Investigating the impacts of crash prediction models on quantifying safety effectiveness of Adaptive Signal Control Systems

Weimin Jin  
*Clemson University*, weiminj@g.clemson.edu

Mashrur Chowdhury  
*Clemson University*, mac@clemson.edu

Sakib Mahmud Khan  
*Clemson University*, sakibk@clemson.edu

Patrick Gerard  
*Clemson University*, pgerard@clemson.edu

Follow this and additional works at: [https://tigerprints.clemson.edu/civileng_pubs](https://tigerprints.clemson.edu/civileng_pubs)

Part of the Civil and Environmental Engineering Commons, and the Statistics and Probability Commons

**Recommended Citation**


This Article is brought to you for free and open access by the Glenn Department of Civil Engineering at TigerPrints. It has been accepted for inclusion in Publications by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
Investigating Hierarchical Effects of Adaptive Signal Control System on Crash Severity

using Random-parameter Ordered Regression Models Incorporating Observed Heterogeneity

Weimin Jin\textsuperscript{1*}, Mashrur Chowdhury\textsuperscript{2}, M Sabbir Salek\textsuperscript{3}, Sakib Mahmud Khan\textsuperscript{4}, and Patrick Gerard\textsuperscript{5}

\textsuperscript{1}Glenn Department of Civil Engineering, Clemson University, Clemson, SC, 29634, USA; email: weiminj@g.clemson.edu

\textsuperscript{2}Ph.D., Professor, Glenn Department of Civil Engineering, Clemson University, Clemson, SC, 29634, USA; email: mac@clemson.edu

\textsuperscript{3}Glenn Department of Civil Engineering, Clemson University, Clemson, SC, 29634, USA; email: msalek@clemson.edu

\textsuperscript{4}Ph.D., Postdoctoral Scholar, Institute of Transportation Studies, U.C. Berkeley 94720, USA; email: sakibk@berkeley.edu

\textsuperscript{5}Ph.D., Professor, School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC, 29634, USA; email: pgerard@clemson.edu

*Corresponding author.
ABSTRACT

By handling conflicting traffic movements and establishing dynamic coordination between intersections in real-time, the Adaptive Signal Control System (ASCS) can potentially improve the operation and safety of signalized intersections on a corridor. This study identifies the hierarchical effects of ASCS on the crash severity by exploring the heterogeneous effect of ASCS on the crash severity. Four different random-parameter ordered regression models (two ordered probit models, and two ordered logit models) are developed and compared. The analysis reveals that the random-parameter ordered probit and logit models (ROP and ROL) with observed heterogeneity perform better than the random-parameter ordered probit and logit models (RP and RL) without observed heterogeneity in terms of the Akaike information criteria and the goodness of fit of the model. The ROP model performs better than the ROL model in terms of classification model performance measures. The ROP model enables parameters (i.e., the coefficients of the explanatory variables) to vary as a function of explanatory variables as well as across observations, thus accounting for both observed (captured by available explanatory variables) and unobserved (not captured by available explanatory variables) heterogeneity. The analysis reveals that the presence of ASCS is associated with lower crash severity. In this study, observed heterogeneity of ASCS effects on the crash severity is captured by variables related to the intersection and corridor features. Other contributing factors besides ASCS, such as annual average daily traffic, speed limit, lighting, peak period, crash type (rear-end, angle), and pedestrian involvements, are also associated with the probability of crash severity. Unobserved heterogeneity of the effect of angle crash type on the crash severity is found to exist across the observations. The findings of this research have practical implications.
for establishing ASCS implementation guidelines in lowering the probability of higher crash severity.

**Keywords**: Adaptive Signal Control System, Safety, Crash Severity, Ordered Regression Model, Unobserved and Observed Heterogeneity

1 Introduction

According to the American Association of State Highway and Transportation Officials (AASHTO) Strategic Highway Safety Plan (Antonucci et al., 2004), intersection-related crashes make up about 23 percent of all fatal crashes. Improving safety at intersections has become one of 22 key areas of focus in this plan. Through traffic control and operational improvement strategies, this plan aims at lowering crash severity and decreasing crash frequency at signalized intersections. Transportation agencies have been seeking new insights and approaches to improve safety at signalized intersections.

