trespassing Crashes by Pre-Crash Behavior Across Highway-Rail Grade Crossings and Non-Crossings	48
Figure 3. Kernel Density Distribution of Rail-Pedestrian and bicyclist Trespassing Crashes.....	49
Figure 4. Chance of Fatality in Rail-Pedestrian and bicyclist Trespassing Crashes for 2005-2014	50

Table 1. Descriptive Statistics for Rail-Pedestrian and bicyclist Trespassing Events Using FRA Data

Variables		Total (N=10146)				Non-crossing (N=8258)		Highway-rail grade crossing (N=1888)		%Diff of Mean
		Mean	Std. Dev	Min	Max	Mean	Std. Dev	Mean	Std. Dev	
Injury	Minor injury (level 1)	0.148	0.355	0	1	0.133	0.340	0.212	0.409	58.76%
	Severe injury (level 2)	0.336	0.472	0	1	0.332	0.471	0.354	0.478	6.44%
	Killed (level 3)	0.516	0.500	0	1	0.534	0.499	0.434	0.496	-18.69%
Personal attributes	Youths (<=16 years old)	0.078	0.268	0	1	0.068	0.251	0.123	0.329	81.99%
	Middle (17-35 years old)	0.378	0.485	0	1	0.391	0.488	0.322	0.467	-17.83%
	Adult (36-64 years old)	0.498	0.500	0	1	0.509	0.500	0.448	0.497	-12.04%
	Seniors (>=65 years old)	0.046	0.208	0	1	0.031	0.175	0.107	0.309	239.82%
Temporal attributes	Darkness (0-no, 1-yes)	0.487	0.500	0	1	0.503	0.500	0.420	0.494	-16.44%
Seasonal attributes	Summer (0-no, 1-yes)	0.288	0.453	0	1	0.298	0.457	0.242	0.428	-18.78%
	Winter (0-no, 1-yes)	0.200	0.400	0	1	0.186	0.390	0.257	0.437	38.03%
	Spring and autumn (0-no, 1-yes)	0.513	0.500	0	1	0.516	0.500	0.501	0.500	-2.90%
Location	Yard (0-no, 1-yes)	0.030	0.171	0	1	0.036	0.186	0.006	0.076	-83.75%
Trespassing pre-crash actions	Climbing, jumping	0.055	0.229	0	1	0.067	0.250	0.005	0.073	-92.06%
	Riding, operation	0.143	0.350	0	1	0.058	0.234	0.515	0.500	789.43%
	Lying, sleeping	0.199	0.400	0	1	0.241	0.428	0.016	0.125	-93.42%
	Running, walking	0.362	0.480	0	1	0.377	0.485	0.295	0.456	-21.71%
	Crossing, crawling	0.024	0.153	0	1	0.025	0.157	0.019	0.135	-26.75%
	Sitting, standing, bending, stooping	0.155	0.362	0	1	0.169	0.374	0.096	0.295	-42.85%
	Other actions	0.062	0.240	0	1	0.063	0.243	0.054	0.226	-14.53%
Crash year (2005-2014)		Not reported				Not reported		Not reported		
	% of bachelor's degree or higher (25+)	26.633	10.141	5.1	71.7	26.893	10.274	25.498	9.457	-64.83%
Census attributes	Median household income (\$/1000)	52.707	13.682	20.972	110.292	52.991	13.931	51.465	12.465	-76.48%
	% of persons below poverty level	16.367	5.520	3.9	44.1	16.309	5.622	16.620	5.042	-69.08%
	Population per square mile	1173.647	3023.090	0.3	69467.5	1214.113	3275.51	996.653	1465.72	20.72%

Note: "Other actions" refers to pre-crash actions that cannot be coded into one of the categories shown in the Table, e. g., other actions include a person exercising on or near tracks;

"% Difference of mean" refers to (Crossing mean – Non-crossing mean) /Non-crossing mean.

"Not reported" statistical descriptive of crash years are shown in Figure 2.

Table 2. Ordered Logit Model for Injury Severity in Rail-Pedestrian and bicyclist Trespassing Crashes at Crossings and Non-Crossings

Variables		Pooled model				Separate model							
						Non-crossing mode				Highway-rail grade crossing model			
		β	Minor	Marg. Eff. Severe	Fatal	β	Minor	Marg. Eff. Severe	Fatal	β	Minor	Marg. Eff. Severe	Fatal
Personal attributes (base: senior)	Youths (<=16 years old)	-0.884**	13.1%	6.9%	-20.0%	-0.740**	10.0%	7.2%	-17.2%	-1.065**	19.4%	1.9%	-21.3%
	Middle (17-35 years old)	-0.635**	7.9%	6.6%	-14.5%	-0.516**	5.9%	6.1%	-12.0%	-0.837**	13.7%	3.8%	-17.5%
	Adult (36-64 years old)	-0.496**	5.9%	5.2%	-11.1%	-0.386**	4.3%	4.6%	-8.8%	-0.669**	10.4%	3.5%	-13.9%
Temporal attributes	Darkness (0-no, 1-yes)	-0.053	0.6%	0.6%	-1.2%	-0.123**	1.4%	1.5%	-2.9%	0.254**	-3.9%	-1.6%	5.5%
Season (base: spring and autumn)	Summer (0-no, 1-yes)	-0.066	0.8%	0.7%	-1.5%	-0.128**	1.4%	1.6%	-3.0%	0.208*	-3.1%	-1.4%	4.5%
	Winter (0-no, 1-yes)	-0.036	0.4%	0.4%	-0.8%	-0.012	0.1%	0.1%	-0.3%	-0.094	1.5%	0.6%	-2.0%
Location	Yard (0-no, 1-yes)	-0.795**	9.4%	8.9%	-18.4%	-0.822**	9.0%	10.2%	-19.2%	-0.602	9.3%	3.7%	-13.0%
Trespassing pre-crash actions (base: other action)	Climbing, jumping	-0.268**	3.4%	2.8%	-6.2%	-0.317**	3.8%	3.7%	-7.5%	-0.404	6.9%	1.6%	-8.5%
	Riding, operation	-0.383**	4.9%	4.1%	-9.0%	-0.344**	4.2%	3.9%	-8.1%	-0.138	2.1%	0.9%	-3.0%
	Lying, sleeping	1.052**	-10.1%	-13.3%	23.5%	1.025**	-9.4%	-13.6%	23.0%	1.616**	-15.9%	-17.5%	33.4%
	Running, walking	0.643**	-7.3%	-7.4%	14.7%	0.552**	-5.9%	-6.7%	12.6%	1.155**	-15.4%	-11.3%	26.8%
	Crossing, crawling	0.107	-1.2%	-1.2%	2.5%	0.008	-0.1%	-0.1%	0.2%	0.770**	-9.7%	-7.1%	16.9%
	Sitting, standing, bending, stooping	0.662**	-6.8%	-8.1%	14.9%	0.693**	-6.5%	-9.0%	15.5%	0.368	-5.3%	-2.8%	8.0%
Crash year (base: 2005 year)		Not reported				Not reported				Not reported			
Census attributes	% of bachelor's degree or higher (25+)	-0.009**	0.1%	0.1%	-0.2%	-0.010**	0.1%	0.1%	-0.2%	-0.001	0.0%	0.0%	0.0%
	Median household income (\$/1000)	0.017**	-0.2%	-0.2%	0.4%	0.017**	-0.2%	-0.2%	0.4%	0.019**	-0.3%	-0.1%	0.4%
	% of persons below poverty level	0.010	-0.1%	-0.1%	0.2%	0.006	-0.1%	-0.1%	0.1%	0.025*	-0.4%	-0.2%	0.5%
	Population per square mile	0.000	0.0%	0.0%	0.0%	0.000	0.0%	0.0%	0.0%	0.000	0.0%	0.0%	0.0%
Constant	cut 1	-1.163				-1.373				-0.024			
	cut 2	0.652				0.471				1.730			
Sample size		10146				8258				1888			
Pseudo-R2		0.0457				0.0394				0.0662			
Log Likelihood at β		-9593.8587				-7693.298				-1866.2424			
Log Likelihood at 0		-10053.083				-8008.8904				-1998.6202			
Prob>ChiSq		<.0001*				<.0001*				<.0001*			
Likelihood ratio test						68.64, <.0001*							

Notes: STATA software (ologit program) was used. "Pseudo-R²" refers to $1 - (\text{Log Likelihood at } \beta / \text{Log Likelihood at } 0)$;

Marginal effects refer to the changes in the response variable with a unit change in the explanatory variable.

"Other actions" is the base for pre-crash behaviors; it refers to pre-crash actions that cannot be coded into one of the categories shown in the Table, e.g., other actions include a person exercising on or near tracks;

Crash years have not shown statistical significant correlation at non-crossings in ordered logit model.

*** means statistical significant associations were found (at 5% level); ** means statistical significant associations were found (at 10% level).

Table 3. Partial Proportional Odds Models for Injury Severity in Rail-Pedestrian and bicyclist Crashes at Crossings and Non-Crossings

Variables		Pooled model					Separate models									
		Coefficient		Marginal Effect			Non-crossing model					Highway-rail grade crossing model				
		β_1	β_2	Minor	Severe	Fatal	β_1	β_2	Minor	Severe	Fatal	β_1	β_2	Minor	Severe	Fatal
Personal attributes (base: senior)	Youths (<=16 years old)	-0.878**	-0.878**	10.6%	9.5%	-20.1%	-0.744**	-0.744**	8.3%	8.8%	-17.1%	-1.052**	-1.052**	16.2%	6.4%	-22.7%
	Young(17-35 years old)	-0.512**	-0.671**	6.2%	9.2%	-15.3%	-0.415**	-0.545**	4.6%	7.9%	-12.5%	-0.838**	-0.838**	12.9%	5.1%	-18.1%
	Adult(36-64 years old)	-0.488**	-0.488**	5.9%	5.3%	-11.2%	-0.381**	-0.381**	4.3%	4.5%	-8.8%	-0.659**	-0.659**	10.2%	4.0%	-14.2%
Temporal attributes	Darkness (0-no, 1-yes)	-0.055	-0.055	0.7%	0.6%	-1.3%	-0.125**	-0.125**	1.4%	1.5%	-2.9%	0.422**	0.167	-6.5%	2.9%	3.6%
Season (base: spring and autumn)	Summer (0-no, 1-yes)	-0.067	-0.067	0.8%	0.7%	-1.5%	-0.128**	-0.128**	1.4%	1.5%	-2.9%	0.209*	0.209*	-3.2%	-1.3%	4.5%
	Winter (0-no, 1-yes)	-0.035	-0.035	0.4%	0.4%	-0.8%	-0.012	-0.012	0.1%	0.1%	-0.3%	-0.089	-0.089	1.4%	0.5%	-1.9%
Location	Yard (0-no, 1-yes)	-0.646**	-1.059**	7.8%	16.4%	-24.2%	-0.697**	-1.047**	7.8%	16.3%	-24.1%	-0.623	-0.623	9.6%	3.8%	-13.4%
Trespassing pre-crash actions (base: other action)	Climbing, jumping	0.170	-0.618**	-2.0%	16.2%	-14.1%	0.127	-0.669**	-1.4%	16.8%	-15.4%	-0.455	-0.455	7.0%	2.8%	-9.8%
	Riding, operation	-0.373**	-0.373**	4.5%	4.0%	-8.5%	-0.335**	-0.335**	3.7%	4.0%	-7.7%	-0.153	-0.153	2.4%	0.9%	-3.3%
	Lying, sleeping	1.048**	1.048**	-12.6%	-11.3%	23.9%	1.011**	1.011**	-11.3%	-12.0%	23.3%	1.635**	1.635**	-25.2%	-10.0%	35.2%
	Running, walking	0.639**	0.639**	-7.7%	-6.9%	14.6%	0.540**	0.540**	-6.0%	-6.4%	12.4%	1.164**	1.164**	-18.0%	-7.1%	25.1%
	Crossing, crawling	0.535**	-0.081	-6.4%	8.3%	-1.8%	0.426**	-0.201	-4.7%	9.4%	-4.6%	0.762**	0.762**	-11.8%	-4.7%	16.4%
	Sitting, standing, bending, stooping	0.417**	0.712**	-5.0%	-11.3%	16.3%	0.464**	0.725**	-5.2%	-11.5%	16.7%	0.026	0.527**	-0.4%	-11.0%	11.4%
Crash year (base: 2005 year)		Not reported					Not reported					Not reported				
	% of bachelor's degree or higher (25+)	-0.003	-0.011**	0.0%	0.2%	-0.3%	-0.005	-0.012**	0.1%	0.2%	-0.3%	-0.001	-0.001	0.0%	0.0%	0.0%
Census attributes	Median household income (\$/1000)	0.018**	0.018**	-0.2%	-0.2%	0.4%	0.017**	0.017**	-0.2%	-0.2%	0.4%	0.019**	0.019**	-0.3%	-0.1%	0.4%
	% of persons below poverty level	0.009	0.009	-0.1%	-0.1%	0.2%	0.006	0.006	-0.1%	-0.1%	0.1%	0.025*	0.025*	-0.4%	-0.2%	0.5%
	Population per square mile	0.000	0.000	0.0%	0.0%	0.0%	0.000	0.000	0.0%	0.0%	0.0%	0.000	0.000	0.0%	0.0%	0.0%
Constant	α_1	0.913**	-0.596**				1.065**	-0.388				-0.012	-1.720**			
Sample size (N)		10146					8794					1888				
Pseudo-R ²		0.0517					0.0459					0.0716				
Log Likelihood at β		-9532.8759					-7641.5635					-1855.5875				
Prob> χ^2		<.0001*					<.0001*					<.0001*				
Likelihood Ratio test		71.45, <0.0001*														

Notes: STATA software (gologit2 program) was used with autofit; “Pseudo-R2” refers to 1 – (Log Likelihood at β /Log Likelihood at 0);

Marginal effects refer to the changes in the response variable with a unit change in the explanatory variable.