Adaptive Signal Control System (ASCS) is an advanced signal control system deployed at intersections. Although each type of ASCS is unique, all types of ASCS use a similar framework to meet fluctuating real-time traffic demand. ASCS usually includes algorithms that optimize and update parameters (i.e., cycle length, offsets, phase splits, and phase sequence) of traffic signals in real-time (Gartner et al., 2002; Lowrie, 1990). ASCS requires detectors such as loop detectors and video detectors, and a communication network enabling the central and/or local controllers to communicate with such detectors. Prior to the deployment of ASCS, conventional signal systems (i.e., pre-timed or actuated signal control) based on Time of Day (TOD) are typically
used at signalized intersections. These TOD based signal systems with pre-set signal plans usually updated every two to three years cannot handle highly variable traffic demand. However, ASCS can better accommodate fluctuating traffic demand or extreme traffic conditions. Previous studies suggest that ASCS has produced significant operational improvements on corridors and at intersections (Eghtedari, 2006; Elkins and Niehus, 2012; Fontaine et al., 2015; Kergaye et al., 2009; Khattak, 2016; Khattak et al., 2019b; So et al., 2014). By handling conflicting traffic movements and establishing dynamic coordination between intersections in real-time, ASCS can smooth traffic flow and reduce traffic congestion, thus potentially yielding safety benefits for signalized intersections (Dutta et al., 2010; Fink et al., 2016; Jin et al., 2019; Khattak et al., 2018; Ma et al., 2016). These studies related to the safety effects of ASCS have emphasized the impact of ASCS on the crash frequency but not its effect on the outcomes of crash severity (i.e., non-injury, possible injury, non-incapacitating injury, incapacitating injury, or fatal).

Only a few previous studies have examined ASCS effects on crash severity outcomes. Out of those studies, two studies (Dutta et al., 2010; Fink et al., 2016) assume that the ASCS effects on crash severity are fixed in all signalized intersections for a certain type of ASCS. One recent study (Khattak et al., 2019a) has identified the disparity of ASCS effects on crash severity with two different types of ASCS and the difference of ASCS effects on crash severity between two states- Pennsylvania and Virginia. However, previous studies typically assume that ASCS effects on crash severity outcomes do not vary across observations (i.e., crashes).

Crash severity outcomes could be associated with a series of variables related to the corridor, intersection, and crash features. A multilevel structure (i.e., hierarchical structure) inherent in the
crash data will be overlooked if all the variables are viewed at one level. Hierarchical modeling is used to represent the multilevel-structure of the crash data. The hierarchical effect on crash estimation has been explored in previous studies (Chen and Persaud, 2014; Huang et al., 2008; Khazraee et al., 2018; Krishnan et al., 2013; Xie et al., 2013). In this study, ASCS is usually deployed at several signalized intersections on corridors; thus, the hierarchical structure exists inherently in the crash data. The crash data structure can be viewed as a two-level hierarchy with Level 1 being an individual crash, and Level 2 being the intersection and corridor (i.e., one individual crash can be associated with one specific intersection and corridor). The ASCS effect on the crash severity that exists in the hierarchical structure can be estimated by implementing hierarchical models. A random-parameter ordered regression model integrating observed heterogeneity (also known as a hierarchical model) allows the ASCS parameter to vary both as a function of explanatory variables related to the intersection and corridor features, and across crashes. This kind of ASCS effect that exists in the hierarchical structure is referred to as “Hierarchical Effects of ASCS on the Crash Severity” in this paper.

In previous studies related to crash severity outcome modeling, random-parameter ordered regression models have been implemented, but they have not been integrated with observed heterogeneity. This study contributes to the literature related to ASCS effects on crash severity by developing random-parameter ordered regression models with observed heterogeneity. Our model is capable of accounting for both observed (i.e., changes in the effect of predictors across the observations that are known and could be captured by available explanatory variables) and unobserved (i.e., changes in the effect of predictors across the observations that are unknown and could not be captured by available explanatory variables) heterogeneity. The objective of this
paper is to determine the hierarchical effects of ASCS on the crash severity. Through accounting for the observed heterogeneity in random-parameter ordered regression models using crash data from six ASCS corridors with 65 signalized intersections, the hierarchical effects of ASCS on the crash severity are identified. The identification of the hierarchical effects of ASCS on the crash severity provides several practical implications on ASCS implementations at the standpoint of safety. The same ASCS type has been deployed at the intersections considered in this study. So, the algorithms of this particular type of ASCS are the same for all the intersections considered in this study. ASCS optimizes the cycle length, phase splits, and offsets in real-time based on current traffic conditions to minimize overall traffic delays at the intersections while guaranteeing a reasonable coordination between the intersections on a corridor. Our study focuses on investigating the effects of a particular ASCS type (which this paper refers to as “ASCS”) on the cash severity without considering the variations between multiple ASCS types as it is not within the scope of this study.