“Other actions” is the base for pre-crash behaviors; it refers to pre-crash actions that cannot be coded into one of the categories shown in the Table, e.g., other actions include a person exercising on or near tracks; “Not reported” coefficients of crash years are shown in Figure 4;

“***” means statistical significant associations were found (at 5% level); “**” means statistical significant associations were found (at 10% level).

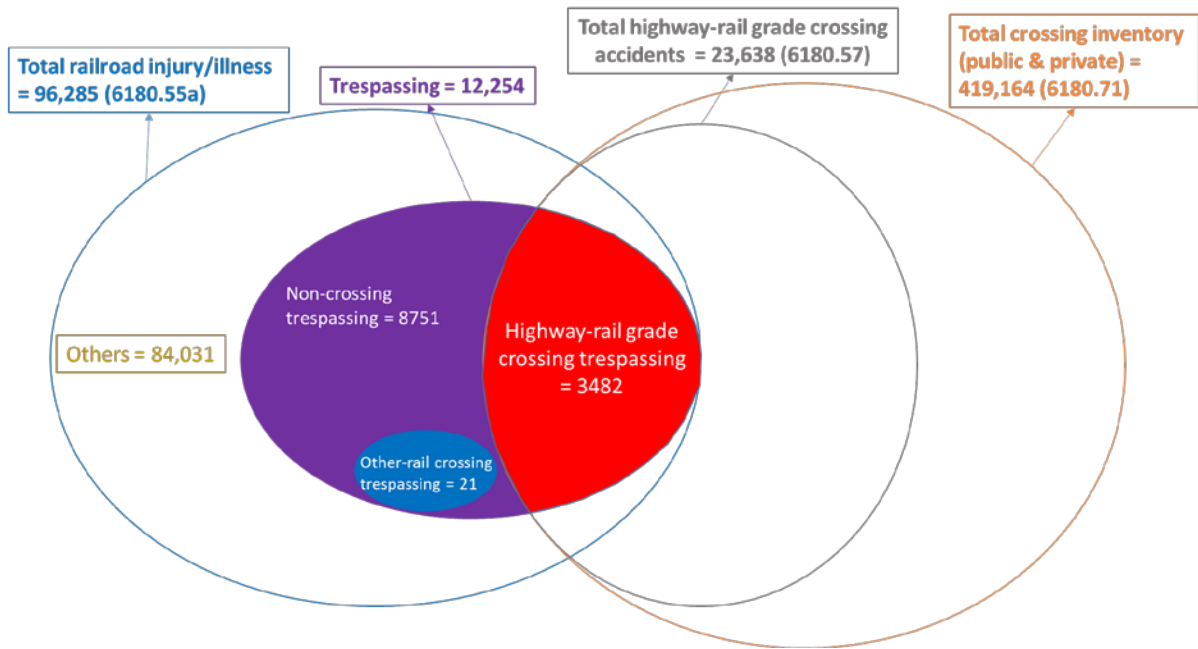
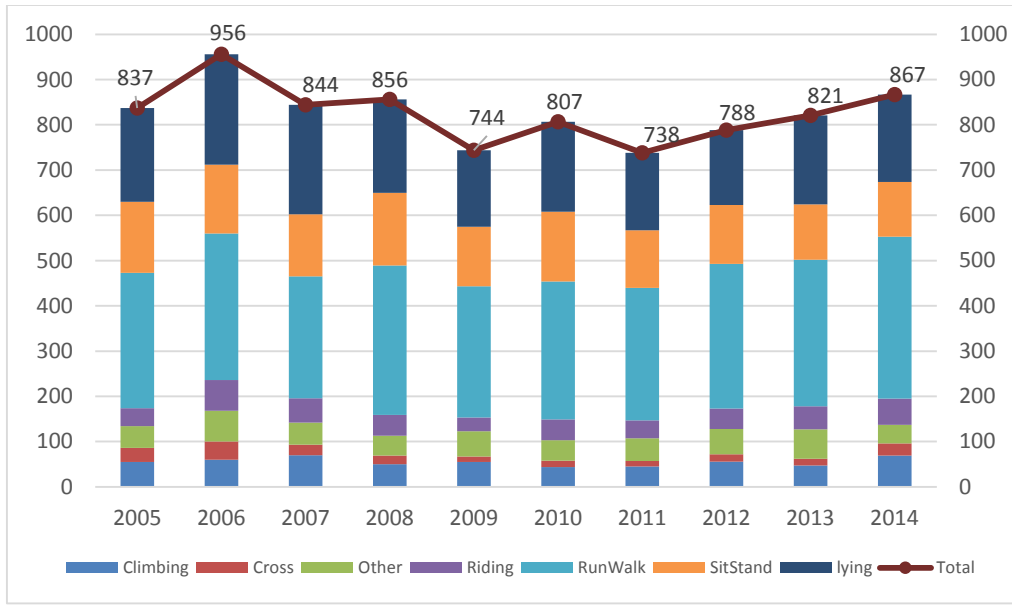
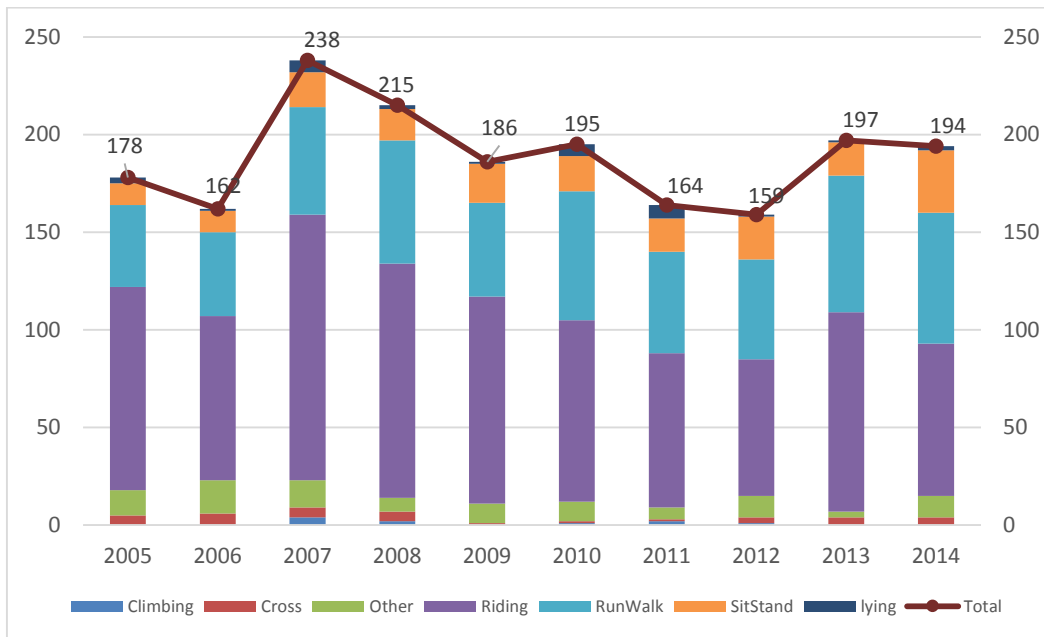


Figure 1. Data Structure of Railway Safety Dataset in FRA

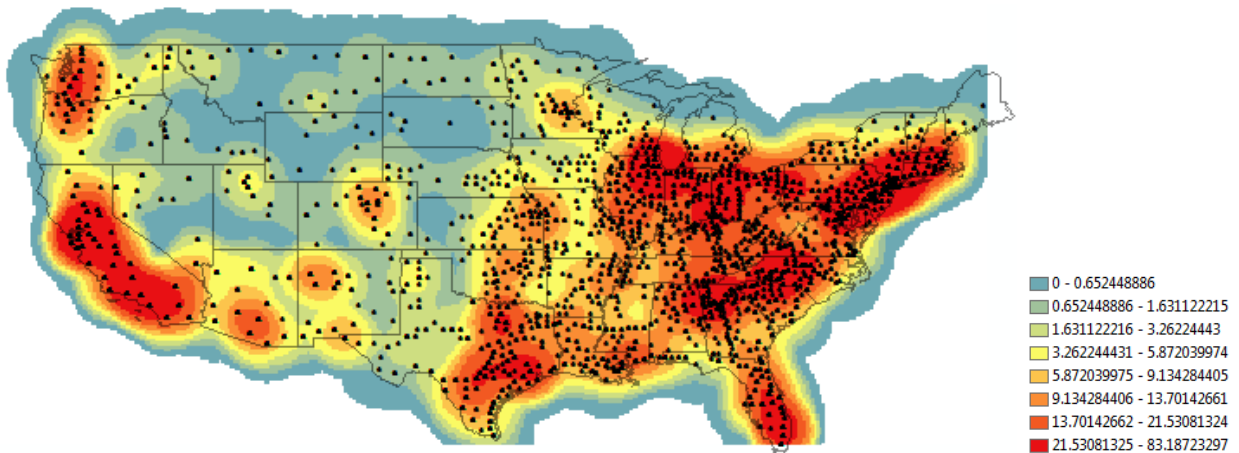


a) Crashes at Non-Crossings (N= 8,258)

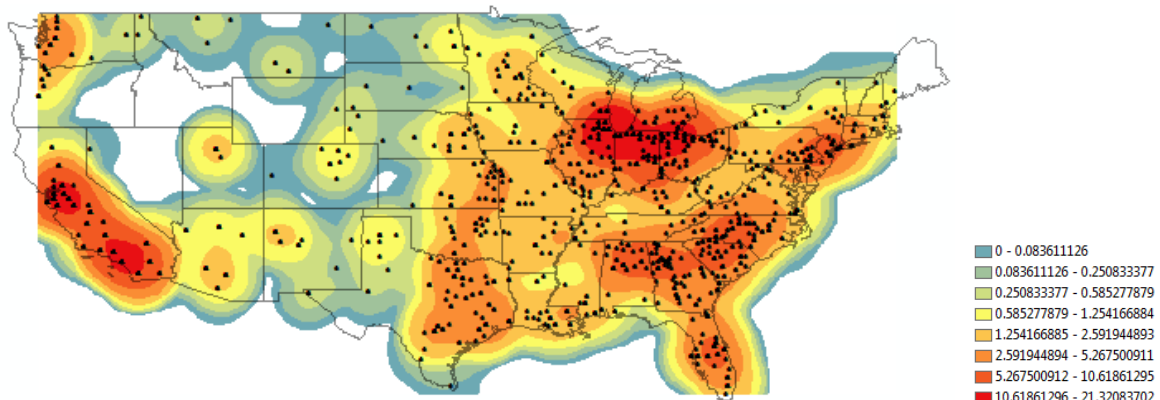


b) Crashes at Grade Crossings (N= 1,888)

Figure 2. Rail-Pedestrian and bicyclist Trespassing Crashes by Pre-Crash Behavior Across Highway-Rail Grade Crossings and Non-Crossings (2005-2014)



a) Total Rail-Pedestrian and bicyclist Trespassing Crashes at Non-Crossings ($N= 8,258$)



b) Total Rail-Pedestrian and bicyclist Trespassing Crashes at Grade Crossings ($N= 1,888$)

Figure 3. Kernel Density Distribution of Rail-Pedestrian and bicyclist Trespassing Crashes

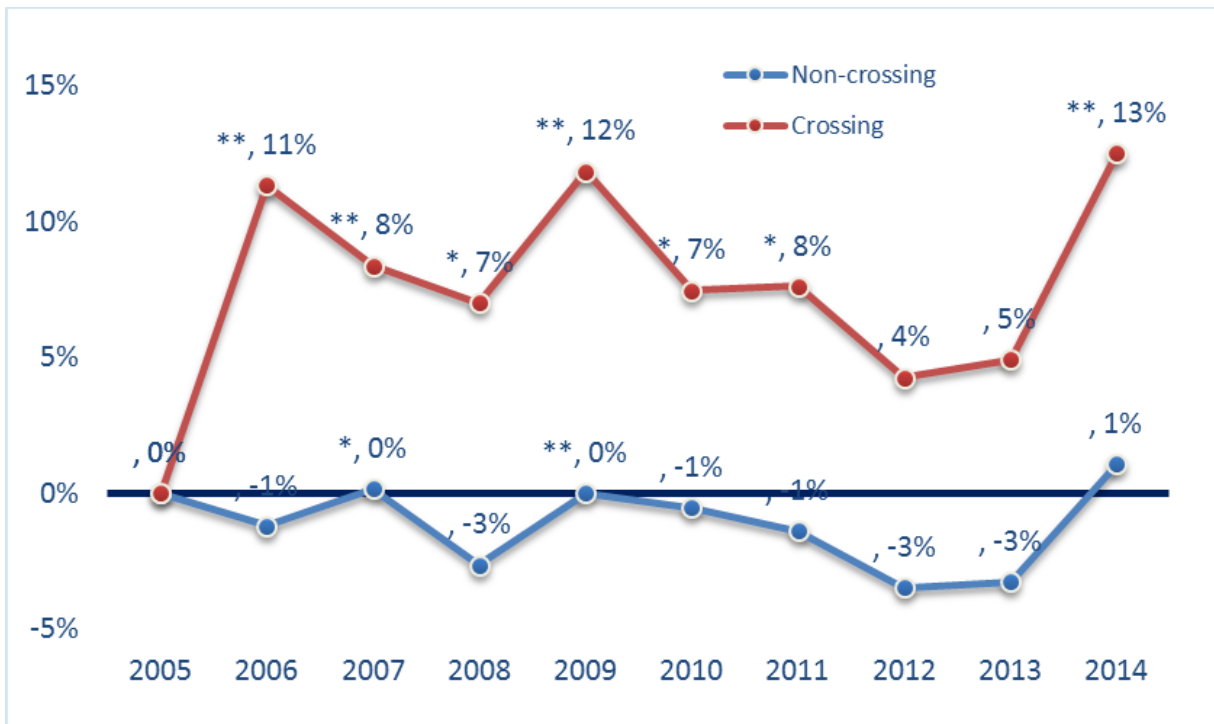


Figure 4. Chance of Fatality in Rail-Pedestrian and bicyclist Trespassing Crashes for 2005-2014

WHAT ARE THE DIFFERENCES IN DRIVER INJURY OUTCOMES AT HIGHWAY-RAIL GRADE CROSSINGS? UNTANGLING THE ROLE OF PRE-CRASH BEHAVIORS

Jun Liu, Asad Khattak, Stephen H. Richards and Shashi Nambisan
The University of Tennessee, Knoxville

Abstract – Crashes at highway-rail grade crossings can result in severe injuries and fatalities to vehicle occupants. Using a crash database from the Federal Railroad Administration (N=15,639 for 2004-2013), this study explores differences in safety outcomes from crashes between passive controls (Crossbucks and STOP signs) and active controls (flashing lights, gates, audible warnings and highway signals). To address missing data, an imputation model is developed, creating a complete dataset for estimation. Path analysis is used to quantify the direct and indirect associations of passive and active controls with pre-crash behaviors and crash outcomes in terms of injury severity. The framework untangles direct and indirect associations of controls by estimating two models, one for pre-crash driving behaviors (e.g., driving around active controls), and another model for injury severity. The results show that while the presence of gates is not directly associated with injury severity, the indirect effect through stopping behavior is statistically significant (95% confidence level) and substantial. Drivers are more likely to stop at gates that also have flashing lights and audible warnings, and stopping at gates is associated with lower injury severity. This indirect association lowers the chances of injury by 16 percent, compared with crashes at crossings without gates. Similar relationships between other controls and injury severity are explored. Generally, crashes occurring at active controls are less severe than crashes at passive controls. The results of study can be used to modify Crash Modification Factors (CMFs) to account for crash injury severity. The study contributes to enhancing the understanding of safety by incorporating pre-crash behaviors in a broader framework that quantifies correlates of crash injury severity at active and passive crossings.

Keywords: grade crossing, pre-crash behavior, injury severity, controls, path analysis, data imputation

SECTION 1: INTRODUCTION

Safety at highway-rail grade crossings remains an important societal concern in the United States, as well as other parts of the world. Crashes occurring at grade crossings can result in severe injuries and fatalities to vehicle occupants. Safety effectiveness of crossing controls is also important in tort liability that results from crashes. According to 2013 Federal Railroad Administration (FRA) crossing inventory database, the United States has 133,825 reported public crossings, as opposed to 82,921 crossings located on private property, highway-railroad (vehicle) grade crossings. Of these public vehicle grade crossings, 64,626 (48.3 percent) are passive crossings. Such crossings are those fitted with only passive warning devices (e.g., Crossbuck and STOP signs, pavement markings and advanced warning signs) that deliver static warnings, guidance, and, in some instances, mandatory action for the driver. The remaining

69,199 (51.7 percent) public grade crossings are active crossings, which are additionally fitted with active traffic control devices (e.g., gates, flashing lights and bells) that provide variable messages to motorists, indicating whether or not a train is approaching or occupying a crossing (Ogden 2007).

During 1981 to 2013, the number of crashes at highway-rail grade crossings has reduced by 77.8 percent (FRA 2013). The decreases are typically attributed to the upgrading from passive to active crossings and the improvements made on active grade crossings (Meeker, Fox et al. 1997, Millegan, Yan et al. 2009, Lenné, Rudin-Brown et al. 2011). Compared with passive controls (STOP signs and Crossbucks), active control devices (flashing lights and gates) have shown lower crash rates (Raub 2009). This is mainly due to the potential for active controls to gain additional driver attention and lead to greater compliance (Meeker, Fox et al. 1997). Although crash frequency has declined over the years, it is notable that fatality rates (per crash) at grade crossings have increased, from 7.7 percent in 1981 to 11 percent in 2013, as have injury rates (FRA 2013). While studies have pointed out clear relationships between crash frequencies and associated factors, and explored correlates of crash frequencies (Oh, Washington et al. 2006), it is still unclear what key factors, especially crossing controls, contribute to the severity of outcomes (e.g., lower or higher injuries) and the behavioral mechanisms that lead to injuries, given a crash.

The main objective of this study is to investigate relationships between safety outcomes and various crossing controls, and answer how differences in safety outcomes between active and passive traffic control at highway-rail grade crossings are associated with drivers' actions prior to the event of a crash, called pre-crash behaviors.

SECTION 2: LITERATURE REVIEW

Researchers aim to reveal what types of controls at grade crossings are effective in improving crossing safety, i.e., reducing crash frequencies/rates, or lowering crash injury severity. Table 1 and 2 summarize relevant studies that focused on the examination of crossing control effectiveness, in terms of crash frequencies/rates or crash injury severity.

2.1 Crash Rates

Many studies have compared crash rate at passive controls and active controls (Austin and Carson 2002, Elvik and Vaa 2004, Mok and Savage 2005, Park and Saccomanno 2005, Raub 2006, Saccomanno, Park et al. 2007, Elvik, Vaa et al. 2009, Millegan, Yan et al. 2009, Raub 2009, Yan, Han et al. 2010, Yan, Richards et al. 2010). They have generally found that crash rates are lower at STOP controlled intersections compared with Crossbuck signs. Furthermore, crossings fitted with active control devices including gates and flashing lights had lower crash rates than those with STOP signs.

The Highway Safety Manual provides Crash Modification Factors (CMFs) to indicate the percent of reduction/increase in crash rates after implementing certain types of traffic control devices at grade crossings (Gross, Persaud et al. 2010). For example, the CMF for upgrading previously passive crossings to active crossings with gates is 0.33, indicating the crash rates are expected to be reduced by 67 percent if the upgrading is made; and for installing gates at crossings that previously had flashing lights and sound signals is 0.55 (Elvik and Vaa 2004).

More CMFs generated by relevant studies are shown in Table 1, as well as the Clearinghouse website (<http://www.cmfclearinghouse.org/>). However, no information is available about the changes of the crash injury severity by implementing certain countermeasures at highway-rail grade crossings. This study explores the levels of the crash injury severity given a crash, focusing on the role of active vs. passive crossings.

2.2 Injury Severity

Comparatively fewer studies have examined crash injury severity (Raub 2006, Hu, Li et al. 2010, Eluru, Bagheri et al. 2012, Russo and Savolainen 2013, Hao and Daniel 2014, Zhao and Khattak 2015). Table 2 summarizes relevant studies on injury severity given a crash. Generally, crashes at crossings are severe, with many involving fatalities, e.g., 31.8 percent collisions at gate crossings were with at least one fatality, 25 percent at crossings with flashing lights were fatal, and only 12.4 percent at crossings with STOP signs were fatal (Raub 2006). Also, gates were found to be associated with mitigated injury severities, compared with those at crossing with only Crossbuck signs, while the flashing lights were found to be related with higher likelihood of injuries and fatalities (Eluru, Bagheri et al. 2012). Studies have also examined the role of socio-demographics in injury severity. Despite studies on the injury severity of crashes at grade crossings (Hu, Li et al. 2010, Russo and Savolainen 2013, Hao and Daniel 2014, Zhao and Khattak 2015), the role of active vs. passive controls remains under-researched.

2.3 Behavioral Considerations

The Highway Safety Manual provides a discussion of factors contributing to crashes, which are helpful in countermeasure selection and implementation. Typically, crashes have multiple contributing factors, e.g., time of day, driver attentiveness, speed, vehicle condition, road design, weather, and glare. Driver factors typically play a large role in a majority of crashes. Therefore, it is critical to consider the role of driver behaviors in crash occurrence and injury severity (AASHTO 2010).

The occurrence of risky driver behaviors may be largely attributed to two factors: 1) inadequate comprehension of the traffic controls at grade crossings and 2) non-compliance with crossing controls. Various studies have evaluated evaluate driver knowledge of traffic control devices at grade crossings. Not all drivers were certain about the required responses at grade crossings with contain controls, e.g., Crossbucks, advance warning signs, and flashing lights (Tidwell Jr and Humphreys 1981, Richards and Heathington 1988).. Some traffic control devices (e.g., STOP signs) that are often used at regular highway intersections may be confusing to some drivers, if used at rail grade crossings (Jeng 2005). In addition to the inadequate comprehension, non-compliance with crossing controls is another issue detrimental to crossing safety. Some drivers may violate crossing controls intentionally if they judge that it is safe to cross the track, even when they are warned about an approaching train (Richards and Heathington 1990, Tenkink and Van der Horst 1990, Abraham, Datta et al. 1998, Witte and Donohue 2000, Gill, Multer et al. 2007). Various reasons, such as long warning times (greater than 60 seconds), avoidance of wait times, and poor sight distance at grade crossings, may also lead motorists not to seeing an oncoming train and engaging in risky crossing behaviors (Richards and Heathington 1990, Abraham, Datta et al. 1998, Liu, Bartnik et al. 2015).

Some risky behaviors, such as “did not stop at crossings,” are associated with an increased likelihood of injury and fatalities (Tidwell Jr and Humphreys 1981, Richards and Heathington 1988, Jeng 2005, Russo and Savolainen 2013). Drivers were more likely to sustain severe injuries if they went around or through the gates, given a crash (Cooper and Ragland 2012, Eluru, Bagheri et al. 2012). Pre-crash driver behaviors, defined as motorist actions prior to a crash, were found to be good indicators of the effectiveness of crossing control devices and they can result in higher or lower injury severity (Liu, Bartnik et al. 2015). This study explores the associations of pre-crash behaviors with crossing controls, and the consequences of motorists engaging in certain behaviors prior to crashes.

2.4 Crash Modeling

Given the ordinal nature of crash severity, ordinary ordered logit/probit models are often estimated (Eluru, Bagheri et al. 2012, Khattak 2013, Russo and Savolainen 2013, Hao and Daniel 2014). Some researchers treat injury severities as discrete outcomes and neglect ordering in the severity, so the multinomial logit model is estimated (Fan and Haile 2014, Zhao and Khattak 2015). Other models such as generalized logit model (Hu, Li et al. 2010), latent segmentation based ordered logit model (Eluru, Bagheri et al. 2012), random parameter logit model (Zhao and Khattak 2015) and mixed generalized ordered logit model (Eluru, Bhat et al. 2008) have been applied to explore the linkages between crash severity and associated factors. Recently, researchers have developed methodologies that predict crash frequency and crash severity simultaneously using multivariate models, such as multivariable Poisson models (Ye, Pendyala et al. 2009, Ye, Pendyala et al. 2013).

Theoretically, driver, vehicle, and roadway factors are correlates of injury severity. However, the relationships between injury severity and correlates are more nuanced and the variables may be inter-related. For example, drivers may be more likely to violate gates (given crashes), resulting in higher injury severity; but at the same time, the presence of gates may serve as a deterrent and associate with lower severity of injury. This method of untangling inter-relationships is known as path analysis, also called Structural Equation Modeling (SEM), which has been used for many driving behavior and traffic crash studies (Khattak and Rocha 2003, Constantinou, Panayiotou et al. 2011, Kim, Pant et al. 2011, Wang and Qin 2014).

Factors, such as driver age, gender, road conditions etc. and the crash severity have been linked directly to the crash severity, and also connected indirectly through latent endogenous variables, such as vehicle speed and collision force (Wang and Qin 2014). Traffic code violations have been used as a bridge between driving outcomes (incidents and offences) and associated factors including driver age, gender and other personality factors (Constantinou, Panayiotou et al. 2011).

In general, most of previous studies on crossing crash injury severity have focused on the direct associations of controls, while the indirect effects of controls through the pre-crash behaviors have not been fully investigated. With the availability of contemporary statistical techniques, it is possible to untangle the chain of behavioral events and other variables that result in more or less severe crash outcomes. Specifically, this study uses path analysis to explore the direct and indirect associations (through driver pre-crash behaviors) of crossing control devices and other factors with crash outcome –injury severity. More specifically, this study quantifies the

additional benefits of installing gates in terms of lowering injury severity for the remaining crashes, noting that some crashes will be eliminated altogether due to gates.

SECTION 3: APPROACH AND METHODOLOGY

3.1 Data Acquisition

The Office of Safety Analysis under FRA maintains a safety website (<http://safetydata.fra.dot.gov/>) to make all railway-related safety data available to the public. Previous studies used data from this source, for different study periods. Most studies used ten-year crash data for safety studies on grade crossings (Raub 2009, Eluru, Bagheri et al. 2012, Russo and Savolainen 2013, Hao and Daniel 2014). Note that for studies on crash frequency/rate, using 10 year data may overcome the low mean number of crashes issue at grade crossings (Lord 2006). However, the reasons for using 10 year data in this study include 1) making full use of the available information, 2) fully capturing the variations in infrastructure, vehicles, and motorists, and 3) with some combinations of crossing controls being rare, using more years of data to increase the size of such combinations. Also, note that the unit of analysis in this study is the crash rather than grade crossing.

For this study, ten-year crash data were extracted from Highway-Rail Grade Crossing Accident/Incident Report (Form 6180.57). The data include crash features (e.g., time, fatality, injury), crossing features (e.g., controls, ID, track class, impact train speed), involved highway user characteristics (e.g., user type, vehicle speed, user age, gender, user behavior), situational factors (e.g., weather, visibility, temperature, view of track) and other information required to be documented by FRA. The data are of reasonably good quality. The data were error checked by the authors and ranges and distributions of key variables were found to be reasonable, compared with equivalent databases used in similar studies (Raub 2009, Eluru, Bagheri et al. 2012, Russo and Savolainen 2013, Hao and Daniel 2014).

3.2 Data Imputation

Crash reports may be incomplete for several reasons, e.g., errors or lack of entry by individuals who code the data, or non-coverage of certain roadways. Observations with missing variables can be removed, if there are few of them or the missing values are randomly distributed. However, if these conditions are not met, e.g., the missing data are non-random, then removing them is inefficient, potentially resulting in estimation and interpretation errors (Gelman and Hill 2006). To improve efficiency of estimation, avoid errors in interpretation, and preserve the population size of available grade crossing crashes, this study applied rigorous data imputation methods for handling missing data (Orchard and Woodbury 1972).

The total number of observations is 15,639. In this study, four variables needed imputation: vehicle speed, train speed, driver age and driver gender. They have missing values for 2.7 percent for train speed (continuous variable), 5.1 percent for vehicle speed (continuous variable), 5.3 percent for driver gender (binary variable) and 16.9 percent for driver age (continuous variable). Missing values for variables are spread across observations. If missing data are simply removed from the dataset, about 23 percent of the observations will be removed, and 12,025 observations will remain. Given the advantages of data imputation technologies, and

to improve the efficiency of estimation and preserve the population size, this study applies a data imputation method for handling the missing information.

The method of multivariate imputation using chained equations (MICE) is applied to impute missing values across several variables (Raghunathan, Lepkowski et al. 2001, Rubin 2004, Royston 2009). The basic idea of MICE is to impute missing values of multiple variables iteratively via a sequential series of univariate imputation models. The MICE imputation procedure can be described in the following equations (StataCorp 2013):

$$\begin{aligned} X_1^{(t+1)} &\sim g_1(X_1|X_2^{(t)}, X_3^{(t)}, \dots, X_n^{(t)}, Z_1^{(t)}, Z_2^{(t)}, Z_3^{(t)}, \dots, Z_k^{(t)}, \varphi_1) \\ X_2^{(t+1)} &\sim g_2(X_2|X_1^{(t+1)}, X_3^{(t)}, \dots, X_n^{(t)}, Z_1^{(t)}, Z_2^{(t)}, Z_3^{(t)}, \dots, Z_k^{(t)}, \varphi_2) \\ X_3^{(t+1)} &\sim g_3(X_3|X_1^{(t+1)}, X_2^{(t+1)}, \dots, X_n^{(t)}, Z_1^{(t)}, Z_2^{(t)}, Z_3^{(t)}, \dots, Z_k^{(t)}, \varphi_3) \\ &\dots \dots \\ X_n^{(t+1)} &\sim g_n(X_n|X_1^{(t+1)}, X_2^{(t+1)}, X_3^{(t+1)}, \dots, X_{n-1}^{(t+1)}, Z_1^{(t)}, Z_2^{(t)}, Z_3^{(t)}, \dots, Z_k^{(t)}, \varphi_n) \end{aligned}$$

Where,

X_1 is the most observed variable except those with complete information;

X_2 is the second most observed variable except those with complete information;

X_3 is the third most observed variable except those with complete information;

X_n is the least observed variable;

$Z_1, Z_2, Z_3, \dots, Z_k$ are variables with complete information;

t is the iteration serial number, $t= 0, 1, 2, \dots$ (normally between 5 and 20);

g_i is the univariate imputation model and can be of a different type (linear, logit, etc.), according to the type of imputing variable X_i , where $i = 0, 1, 2, 3 \dots n$;

φ_i is the corresponding model parameter with a uniform prior, where $i = 0, 1, 2, 3 \dots n$.

MICE simultaneously imputes variables of different types by choosing the appropriate univariate imputation model specifications. Except for complete variables, the most observed variable is train speed that should be imputed first. Since it is a continuous variable, a linear imputation model is needed. The next variable to be imputed is the vehicle speed, for which a linear imputation model was applied. The imputation of missing gender information needs a logit model and driver age was imputed through a linear imputation model. After data imputation, all variables have complete information for 15,639 crashes used for data analysis.

3.3 Conceptual Framework and Data Analysis

Railroad grade crossing controls along with other associated factors can have direct associations with safety outcomes, i.e., injury severity in crashes. Such outcomes have been investigated in previous studies (Eluru, Bagheri et al. 2012, Russo and Savolainen 2013, Fan and Haile 2014, Hao and Daniel 2014, Zhao and Khattak 2015), along with associations of pre-crash behaviors (Eluru, Bagheri et al. 2012, Russo and Savolainen 2013). The pathways embedded in crossing controls, pre-crash behaviors and crash injury severity are generally underexplored. A study by Eluru et al. showed that gate controls were associated with the lowest injury severity given a crash, which is a direct association. The study also revealed that drivers who drove around or through the gates were most likely to be injured severely (Eluru, Bagheri et al. 2012). Given that driving around or through the gates can only occur at crossings equipped with gates, the presence of gates may induce this type of dangerous behavior, if some motorists approaching active controls are likely to violate the intent of the control. Such violations are positively associated with higher injury

severity, which is the indirect association between gates and the injury severity through pre-crash behavior.

The conceptual framework presented in Figure 1 shows that crossing controls can have both a direct association with the safety outcomes and indirect association (dashed line) through dangerous approaching behaviors (captured by pre-crash behaviors). The path analysis helps untangle the two associations, while controlling for other key variables such as train and vehicle speeds, control activation warning time, and driver demographics (Alwin and Hauser 1975).

This study develops a system of 2 equations that specify the linkages between correlates of pre-crash behaviors and injury severity. The intent of the first linkage $\beta_{(i)}$ is to explore the correlates of pre-crash behaviors, defined as drivers' actions prior to the event of a crash. Using their judgments, drivers can choose from a set of actions that can include stopping or not stopping at gates. Their actions are discrete choices that a driver makes prior to the event of a crash, justifying the use of discrete choice modeling techniques (Anas 1983, Abdel-Aty and Abdalla 2004). Specifically, due to its simplicity and ease of interpretation, the multinomial logit model is used for estimation, i.e., to quantify correlates of pre-crash behaviors.

The pre-crash behaviors available in the database include five categories:

1) "Drove around or thru the gate." This pre-crash behavior is clearly in non-compliance with controls; this can be due to a number of reasons including a desire to reduce delay, and/or poor sight distance at grade crossings. In such cases, this will be a deliberate attempt to cross before the arrival of a train and it will be coded as trespassing (FRA 2011). In crashes, males were found to be much more likely to engage in driving around or thru gates than females (Cooper and Ragland 2012).

2) "Stopped and then proceeded." This pre-crash behavior indicates initial compliance with controls.

3) "Did not stop." In this case, the vehicle should be moving over the crossing, with at least some speed at the time of collision. This pre-crash behavior also represents non-compliance with grade crossing controls.

4) "Stopped on crossing." This pre-crash action can be attributed to factors such as vehicle malfunction, inability of vehicle to clear the tracks (getting stuck), or vehicle getting trapped as gates come down, and deliberate action by the motorist.. Whether the person is in the vehicle at the time of collision is critical to injury severity.

5) "Other behaviors." These behaviors can include suicide or attempting suicide, going thru/around temporary construction barriers, and shoving onto the track. Ideally, such behaviors should be analyzed separately from the other behaviors as they are a mixture of intentional and unintentional harm. However, data availability precludes this.

The second model uncovers the correlations between the injury severity and associated factors including the pre-crash behaviors. Given the ordinal nature of the injury severity, the ordered logit model is applied due to its simplicity and ease of interpretation (Eluru, Bagheri et al. 2012, Russo and Savolainen 2013). The following equations describe the path analysis structure, as shown in Figure 1 (Khattak and Rocha 2003):

$$Y_1 = X\beta_{(i)} + \varepsilon_1 \tag{1}$$

$$Y_2 = X\beta + Y_1\gamma + \varepsilon_2 \tag{2}$$

Where,

Y_1 is the pre-crash behavior;

$\beta_{(i)}$ is a set of coefficients corresponding to the i^{th} behavior in the multinomial logit model;

X is a vector of explanatory variables that include grade crossing controls;
 Y_2 is the safety outcome, i.e., injury severity measured on ordinal scale;
 β is a set of coefficients of X in the ordered logit model;
 γ is the association of driver behavior with injury severity, estimated in the ordered logit model;

ε_1 and ε_2 are error terms, which are assumed to be uncorrelated.

In the multinomial logit model, the corresponding probability of each outcome (i.e., one type of pre-crash behaviors) is given by (StataCorp 2013):

$$\Pr(Y_1 = 1) = \frac{\exp(X\beta_{(1)})}{\exp(X\beta_{(1)}) + \exp(X\beta_{(2)}) + \dots + \exp(X\beta_{(n)})} \quad (3)$$

$$\Pr(Y_1 = 2) = \frac{\exp(X\beta_{(2)})}{\exp(X\beta_{(1)}) + \exp(X\beta_{(2)}) + \dots + \exp(X\beta_{(n)})} \quad (4)$$

.....

$$\Pr(Y_1 = i) = \frac{\exp(X\beta_{(i)})}{\exp(X\beta_{(1)}) + \exp(X\beta_{(2)}) + \dots + \exp(X\beta_{(n)})} \quad (5)$$

Where,

Y_1 is the outcome variable, i.e., pre-crash behaviors such as driving around gates;

X is a vector of explanatory variables, e.g., grade crossing controls;

$\beta_{(i)}$ is a set of coefficients corresponding to the i^{th} behavior, $i = 1, 2, \dots, n$.

Based on the findings from injury severity studies, the ordered logit/probit models were found to be appropriate (Savolainen, Mannering et al. 2011, Yasmin and Eluru 2013, Mannering and Bhat 2014, Yasmin, Eluru et al. 2014, Ye and Lord 2014). Therefore, this study uses an ordered logit regression model to explore the relationships between injury severity and pre-crash behaviors and crossing controls. In this case, $Y_2 = 1$ represents a non-injury crash ($Y_2 = 2$ is non-fatal injury), and $Y_2 = 3$ is fatal injury; the predicted probabilities for each category are (StataCorp 2013) :

$$\Pr(Y_2 = 1) = \frac{\exp(\alpha_1 - X\beta)}{1 + \exp(\alpha_1 - X\beta)} \quad (6)$$

$$\Pr(Y_2 = 2) = \frac{\exp(\alpha_2 - X\beta)}{1 + \exp(\alpha_2 - X\beta)} - \frac{\exp(\alpha_1 - X\beta)}{1 + \exp(\alpha_1 - X\beta)} \quad (7)$$

$$\Pr(Y_2 = 3) = 1 - [\Pr(Y_2 = 1) + \Pr(Y_2 = 2)] = 1 - \frac{\exp(\alpha_2 - X\beta)}{1 + \exp(\alpha_2 - X\beta)} \quad (8)$$

Where,

Y_2 is the injury severity level;

X is a vector of explanatory variables, e.g., crossing controls and pre-crash behaviors;

β is a set of coefficients of X .

Unlike the multinomial logit model where each outcome can have its own set of coefficients, the ordered outcomes model shares one set of coefficients but with different intercepts which are the cut-points α_1 and α_2 in above model equations (StataCorp 2013).

Path analysis allows two or more models to be estimated simultaneously. The estimated parameters can be used to calculate the direct, indirect and total effects. Marginal effects provide the change in response values associated with one unit change in explanatory variables. For

ordinary least squares models, the parameters are the marginal effects. However, for other models estimated in this study, i.e., the multinomial and ordered logit regression models, the estimated coefficients are the log of odds ratio. To interpret modeling results and obtain the direct, indirect and total effects of variables, the marginal effects for multinomial logit model must be calculated as follows (Wang 2008):

$$\frac{\partial \Pr(Y_1 = i | \beta_{(i)} X)}{\partial X} = \frac{\exp(X\beta_{(i)}) \beta'_{(i)} + \exp(X\beta_{(i)}) [\sum_{j=1}^n \exp(X\beta_{(j)}) (\beta'_{(i)} - \beta'_{(j)})]}{[1 + \sum_{j=1}^n \exp(X\beta_{(j)})]^2} \quad (9)$$

Where,

Y_1 is the outcome, i.e., pre-crash behaviors;

$\beta_{(i)}$ is a set of coefficients corresponding to the i^{th} outcome, $i = 1, 2, \dots, n$;

$\beta_{(j)}$ is a set of coefficients corresponding to the j^{th} outcome, $j = 1, 2, \dots, n$;

X is a vector of explanatory variables.

Note that in multinomial logit models, the signs of marginal effects are not always consistent with the sign of the coefficients (β), while in some logit regressions (e.g., binary logit) they are consistent. This is because the marginal effect depends on the values and levels of other variables. As the values of other variables and the variables in equation change, the signs of the marginal effect can also change. Note that the signs of marginal effects are determined by $\beta'_{(i)} - \beta'_{(j)}$, which is the differences of coefficients under different outcomes. The marginal effects show the probability change of one outcome compared with the base level chosen among the five types of pre-crash behaviors (Wang 2008).

The marginal effects for the ordered logit model can be computed as follows (Wang 2008):

$$\frac{\partial \Pr(Y_2 = 1 | \beta X)}{\partial X} = \frac{\exp(\alpha_1 - X\beta)}{[1 + \exp(\alpha_1 - X\beta)]^2} \beta' \quad (10)$$

$$\frac{\partial \Pr(Y_2 = 2 | \beta X)}{\partial X} = -\frac{\exp(X\beta - \alpha_2)}{[1 + \exp(X\beta - \alpha_2)]^2} \beta' - \frac{\exp(\alpha_1 - X\beta)}{[1 + \exp(\alpha_1 - X\beta)]^2} \beta' \quad (11)$$

$$\frac{\partial \Pr(Y_2 = 3 | \beta X)}{\partial X} = \frac{\exp(X\beta - \alpha_2)}{[1 + \exp(X\beta - \alpha_2)]^2} \beta' \quad (12)$$

Where,

Y_2 is the injury severity level;

X is a vector of explanatory variables, e.g., crossing controls and pre-crash behaviors;

β is a set of coefficients of X ;

α_1 is the constant term for predicting level-1 outcome;

α_2 is the constant term for level-2 outcome.

Marginal effects in ordered logit model are the changes of probability from a lower level to a higher level associated with a unit change in the independent variable.

SECTION 4: FINDINGS

4.1 Descriptive Analysis

Table 3 provides the descriptive statistics, including before and after data imputation. The imputation results show that imputed variables have a similar distribution with their original

distributions. The data were error checked. The ranges and distributions of key variables were reasonable, compared with equivalent databases used in similar studies (Raub 2009, Eluru, Bagheri et al. 2012, Russo and Savolainen 2013, Hao and Daniel 2014).

Gates, flashing lights and audible warnings can be individually or simultaneously present at railroad grade crossings. This study uses the following combinations of active controls: 1) gates only, 2) flashing lights only, 3) audible warning device only, 4) gates and flashing lights, 5) gates and audible, 6) flashing lights and audible, and 7) gates and flashing lights and audible warning. Among the 15,639 collisions, 26.18 percent occurred at crossings equipped with gates and flashing lights and audible warning devices, 10.72 percent with flashing lights and audible warning devices, and 13.79 percent with STOP signs. About 53.6 percent were active crossings with a recommended a 20 second minimum warning time, 1.4 percent with a warning time longer than 60 seconds and 2.18 percent with a short warning time (less than 20 seconds, or not activated when collisions occurred).

Fatal injury, non-fatal injury and property-damage-only (PDO) crashes constitute 8.02%, 25.52% and 66.46% of the 15,639 reported crashes respectively. Drivers' pre-crash behaviors were extracted specifically from the database. This study treats pre-crash behaviors as an intermediate safety outcome and uses the path analysis to link crossing controls and injury severity using driver behavior as a bridge. Pre-crash behaviors were reported in five categories: 1) "Drove around or thru the gate," 2) "Stopped and then proceeded," 3) "Did not stop," 4) "Stopped on crossing," 5) and other behaviors, e.g., suicide or attempting suicide, going thru/around temporary construction barriers, and shoving onto the track. Note that nearly one-half of drivers were out of their vehicle at the time their vehicle was "Stopped on crossing". Therefore, some of these drivers were able to get out of harm's way and did not receive severe injuries. A study by Eluru et al. did not split the cases of drivers stopping on crossing and thus found that drivers who stopped on crossings were the least likely to be injured or killed, which is counterintuitive (Eluru, Bagheri et al. 2012). By splitting the pre-crash behavior of "Stopped on crossing" further into driver inside or out of their vehicle, this study was able to obtain deeper insights into the role of pre-crash behaviors.

4.2 Driver Behavior and Crossing Controls

Table 4 presents the results of pre-crash behavioral propensities, given a set of factors. The goodness of fit for the final model seems reasonable, and the parameter signs are as expected. The models presented in this paper were selected based on theoretical considerations, i.e., hypotheses shown in the conceptual framework) and statistical properties that include goodness of fit, statistical significance of the overall model and specific variables, as well as their magnitude and signs. The selected models contain some statistically insignificant variables, either because there is strong theoretical justification to include them or they are part of a group of variable.

Note that the pre-crash driver behavior – "Drove around or thru the gate" requires the presence of gates. Therefore, cases without gate controls are not relevant and excluded from the model. The results show strong associations between behaviors and different settings of gate controls. Compared with the base of "gates, flashing lights, and audible warnings," the chances of driving around or thru gates were higher when gates did not have flashing lights and/or audible warnings. The marginal effects show a higher possibility (3.7%, 4.0% and 6.5%) of

driving around or thru the gate when crossings were equipped with gates only, gates and flashing lights, and gates and audible warnings, respectively.

Importantly, the results show that compared with the base (a fully controlled crossing) there is a substantially higher possibility of not stopping at crossings without gates, as shown in Table 4. For example, when only flashing lights and audible warnings are present, the chances of not stopping were higher by 43.9 percent, given crashes. Furthermore, the chances of “Stopped on crossing (and driver in vehicle)” were higher when crossings were passively controlled, i.e. when STOP sign was present (5.0 percent higher) or only Crossbuck was present (7.1 percent higher).

Advanced warning time had significant correlations with pre-crash behaviors. Note that warning time refers to how many seconds the controls are activated before the train enters the crossing. The base for warning time in the model is minimum warning time of 20 seconds, recommended by the Federal Railroad Administration (Bowman 1987, Richards, Heathington et al. 1990). Results show that active controls with warning time less than 20 seconds were associated with a lower chance of first two behaviors. Notably, if the warning time is less than 20 seconds, there is a 25.8 percent higher chance of not stopping (“Did not stop” behavior).

Higher train speeds were related to increasing occurrence of “Stopped on crossing” and decreasing occurrences of other behaviors. The first three behaviors were associated with higher vehicle speeds. Before crashes, male drivers were more likely to drive around or through the gates, less likely to stop first and then proceed, and more likely to stop on crossings but get out of vehicle. Older drivers were more likely to stop first and then proceed, less likely to not stop, and more likely to be in the vehicle when the vehicle stopped on the rail crossing. In addition, pick-up truck drivers are less likely to go around the gates compared with passenger car drivers.

4.3 Injury Severity and Crossing Controls

Table 5 presents the correlates of injury severity tested by estimating ordered logit regression model. The final model specification was selected based on theoretical considerations and empirical properties of the model. All pre-crash behaviors were significantly correlated to driver injury severity and their effects are substantial. When drivers “Drove around or thru the gate” or “Did not stop” they had a substantially higher chance injury in crashes, at 40.0 and 39.5 percent respectively, than the base behavior, which was “Stopped on crossing (and driver out of vehicle).” For “Stopped and then proceeded” and “Stopped on crossing (driver in vehicle)” the chances of injury were higher by 33.2 and 23.7 percent, respectively.

Crossing controls did not have a strong direct relationship with driver injury severity. For crashes occurring at active crossings with warning time less than 20 seconds, there was a 9.4 percent higher chance of injury, compared with crossings given at least 20 second warning time.

As expected, train speed and vehicle speed had a strong association with the injury severity. A 10 mph increase in trains approaching speed was associated with 8 percent higher possibility of drivers getting injured. A 10 mph increase in vehicle collision speed was related to a 4 percent lower possibility of drivers getting injured. The model also shows that male drivers were less likely to get injured in crashes than female drivers. Older drivers were more likely to get injured, given a crash.

4.4 Path Analysis

From the above discussion, we know that the direction associations between controls and injury severity were not very significant. Nevertheless, some crossing controls have very significant correlations with pre-crash behaviors which were in turn significantly associated with injury severity. Crossing controls have substantial indirect impacts on injury severity through pre-crash behaviors. Path analysis allows us to compare both direct and indirect effects and combine them to provide total impacts of controls on injury severity, as outlined in the structure presented in Figure 1. The impact of each variable on pre-crash behaviors was calculated by using the marginal effects from the multinomial logit regression model, while the impact of each variable on levels injury severity was calculated by using the marginal effects from the ordered logit regression model.

Table 6 shows the associations—the total effect is the sum of the direct and indirect effects manifested through pre-crash behaviors. Note that, non-statistically significant effects were omitted for clarity. The injury severity model in Table 5 shows that “Did not stop” was associated with a 39.5 percent higher probability of driver getting injured, while the presence of Crossbuck-only was related to a 45 percent higher chance of this behavior. Thus, through this behavior, the indirect impact of Crossbuck-only on increasing the probability of driver getting injured is $45\% \times 39.5\% = 17.8\%$. The indirect effect of Crossbuck-only through other behaviors on increasing/decreasing injury probability was 0%, -0.5%, 1.7% and 0% respectively. The model in Table 5 shows that the direct effect of Crossbuck-only on injury severity was not statistically significant (at p-value = 0.05 level), thus the direct impact of gates on injury severity was zero percent. The net or total impact of Crossbuck-only at crossings on the probability of driver getting injured is $0\% - 0.5\% + 17.8\% + 1.7\% + 0\% = 18.9\%$ (or an overall 18.9 percent higher chance of injury). In the same way, relationship between other controls and injury severity can be interpreted through path analysis. In general, the cases with gates are associated with lower chance of injury compared with other cases, especially the passive controls (i.e., STOP sign and Crossbuck-only). This makes sense because gates more effectively prevent dangerous driving behaviors. Note that, this study also estimated the ordered logit model for injury severity without the pre-crash behaviors (modeling results are available from authors). The results show that not all direct effects of crossing controls on injury severity are statistically significant. Clearly, path analysis better connects the associations of controls on injury through pre-crash behaviors.

The results also show that the warning time less than 20 seconds was related to a 19 percent higher chance of injury, compared with a minimum of 20 second warning time. Higher train approaching speed and vehicle collision speed also had an association with higher injury severity. Male drivers were less likely to be injured in crashes perhaps because of their better survivability. Old drivers were more likely to get injured in crashes.

SECTION 5: DISCUSSION

By investigating the role of pre-crash behaviors, i.e., their correlations with injury severity in crashes, this study uncovered the correlations between injury severity and the type of crossing controls. The study found that outcomes of crashes at passive crossings (e.g., STOP signs or only Crossbuck signs) are more severe than those of crashes at active crossings (i.e., gates, flashing lights and audible warning devices). Some researchers have argued against the use of passive controls (e.g., highway STOP signs) at grade crossings. Burnham’s study found that only 18 percent motorists were alerted to the STOP signs and 82 percent were either confused or semi-

confused about the presence of STOP signs at grade crossings (Burnham 1995). Among all types of crossings, those with only Crossbuck signs are associated with the highest probability of injuries given a crash. It is notable that compared with STOP signs and active controls, Crossbucks may not be able to gain a drivers' attention (Sanders, McGee et al. 1978) and they also have higher crash rates (Eck and Shanmugam 1987).

Within actively controlled crossings, crashes at crossings with gate controls are related to a lower chance of injury, compared with crossing that do not have gates (i.e., only flashing lights, only audible devices, or flashing lights plus audible devices). The relatively large difference can be explained by the impact of gates on pre-crash behaviors. Flashing lights and audible warnings are used to warn drivers about the presence of a train. However, they do not deter dangerous behaviors as much as gates do, by physically stopping drivers from crossing when gates are lowered. Nevertheless, if drivers notice that the train is still far away from the crossing (i.e., they have longer warning time), they may still decide to attempt to negotiate the crossing maneuver. Meeker et al. found that the addition of automatic gates reduced the likelihood of drivers crossing in front of trains from 67 to 38 percent (Meeker, Fox et al. 1997). Correspondingly, this study revealed that gates are related to an average 30 percent lower chance of the "Did not stop" behavior. Within gate crossings, the differences in injury severities are not substantial but they cannot be ignored. The study revealed that gate crossings with flashing lights but not bells are more likely to be associated with driving around or through the gates, which is intentional trespassing behavior (FRA 2011). The violation of gates is in turn related to 40 percent higher chances of injuries than other pre-crash behaviors. As a result, the presence of "Gates plus Flashing lights" is associated with 2.80 percent higher levels of injuries. The two other types of gate crossings missing flashing lights, or bells, or both have the same gate violation issue.

As mentioned in the literature review, the current Crash Modification Factors (CMFs) documented in the Clearinghouse website (<http://www.cmfclearinghouse.org/>) are for the crash rate reductions. Information about the changes of the crash injury severity by implementing certain countermeasures at highway-rail grade crossings is not available. Lack of information about severity may miss some of the benefits if reductions in injury severity are not accounted for. While this study is correlational, it provides information about the injury severity, e.g., with only flashing lights and audible devices there was a 16.2 percent higher chance of injury, compared to crossing with gates, flashing lights and audible devices. Thus, adding gates to a crossing with flashing lights and audible warnings, on average, is correlated with 16.2 percent lower chances of injury, given a crash (in addition to reducing crash frequency). Note that the initial CMF for installing gates at crossings that previously had flashing lights and audible devices is 0.55 (Elvik and Vaa 2004), indicating that adding gates can reduce the crashes by 45 percent. Using the results from this study, a CMF related to injury severity can be developed.

Table 7 presents a simple illustrative example that integrates the changes in crash rates and injury severity. Suppose there are 100 crashes in a year at a crossing with flashing lights and bells. By adding gates the expected number of crashes will be 55. The initial distribution of injury crashes at crossings with flashing lights and bells is 36.87 percent, by examining the raw crash data used in this study. We can assume that out of 100 crashes, 36.87 are injury crashes and the rest (63.13) are PDO crashes; with the installation of gates the distribution will be 20.28 injury crashes and 34.72 PDO crashes. Installation of gates is associated with 16.2 percent lower injury severity, meaning that the new distribution of injury crashes would be 30.90% ($= 36.87\% \times (1 - 0.162)$). Therefore the modified distribution will be 16.99 injury ($= 55 \times 30.90\%$) and

38.01 PDO crashes (= 55 – 16.99). According to the raw crash data used in this study, the observed probability of injury crashes at crossings with gates, flashing lights and bells, is 30.18 percent, which is very close to the expected distribution of injury crashes (i.e., 30.9 percent). Generally, the distribution of expected injury and PDO crashes after countermeasure implementation will be given by the following equations:

$$\text{Expected Number of Injury Crashes-After} = (CMF \times N) \times P_{inj} \times (1 - CMFSEV) \quad (13)$$

$$\text{Expected Number of PDO Crashes-After} = (CMF \times N) - (CMF \times N) \times P_{inj} \times (1 - CMFSEV) \quad (14)$$

Where,

CMF is the Crash Modification Factor;

N is the number of crashes before implementation of countermeasure;

P_{inj} is the probability of injury, between 0 and 1 before countermeasure implementation;

CMFSEV is the Crash Modification Factor for injury severity.

SECTION 6: LIMITATIONS

While several key factors were considered in this study, there might be other factors (such as weather or terrain) that may be important to injury severity. Thus, this study is limited by the number of factors considered in the analysis. Hauer et al. has pointed out that the safety of a rail-highway grade crossing should be estimated by mixing information about causal factors (Hauer and Persaud 1987). However, firm causal inferences could not be drawn in this study given the cross-sectional nature of the data.

Another limitation of this study is the inherent unpredictability of human behavior. Even though certain outcomes may be correlated with certain human characteristics, driver behavior is a moving target and it changes over time.

The data were taken from an outside source and analyzed using parametric statistical methods. As a non-observational study, researchers did not control the variation in the variables and also some of the variables had missing data. The data imputation model was applied to recover missing data. However, the imputed data cannot be independently verified.

Since the database spans over 10 years, there may be variation in reporting procedures. There may be errors or discrepancies between reported data values, e.g., coding errors, coverage and geo-coding errors, or measurement errors since crash inspectors/investigators are typically not on the scene when a crash occurs.

SECTION 7: CONCLUSIONS AND CONTUNING RESEARCH

Many previous studies have not included crossing controls in models of injury severity, indicating an assumption that crossing controls are not significantly correlated with injury severity, given a crash. This study explicitly investigated the role of controls and found that crossing controls do not have significant direct association with injury severity. However, crossing controls were found to be significantly correlated with pre-crash behaviors and also pre-crash behaviors were significantly associated with injury severity. Thus through path analysis, this study uncovered the indirect role of crossing controls with injury severity. Furthermore,

using path analysis, this study answers why some controls are associated with higher injury severity, owing to their impacts on motorist behaviors prior to the event of a crash.

This study found that crashes occurring at rail crossings equipped with gates are generally less injurious. More broadly, active controls were associated with lower driver injuries compared with passive controls, given a crash. Also, active crossings with short warning times pose higher risks of injuries in crashes at grade crossings. A minimum of 20 second warning time seems appropriate at grade crossings, based on evidence from injury data. Relationships between driver behavior, injury severity and other factors (trains speed, vehicle speed, and demographics) were also quantified in this study. Higher train speed and vehicle speeds were statistically significantly associated with higher injury severity, and females and older drivers were more likely to be injured in railroad grade crossing crashes, consistent with the findings from previous studies (Russo and Savolainen 2013,Zhao and Khattak 2015).

For practicing engineers and planners, the study provides a quantification of injury risk factors and points to risky behaviors that are associated with higher injury severity. Such behaviors (e.g., not stopping or driving around gates) may be minimized through additional countermeasures that can further lower injury severity. This study helps us understand that potential effectiveness of control treatments is likely to depend on how drivers respond to treatments and how they change their behaviors. Specifically, this study generates information that quantifies reductions in injury severity associated with various controls. It provides valuable information for reductions in severity that is relevant to Crash Modification Factors used in Highway Safety Manual. For example, adding gates to a crossing with flashing lights and audible warnings is associated with 16.2 percent lower chances of injury, given a crash, in addition to reducing crash frequency. Integrating the initial CMF from the Highway Safety Manual, this study provides information that is useful for quantifying changes in injury severity distributions, post implementation. The results of this study are also of national interest because the crashes analyzed in the study are dispersed throughout the country.

This study reveals the correlations between crossing controls and pre-crash behaviors, which is useful in explaining safety outcomes. More research is needed to explain why drivers chose their pre-crash actions using the narrative information recorded in the database (Form 6180.57). Future research should do a deeper exploration of spatial distributions, injury severity, and the role of dangerous pre-crash driver behaviors, especially of intentional trespassing behaviors such as driving around or through gates. While this study found that creating new variables—whether a vehicle stopped on the tracks had the driver or not was critical to injury outcomes, exploring why vehicles stop on crossings and how to prevent such stoppings can be helpful. Also, other aspects of safety outcomes can be observed at railroad grade crossings by integrating more databases, such as Railroad Injury and Illness Summary (Form 6180.55a) and Rail Equipment Accident/Incident Report (Form 6180.54). These two databases provide more details of injuries, e.g., injury nature, medication care, and the rail equipment damage caused by crashes. The data can provide a broader sense of the seriousness and cost of train-car crashes. Collisions between trains and other types of vehicles, such as trucks and buses, can also be explored using the methodology developed by this study. More research is also needed to better support the decisions of closing railroad grade crossings or going to passive or active controls, and providing insights to agencies and railroad companies regarding countermeasures for improving safety at highway-rail grade crossings.

SECTION 8: REFERENCES

- AASHTO. 2010. Highway Safety Manual. American Association of State Highway and Transportation Officials, Washington, DC.
- Abdel-Aty, M., and M. F. Abdalla. 2004. Modeling drivers' diversion from normal routes under ATIS using generalized estimating equations and binomial probit link function. *Transportation* 31, 327-348.
- Abraham, J., T. K. Datta, and S. Datta. 1998. Driver behavior at rail-highway crossings. *Transportation Research Record: Journal of the Transportation Research Board* 1648, 28-34.
- Alwin, D. F., and R. M. Hauser. 1975. The decomposition of effects in path analysis. *American Sociological Review*, 37-47.
- Anas, A. 1983. Discrete choice theory, information theory and the multinomial logit and gravity models. *Transportation Research Part B: Methodological* 17, 13-23.
- Austin, R. D., and J. L. Carson. 2002. An alternative accident prediction model for highway-rail interfaces. *Accident Analysis & Prevention* 34, 31-42.
- Bowman, B. L. 1987. The effectiveness of railroad constant warning time systems. *Transportation Research Record*.
- Burnham, A. 1995. Stop Sign Effectiveness at Railroad Grade Crossings (Abuse without Excuse). In: *Third International Symposium on Railroad-Highway Grade Crossing Research and Safety*
- Constantinou, E., G. Panayiotou, N. Konstantinou, A. Loutsiou-Ladd, and A. Kapardis. 2011. Risky and aggressive driving in young adults: Personality matters. *Accident Analysis & Prevention* 43, 1323-1331.
- Cooper, D. L., and D. R. Ragland. 2012. Applying Safety Treatments To Rail-Highway At-Grade Crossings.
- Eck, R., and R. Shanmugam. 1987. Physical and operational characteristics of rail-highway grade crossings on low-volume roads. *Transportation Research Record*.
- Eluru, N., M. Bagheri, L. F. Miranda-Moreno, and L. Fu. 2012. A latent class modeling approach for identifying vehicle driver injury severity factors at highway-railway crossings. *Accident Analysis & Prevention* 47, 119-127.
- Eluru, N., C. R. Bhat, and D. A. Hensher. 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis & Prevention* 40, 1033-1054.
- Elvik, R., and T. Vaa. 2004. The handbook of road safety measures. Amsterdam: Elsevier, ISBN 0-08-044091-6.
- Elvik, R., T. Vaa, A. Erke, and M. Sorensen. 2009. The handbook of road safety measures. Emerald Group Publishing.
- Fan, W., and E. W. Haile. 2014. Analysis of Severity of Vehicle Crashes at Highway-Rail Grade Crossings: Multinomial Logit Modeling. In: *Transportation Research Board 93rd Annual Meeting*
- FRA. 2013. All Highway-Rail Incidents at Public and Private Crossings, 1981-2013. <http://oli.org/about-us/news/collisions-casulties> Accessed March 26 2014.
- Gelman, A., and J. Hill. 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.

- Gill, M., J. Multer, and M. Yeh. 2007. Effects of active warning reliability on motorist compliance at highway-railroad grade crossings. DOT-VNTSC-FRA-09-04.
- Gross, F., B. Persaud, and C. Lyon. 2010. A guide to developing quality crash modification factors.
- Hao, W., and J. Daniel. 2014. Motor vehicle driver injury severity study under various traffic control at highway-rail grade crossings in the United States. *Journal of Safety Research* 51, 41-48.
- Hauer, E., and B. Persaud. 1987. How to estimate the safety of rail-highway grade crossings and the safety effects of warning devices. Transportation Research Board.
- Hu, S.-R., C.-S. Li, and C.-K. Lee. 2010. Investigation of key factors for accident severity at railroad grade crossings by using a logit model. *Safety science* 48, 186-194.
- Jeng, O.-J. 2005. Survey of Driver Perceptions of Railroad and Light Rail Warning Devices/Grade Crossings, New Jersey Institute of Technology.
- Khattak, A. 2013. Severity of Pedestrian Crashes at Highway-Rail Grade Crossings. In: *Transportation Research Board 92nd Annual Meeting*
- Khattak, A., and M. Rocha. 2003. Are SUVs" Supremely Unsafe Vehicles"?: Analysis of Rollovers and Injuries with Sport Utility Vehicles. *Transportation Research Record: Journal of the Transportation Research Board* 1840, 167-177.
- Kim, K., P. Pant, and E. Yamashita. 2011. Measuring influence of accessibility on accident severity with structural equation modeling. *Transportation Research Record: Journal of the Transportation Research Board* 2236, 1-10.
- Lenné, M. G., C. M. Rudin-Brown, J. Navarro, J. Edquist, M. Trotter, and N. Tomasevic. 2011. Driver Behaviour at Rail Level Crossings: Responses to Flashing Lights, Traffic Signals and Stop Signs in Simulated Rural Driving. *Applied Ergonomics* 42, 548-554. doi:<http://dx.doi.org/10.1016/j.apergo.2010.08.011>.
- Liu, J., B. Bartnik, S. H. Richards, and A. J. Khattak. 2015. Driver Behavior at Railway-Highway Grade Crossings with Passive Traffic Controls: A Driving Simulator Study. *Journal of Transportation Safety and Security*. doi:<http://dx.doi.org/10.1080/19439962.2015.1043478>.
- Lord, D. 2006. Modeling motor vehicle crashes using Poisson-gamma models: Examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. *Accident Analysis & Prevention* 38, 751-766.
- Mannering, F. L., and C. R. Bhat. 2014. Analytic methods in accident research: Methodological frontier and future directions. *Analytic methods in accident research* 1, 1-22.
- Meeker, F., D. Fox, and C. Weber. 1997. A Comparison of Driver Behavior at Railroad Grade Crossings with Two Different Protection Systems. *Accident Analysis & Prevention* 29, 11-16. doi:[http://dx.doi.org/10.1016/S0001-4575\(96\)00055-3](http://dx.doi.org/10.1016/S0001-4575(96)00055-3).
- Millegan, H., X. Yan, S. Richards, and L. Han. 2009. Evaluation of Effectiveness of Stop-Sign Treatment at Highway–Railroad Grade Crossings. *Journal of Transportation Safety & Security* 1, 46-60.
- Mok, S. C., and I. Savage. 2005. Why Has Safety Improved at Rail - Highway Grade Crossings? *Risk Analysis* 25, 867-881.
- Ogden, B. D. 2007. Railroad-Highway Grade Crossing Handbook FHWA-SA-07-010 Institute of Transportation Engineers.
- Oh, J., S. P. Washington, and D. Nam. 2006. Accident prediction model for railway-highway interfaces. *Accident Analysis & Prevention* 38, 346-356.

- Orchard, T., and M. A. Woodbury. 1972. A missing information principle: theory and applications. In: *Proceedings of the 6th Berkeley Symposium on mathematical statistics and probability*. p 697-715.
- Park, Y.-J., and F. F. Saccomanno. 2005. Collision frequency analysis using tree-based stratification. *Transportation Research Record: Journal of the Transportation Research Board* 1908, 121-129.
- Raghunathan, T. E., J. M. Lepkowski, J. Van Hoewyk, and P. Solenberger. 2001. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey methodology* 27, 85-96.
- Raub, R. A. 2006. Examination of Highway-Rail Grade Crossing Collisions Over 10 Years in Seven Midwestern States. *ITE Journal* 76, 16-26.
- Raub, R. A. 2009. Examination of highway-rail grade crossing collisions nationally from 1998 to 2007. *Transportation Research Record: Journal of the Transportation Research Board* 2122, 63-71.
- Richards, S. H., and K. W. Heathington. 1988. Motorist understanding of railroad-highway grade crossing traffic control devices and associated traffic laws. Transportation Research Board.
- Richards, S. H., and K. W. Heathington. 1990. Assessment of warning time needs at railroad-highway grade crossings with active traffic control. *Transportation Research Record*.
- Richards, S. H., K. W. Heathington, and D. B. Fambro. 1990. Evaluation of constant warning times using train predictors at a grade crossing with flashing light signals. *Transportation Research Record*.
- Royston, P. 2009. Multiple imputation of missing values: further update of ice, with an emphasis on categorical variables. *Stata Journal* 9, 466.
- Rubin, D. B. 2004. Multiple imputation for nonresponse in surveys. John Wiley & Sons.
- Russo, B., and P. Savolainen. 2013. An Examination of Factors Affecting the Frequency and Severity of Crashes at Rail-Grade Crossings. In: *Transportation Research Board 92nd Annual Meeting*
- Saccomanno, F. F., P. Y.-J. Park, and L. Fu. 2007. Estimating countermeasure effects for reducing collisions at highway-railway grade crossings. *Accident Analysis & Prevention* 39, 406-416.
- Sanders, J. H., H. W. McGee, and C. S. Yoo. 1978. Safety Features of Stop Signs at Rail-highway Grade Crossings. FHWA-RD-78-40 Final Rpt.
- Savolainen, P. T., F. L. Mannering, D. Lord, and M. A. Quddus. 2011. The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis & Prevention* 43, 1666-1676.
- StataCorp. 2013. Stata Multiple-Imputation Reference Manual.
- Tenkink, E., and R. Van der Horst. 1990. Car driver behavior at flashing light railroad grade crossings. *Accident Analysis & Prevention* 22, 229-239.
- Tidwell Jr, J. E., and J. B. Humphreys. 1981. Driver knowledge of grade-crossing information. *Transportation Research Record*.
- Wang, K., and X. Qin. 2014. Use of Structural Equation Modeling to Measure Severity of Single-Vehicle Crashes. *Transportation Research Record: Journal of the Transportation Research Board* 2432, 17-25.
- Wang, P. 2008. Financial Econometrics 2nd Edition. Routledge.

- Witte, K., and W. A. Donohue. 2000. Preventing vehicle crashes with trains at grade crossings: the risk seeker challenge. *Accident Analysis & Prevention* 32, 127-139.
- Yan, X., L. D. Han, S. Richards, and H. Millegan. 2010a. Train–Vehicle Crash Risk Comparison Between Before and After Stop Signs Installed at Highway–Rail Grade Crossings. *Traffic injury prevention* 11, 535-542.
- Yan, X., S. Richards, and X. Su. 2010b. Using Hierarchical Tree-Based Regression Model to Predict Train–Vehicle Crashes at Passive Highway-Rail Grade Crossings. *Accident Analysis & Prevention* 42, 64-74. doi:<http://dx.doi.org/10.1016/j.aap.2009.07.003>.
- Yasmin, S., and N. Eluru. 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. *Accident Analysis & Prevention* 59, 506-521.
- Yasmin, S., N. Eluru, and S. V. Ukkusuri. 2014. Alternative ordered response frameworks for examining pedestrian injury severity in New York City. *Journal of Transportation Safety & Security* 6, 275-300.
- Ye, F., and D. Lord. 2014. Comparing three commonly used crash severity models on sample size requirements: multinomial logit, ordered probit and mixed logit models. *Analytic methods in accident research* 1, 72-85.
- Ye, X., R. M. Pendyala, V. Shankar, and K. C. Konduri. 2013. A simultaneous equations model of crash frequency by severity level for freeway sections. *Accident Analysis & Prevention* 57, 140-149.
- Ye, X., R. M. Pendyala, S. P. Washington, K. Konduri, and J. Oh. 2009. A simultaneous equations model of crash frequency by collision type for rural intersections. *Safety Science* 47, 443-452.
- Zhao, S., and A. Khattak. 2015. Motor vehicle drivers' injuries in train–motor vehicle crashes. *Accident Analysis & Prevention* 74, 162-168.

LIST OF TABLES AND FIGURES

Table 1. Selected Studies on Crash Frequencies/Rates.....	71
Table 2. Selected Studies on Crash Injury Severity and Behavioral Considerations.....	71
Table 3. Descriptive Statistics for Railroad Grade Crossing Crashes-Fra Database.....	72
Table 4. Models for Driver Behavior Before Crashes on Railroad Grade Crossings	74
Table 5. Ordered Logit Regression Model for Injury Severity in Railroad Grade Crossing Crashes.....	75
Table 6. Direct, Indirect, and Total Effects of Variables on Injury Severity	76
Table 7. Illustrative Example Integrating Changes in Crash Rates and Injury Severity.....	77
Figure 1. Conceptual Framework for Estimating Associations of Crossing Controls with Injury Severity	78

Table 1. Selected Studies on Crash Frequencies/Rates

Authors/Year	Methodology	Crossing controls	Key findings
Millegan et al., 2009	Negative binomial	STOP signs vs. Crossbucks	STOP signs → 46.95% lower crash rates
Yan et al., 2010	Logistic regression	STOP signs vs. Crossbucks	STOP signs → less “did not stop” and “stopped on crossing” behaviors STOP signs → higher crash rate reductions at higher volume crossings
Yan et al., 2010	Hierarchical tree-based regression model	STOP signs vs. Crossbucks	STOP signs → varying crash rate reductions at different conditions
Raub, 2006	Descriptive statistics	STOP signs vs. Crossbucks	STOP signs → 60% higher crash rates by cross-sectional comparison STOP signs → 28% higher crash rates by before-and-after comparison
Mok and Savage, 2005	Negative binomial	Active vs. passive controls	Active controls → Lower crash frequencies
Raub, 2009	Descriptive statistics	Active (gates and flashing lights) vs. passive controls (STOP and Crossbuck signs)	Gates → 4.1 crashes per 10 MCV Flashing lights → 5.1 crashes per 10 MCV STOP signs → 37.4 crashes per 10 MCV
Park and Saccomanno, 2005	Poisson regression	Gates vs. passive signs Flashing lights vs. passive signs	Gates (vs. passive signs) → CMF = 0.34 Flashing lights (vs. passive signs) → CMF = 0.26
Saccomanno et al., 2007	Empirical Bayesian	Bells vs. flashing light crossings	The addition of bells → CMF = 0.45
Austin and Carson, 2002; Elvik and Vaa, 2004; Elvik et al., 2009	Poisson regression Negative binomial Meta-analysis	Flashing lights vs. passive signs Gates vs. flashing lights and bells Gates vs. passive signs	Flashing lights (vs. passive signs) → CMF = 0.49 Gates (vs. flashing lights and bells) → CMF = 0.55 Gates (vs. passive signs) → CMF = 0.33

NOTES: CMF = Crash Modification Factor

Table 2. Selected Studies on Crash Injury Severity and Behavioral Considerations

Authors/Year	Methodology	Crossing controls	Behavioral considerations	Key finding
Raub, 2006	Descriptive statistics	Yes	No	STOP signs → 12.4% crashes were fatal Gates → 31.8% crashes were fatal Flashing lights → 25% crashes were fatal
Eluru et al., 2012	Ordered logit model (Latent segmented)	Yes	Yes	Gates → Lowest injury severity Flashing lights (vs. STOP signs) → Higher injury severity Drove around or through the gates → Higher injury severity
Cooper et al., 2012	Descriptive statistics	Yes	Yes	Gates → 8.8% crashes were fatal Drove around gates → 20.6% crashes were fatal
Hao and Daniel, 2014	Descriptive statistics for the control and injury severity Ordered probit model for other factors	Yes	No	Active controls → 9.11% crashes were fatal Passive controls → 6.82% crashes were fatal Higher train/vehicle speed → higher injury severity
Hu et al., 2010	Generalized logit model	No	No	No findings on crossing controls Law enforcement cameras → lower injury severity
Russo and Savolainen, 2013	Ordered logit model	No	Yes	No findings on crossing controls Did not stop → Higher injury severity Higher train/vehicle speed → Higher injury severity Older drivers, females → Higher injury severity
Fan and Haile, 2014	Multinomial logit model	No	No	Higher train/vehicle speed → Higher injury severity
Zhao and Khattak, 2015	Multinomial logit model Ordered probit model Random parameter logit model	No	No	No findings on crossing controls Ordered probit model is less suitable for modeling injury severity than other two models Higher train/vehicle speed → Higher injury severity Older drivers, females → Higher injury severity

Table 3. Descriptive Statistics for Railroad Grade Crossing Crashes-Fra Database

Variables		N	Frequency	Mean or Percent	Std. Dev.	Min	Max
Crossing Control							
Active Crossing	Gates	15639	291	1.86%	0.135	0	1
	Flashing lights	15639	1145	7.32%	0.260	0	1
	Audible - Only	15639	45	0.29%	0.054	0	1
	Gates + Flashing lights	15639	1474	9.43%	0.292	0	1
	Gates + Audible	15639	175	1.12%	0.105	0	1
	Flashing lights + Audible	15639	1676	10.72%	0.309	0	1
Passive Crossing	Gates + Flashing lights + Audible	15639	4095	26.18%	0.440	0	1
	STOP sign	15639	2157	13.79%	0.345	0	1
	Crossbuck - Only	15639	3286	21.01%	0.407	0	1
	Other	15639	896	5.73%	0.232	0	1
	No controls	15639	399	2.55%	0.158	0	1
Advanced Warning Time							
Passive crossing (no advanced warning time)		15639	6692	42.79%	0.495	0	1
Warning time (>60 seconds)		15639	224	1.43%	0.119	0	1
Warning time (20~60 seconds)		15639	8382	53.60%	0.499	0	1
Warning time (<20 seconds)		15639	341	2.18%	0.146	0	1
Collision Speed							
Vehicle speed (Tidwell Jr and Humphreys) - before imputation		14838		9.607	13.200	0	100
Vehicle speed (Tidwell Jr and Humphreys) - after imputation		15639		9.724	13.083	0	100
Train speed (Tidwell Jr and Humphreys) - before imputation		15219		29.214	18.623	1	95
Train speed (Tidwell Jr and Humphreys) - after imputation		15639		29.131	18.580	1	95
Socio Demographics							
Driver gender (Male) - before imputation		14811	9996	67.49%	0.468	0	1
Driver gender (Male) - after imputation		15639	10540	67.40%	0.469	0	1
Driver age (Year) - before imputation		13002		40.955	18.719	10	99
Driver age (Year) - after imputation		15639		40.873	18.523	10	99
Vehicle type (Auto)		15639		0.701	0.458	0	1
Vehicle type (Pick-up)		15639	3741	23.92%	0.427	0	1
Vehicle type (Van)		15639	933	5.97%	0.237	0	1
Pre-crash Driver Behavior							
Drove around or thru the gate		15639	2308	14.76%	0.355	0	1
Stopped and then proceeded		15639	979	6.26%	0.242	0	1
Did not stop		15639	6455	41.28%	0.492	0	1
Stopped on crossing (in vehicle)		15639	1549	9.90%	0.299	0	1
Stopped on crossing (out of vehicle)		15639	2200	14.07%	0.348	0	1
Other		15639	2148	13.73%	0.344	0	1
Driver Injury Severity							
PDO		15639	10394	66.46%	0.472	0	1
Injury		15639	3991	25.52%	0.436	0	1
Fatal		15639	1254	8.02%	0.272	0	1

Table 4. Models for Driver Behavior Before Crashes on Railroad Grade Crossings

Pre-crash Behaviors	Drove around or thru the gate		Stopped and then proceeded		Did not stop		Stopped on crossing (in vehicle)		Other behaviors	
	$\beta_{(1)}$	Marg. Effect	$\beta_{(2)}$	Marg. Effect	$\beta_{(3)}$	Marg. Effect	$\beta_{(4)}$	Marg. Effect	$\beta_{(5)}$	Marg. Effect
Constant	-2.783 ***		-4.618 ***		-8.439 ***		-1.750 ***		-1.418 ***	
Crossing Controls										
Gates	0.532 **	0.037	0.674 *	0.031	0.127	-0.030	0.536 **	0.039	-0.639 **	-0.056
Flashing lights	-		3.753 ***	-0.018	8.147 ***	0.419	2.196 ***	-0.020	-0.091	-0.099
Audible warning devices	-		3.383 ***	-0.030	7.470 ***	0.362	2.279 ***	-0.015	0.207	-0.078
Gates + Flashing lights	1.271 ***	0.040	0.713 ***	-0.021	1.543 ***	0.077	0.262 **	-0.039	0.629 ***	-0.007
Gates + Audible warning devices	1.598 ***	0.065	0.368	-0.033	1.405 *	0.077	0.576 **	-0.017	0.768 ***	-0.011
Flashing lights + Audible warn.	-		3.682 ***	-0.014	8.136 ***	0.439	1.960 ***	-0.028	-0.045	-0.103
STOP sign	-		4.372 ***	0.008	8.105 ***	0.378	3.202 ***	0.050	0.710	-0.089
Crossbuck only	-		4.096 ***	-0.016	8.426 ***	0.450	3.566 ***	0.071	1.162	-0.087
Other control	-		4.910 ***	-0.014	8.893 ***	0.393	3.677 ***	0.010	1.751	-0.075
No control	-		3.290 ***	-0.029	7.451 ***	0.359	2.100 *	-0.026	0.901	-0.057
Advanced Warning Time										
Passive crossing	-		-0.706	-0.020	-0.049	0.059	-1.262	-0.072	-1.384	-0.070
Warning time (>60 seconds)	0.203	0.021	-0.630	-0.027	-0.035	0.013	-0.060	0.000	-0.149	-0.013
Warning time (<20 seconds)	-0.362	-0.127	2.462 ***	-0.027	5.172 ***	0.258	0.913 **	-0.081	1.713 ***	0.065
Collision Speed										
Train speed (Tidwell Jr and Humphreys)	-0.007 ***	-0.000	-0.027 ***	-0.001	-0.016 ***	-0.000	-0.011 ***	-0.000	-0.007 ***	-0.000
Vehicle speed (Tidwell Jr and Humphreys)	0.794 ***	0.019	0.746 ***	0.006	0.867 ***	0.035	-0.177 ***	-0.052	0.635 ***	0.017
Socio-Demographics										
Driver gender (Male)	0.182 **	0.027	-0.509 ***	-0.019	-0.190 **	0.010	-0.418 ***	-0.024	-0.258 ***	-0.014
Driver age (Years)	0.017 ***	-0.000	0.030 ***	0.000	0.022 ***	-0.000	0.034 ***	0.002	0.014 ***	-0.000
Vehicle type (Pick-up)	0.313 ***	-0.007	0.736 ***	0.015	0.546 ***	0.003	0.460 ***	0.012	0.455 ***	0.011
Vehicle type (Van)	0.245	-0.003	0.318	-0.003	0.443 **	0.012	0.416 ***	0.017	0.279 *	0.003
SUMMARY STATISTICS										
Number of Observations	15639									
Pseudo-R ²	0.462									
Log Likelihood at 0	-25001.7									
Log Likelihood at β	-13451.2									
Prob > χ^2	0.000									

NOTES: Pseudo-R²= 1 – (Log Likelihood at β /Log Likelihood at 0).

Marginal effects are the changes in the dependent variable with a unit change in the independent variable.

Base level for pre-crash behavior is "stopped on crossing (out of vehicle)".

Base level for crossing control is fully protected crossing, i.e., "gates + flashing lights + audible warnings".

Base level for warning time is "recommended 20 second minimum warning time".

Base level for vehicle type is "automobile".

Other behaviors include suicide or attempting suicide, going thru/around temporary construction barriers, and shoving onto the track, etc.

*** = significant at a 99% confidence level; ** = significant at a 95% confidence level; * = significant at a 90% confidence level.

Table 5. Ordered Logit Regression Model for Injury Severity in Railroad Grade Crossing Crashes

Independent Variables	β		P-value	Marginal Effect		
				PDO	Injury	Fatal
Pre-crash Behavior						
Drove around or thru the gate	3.691	***	0.000	-0.400	0.297	0.103
Stopped and then proceeded	3.357	***	0.000	-0.332	0.255	0.077
Did not stop	3.669	***	0.000	-0.395	0.294	0.101
Stopped on crossing (in vehicle)	2.839	***	0.000	-0.237	0.189	0.048
Other behavior	3.420	***	0.000	-0.344	0.263	0.081
Crossing Control						
Gates	-0.535	*	0.051	0.088	-0.058	-0.030
Flashing lights	-0.059		0.552	0.010	-0.006	-0.004
Audible warning devices	-0.097		0.782	0.017	-0.010	-0.006
Gates + Flashing lights	0.136	*	0.062	-0.024	0.015	0.010
Gates + Audible warning devices	-0.137		0.453	0.024	-0.015	-0.009
Flashing lights + Audible warn.	0.156	*	0.083	-0.028	0.017	0.011
STOP sign	-0.675		0.121	0.110	-0.070	-0.040
Crossbuck only	-0.740	*	0.089	0.122	-0.077	-0.045
Other control	-0.642		0.134	0.105	-0.069	-0.036
No control	-0.819	*	0.070	0.130	-0.088	-0.042
Advanced Warning Time						
Passive crossing (no advanced warning time)	0.815	*	0.057	-0.141	0.083	0.058
Warning time (>60 seconds)	0.093		0.567	-0.016	0.010	0.007
Warning time (<20 seconds)	0.522	***	0.000	-0.094	0.052	0.042
Collision Speed						
Train speed (Tidwell Jr and Humphreys)	0.045	***	0.000	-0.008	0.005	0.003
Vehicle speed (Tidwell Jr and Humphreys)	0.021	***	0.000	-0.004	0.002	0.001
Socio Demographics						
Driver gender (Male)	-0.290	***	0.000	0.051	-0.031	-0.020
Driver age (Years)	0.012	***	0.000	-0.002	0.001	0.001
Vehicle type (Pick-up)	-0.031		0.490	0.005	-0.003	-0.002
Vehicle type (Van)	-0.053		0.500	0.009	-0.006	-0.004
Constant (1)	5.893					
Constant (2)	7.977					
SUMMARY STATISTICS						
Number of Observations				15639		
Pseudo-R ²				0.155		
Log Likelihood at 0				-12861.35		
Log Likelihood at β				-10864.80		
Likelihood Ratio χ^2				3993.09		
Prob > χ^2				0.000		

NOTE:

Marginal effects are the changes of probability from a lower level to a higher level with a unit change in the independent variable. Pseudo-R² = 1 - (Log Likelihood at β /Log Likelihood at 0)

Base level for pre-crash behavior is "stopped on crossing (out of vehicle)."

Base level for crossing control is fully protected crossing, i.e., "gates & flashing lights & audible warnings."

Base level for warning time is "recommended 20 second minimum warning time".

Base level for vehicle type is "automobile."

Other behaviors include suicide or attempting suicide, going thru/around temporary construction barriers, and shoving onto the track, etc.

*** = significant at a 99% confidence level; ** = significant at a 95% confidence level; * = significant at a 90% confidence level.

Table 6. Direct, Indirect, and Total Effects of Variables on Injury Severity

Independent variables	Direct effect-Injury β	Effect on pre-crash behaviors					Effect of behaviors on injury					Indirect effect on injury					Total effect-Injury
		$\beta_{(1)}$	$\beta_{(2)}$	$\beta_{(3)}$	$\beta_{(4)}$	$\beta_{(5)}$	γ_1	γ_2	γ_3	γ_4	γ_5	$\beta_{(1)}$ *	$\beta_{(2)}$ *	$\beta_{(3)}$ *	$\beta_{(4)}$ *	$\beta_{(5)}$ *	
Crossing Control																	
Gates		3.7%		3.9%	-5.6%	40.0%	33.2%	39.5%	23.7%	34.4%	1.5%	0.0%	0.0%	0.9%	-1.9%	0.5%	
Flashing lights			-1.8%	41.9%	-2.0%	40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	-0.6%	16.6%	-0.5%	0.0%	15.5%	
Audible warning devices			-3.0%	36.2%	-1.5%	40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	-1.0%	14.3%	-0.4%	0.0%	12.9%	
Gates + Flashing lights		4.0%	-2.1%	7.7%	-3.9%	-0.7%	40.0%	33.2%	39.5%	23.7%	34.4%	1.6%	-0.7%	3.1%	-0.9%	-0.3%	2.8%
Gates + Audible warning devices		6.5%			-1.7%	-1.1%	40.0%	33.2%	39.5%	23.7%	34.4%	2.6%	0.0%	0.0%	-0.4%	-0.4%	1.8%
Flashing lights + Audible warn.			-1.4%	43.9%	-2.8%		40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	-0.4%	17.4%	-0.7%	0.0%	16.2%
STOP sign			0.8%	37.8%	5.0%		40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	0.3%	14.9%	1.2%	0.0%	16.4%
Crossbuck only			-1.6%	45.0%	7.1%		40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	-0.5%	17.8%	1.7%	0.0%	18.9%
Other control			-1.4%	39.3%	1.0%		40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	-0.5%	15.5%	0.2%	0.0%	15.3%
No control			-2.9%	35.9%			40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	-1.0%	14.2%	0.0%	0.0%	13.2%
Gates + Flashing L.+ Audible - Base																0.0%	
Advanced Warning Time																	
Passive crossing							40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Warning time (>60 seconds)							40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Warning time (<20 seconds)	9.4%		-2.7%	25.8%	-8.1%	6.5%	40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	-0.9%	10.2%	-1.9%	2.2%	19.0%
Warning time (20~60 seconds) -																	0.0%
Collision Speed																	
Train speed (Tidwell Jr and	0.8%	0.0%	-0.1%	0.0%	0.0%	0.0%	40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.8%
Vehicle speed (Tidwell Jr and	0.4%	1.9%	0.6%	3.5%	-5.2%	1.7%	40.0%	33.2%	39.5%	23.7%	34.4%	0.8%	0.2%	1.4%	-1.2%	0.6%	2.1%
Socio-Demographics																	
Driver Gender (Male)	-5.1%	2.7%	-1.9%	1.0%	-2.4%	-1.4%	40.0%	33.2%	39.5%	23.7%	34.4%	1.1%	-0.6%	0.4%	-0.6%	-0.5%	-5.3%
Driver age (Year)	0.2%	0.0%	0.0%	0.0%	0.2%	0.0%	40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%
Vehicle type (Pick-up)		-0.7%	1.5%	0.3%	1.2%	1.1%	40.0%	33.2%	39.5%	23.7%	34.4%	-0.3%	0.5%	0.1%	0.3%	0.4%	1.0%
Vehicle type (Van)				1.2%	1.7%		40.0%	33.2%	39.5%	23.7%	34.4%	0.0%	0.0%	0.5%	0.4%	0.0%	0.9%
Vehicle type (Auto) - Base																	0.0%

Note: Pre-crash behaviors - 1: Drove around or thru the gate; 2: Stopped and then proceeded; 3: Did not stop; 4: Stopped on crossing (in vehicle);

Table 7. Illustrative Example Integrating Changes in Crash Rates and Injury Severity

Crossing controls	Flashing lights + Bells	→	Gates + Flashing lights + Bells
<i>CMF</i> (Elvik and Vaa 2004)	0.55		
<i>CMFSEV</i>	0.162		
Steps for crash injury distribution modification	Number of crashes before implementation of gates	Number of crashes reduced after implementation of gates (not-modified for injury severity)	Number crashes reduced with modified injury severity distribution
Number of Crashes	100	55	55
P_{Inj}	36.87%	36.87%	30.90%
Number of Injury Crashes	36.87	20.28	16.99
Number of PDO Crashes	63.13	34.72	38.01
<i>Expected Number of Injury Crashes-After</i>	$(CMF \times N) \times P_{Inj} \times (1 - CMFSEV)$		
<i>Expected Number of PDO Crashes-After</i>	$(CMF \times N) - (CMF \times N) \times P_{Inj} \times (1 - CMFSEV)$		

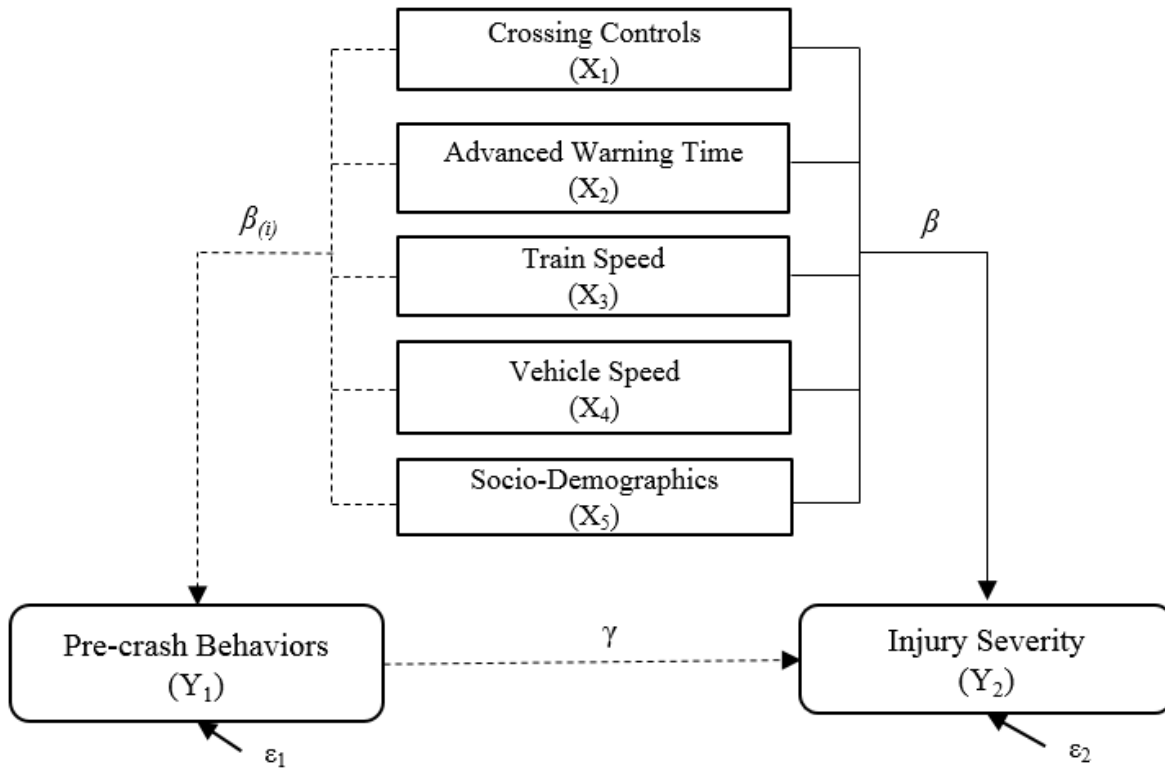


Figure 1. Conceptual Framework for Estimating Associations of Crossing Controls with Injury Severity

PUBLICATIONS & PRESENTATIONS GENERATED BY THIS NURAIL PROJECT

1. Wang, X., Liu, J., Khattak, A. J., & Clarke, D. (2016). Non-crossing rail-trespassing crashes in the past decade: A spatial approach to analyzing injury severity. *Safety Science*, 82, 44-55. (TRB paper # 15-0955, Presented at the Transportation Research Board, National Academies, Washington, D.C., 2015.)
2. Zhang, M., A. Khattak, J. Liu, & D. Clarke, A comparative study of rail-pedestrian/bicyclist trespassing crash injury severity at highway-rail grade crossings and non-crossings. Presented at 2015 Road Safety & Simulation International Conference, Orlando, FL, 2015.
3. Liu, J., Khattak, A. J., Richards, S. H., & Nambisan, S. (2015). What are the differences in driver injury outcomes at highway-rail grade crossings? Untangling the role of pre-crash behaviors. *Accident Analysis & Prevention*, 85, 157-169. (TRB paper # 15-0959, Presented at the Transportation Research Board, National Academies, Washington, D.C., 2015.)
4. Liu J., A. Khattak & S. Richards, Examining motorist gate-violation behaviors at highway-rail grade crossings using geo-spatial modeling integrated with path analysis: a screening tool for safety improvement zones. Presented at 2015 Road Safety & Simulation International Conference, Orlando, FL, 2015.

Primary Contacts

Principal Investigator

Asad J. Khattak

Beaman Professor & Transportation Program Coordinator

Department of Civil and Environmental Engineering

The University of Tennessee, Knoxville

Email: akhattak@utk.edu

Other Faculty and Students Involved

Jun Liu, Ph.D.

Research Associate

Department of Civil & Environmental
Engineering

The University of Tennessee

Email: jliu34@utk.edu

Meng Zhang

Research Assistant

Department of Civil and Environmental
Engineering

University of Tennessee, Knoxville

Email: mzhang27@vols.utk.edu

Xin Wang, Ph.D.

Senior Modeling Systems Analyst

Virginia Department of Transportation

Email: xwang757@gmail.com

NURail Center

217-244-4999

nurail@illinois.edu

<http://www.nurailcenter.org/>