The remainder of the paper is organized as follows: literature review, research method, data description, results and analysis, and conclusions. The literature review section entails the safety effects of the ASCS on the crash frequency, the crash severity outcomes associated with ASCS, and models developed for the crash severity outcomes. A discussion of the ordered regression models and model implementation are included in the research method section. The data description details the description of variables contributing to the crash severity. The model comparison and estimation results on the association between contributing factors and crash severity are detailed in the results and analysis section. Finally, the summary findings and practical implications are detailed in the conclusions section.
2 Literature review

2.1 Safety effects of adaptive signal control systems on crash frequency

As ASCS uses real-time traffic data to maximize vehicle progression and reduce vehicle stops, it is expected to improve safety by reducing rear-end crash probability compared to pre-timed or actuated signal control systems. Based on multiple evaluations across the nation, ASCS has been found to improve intersection safety, with a few exceptions. In one study conducted in Virginia, Fontaine et al. (Fontaine et al., 2015) note that a 17% reduction in the total number of crashes is associated with ASCS using the Empirical Bayes (EB) method. In Pennsylvania, Khattak (Khattak, 2016) notes that 34% and 45% reductions in the total number of crashes and fatal and injury crashes are associated with ASCS, respectively. Based on the analysis of ten years of crash data (1999 - 2008) in Michigan, ASCS is found to reduce the total crash rate by 6% (Dutta et al., 2010). Fink et al. (Fink et al., 2016) have examined the safety effects of signalized intersections with Sydney Coordinated Adaptive Traffic System (SCATS), an ASCS type, in Michigan. Fink et al. (Fink et al., 2016) have implemented a cross-sectional method by using 498 signalized intersections and found that a 19.3% reduction in angle crashes is associated with SCATS.

2.2 Safety effects of adaptive signal control systems on crash severity outcomes

Evaluation of the impact of ASCS deployment on crash severity outcomes is predominantly absent from the literature. Only a few studies related to crash severity effects of ASCS are identified. Dutta et al. (Dutta et al., 2010) have used before period (1999 - 2001) and after period (2003 - 2008) crash data from one corridor with SCATS and another with the pre-timed signal. They perform a t-test analysis and find that a definite change in severity from incapacitating
injury and non-incapacitating injury to possible injury. But, the t-test fails to prove the superiority of SCATS over the pre-timed signal control system in lowering the crash severity at a 0.05 significance level.

In their analysis, Fink et al. (Fink et al., 2016) have used data from 498 signalized intersections in Michigan. They have found that a statistically significant reduction in non-incapacitating injuries is associated with SCATS deployment. However, they have not found that a statistically significant reduction in fatal and incapacitating injuries is associated with SCATS.

Similarly, in the examination of the effect of ASCS on the crash severity, Khattak et al. (Khattak et al., 2019a) have identified the disparity between two different types of ASCS and between two states- Pennsylvania and Virginia. They have found that both types of ASCS are associated with lower crash severity, and ASCS implemented in these two states are also associated with lower crash severity.

2.3 **Modeling of crash severity outcomes**

Of the various modeling approaches developed for modeling the crash severity outcomes, the most widely used models are ordered regression models (i.e., ordered probit and logit models), which account for the ordinal nature reflected in crash severity levels.

Khattak and Tung (Khattak and Tung, 2015) have used an ordered probit model to quantify the impact of various contributing factors (e.g., train speed, weather condition, area types, etc.) on three different severity levels (i.e., no injury, injury, and fatal) of pedestrian injuries. Zhao and Khattak (Zhao and Khattak, 2015) have used an ordered probit model to investigate contributing
factors (e.g., the involvement of freight trains, older drivers, high train speeds, high vehicle speeds, and female drivers) correlated with crash severity levels of motorists in the case of the train-vehicle crashes. Eluru et al. (Eluru et al., 2012) have used a latent segmentation based ordered logit model to identify the contributing factors (e.g., time of the collision, motorist age, snow and/or rain condition, the presence of a vehicle struck by a train, and driver operation in the pre-crash) that influence injury severity of vehicle drivers in train-motor crashes at highway-railway grade crossings. In some studies, researchers have not considered crash severity levels as ordinal outcomes and therefore used multinomial logit models (Fan et al., 2015; Fink et al., 2016). Based on a comparative analysis of the three most common models (i.e., ordered probit, mixed logit model, and multinomial logit), Ye, and Lord (Ye and Lord, 2014) have concluded that each can be the best choice based on the sample size requirement. Among the three models, the mixed logit model requires the largest sample size, and the multinomial logit model requires the second-largest sample size. In contrast, the ordered probit model requires the smallest sample size. The minimum sample size recommended by (Ye and Lord, 2014) for the mixed logit model, multinomial logit model, and ordered probit models is 5000, 2000, and 1000, respectively.

3 Research method

Ordered regression models (i.e., ordered probit and logit models) are implemented to account for the ordinal nature (e.g., ranging from non-injury to possible injury, to non-incapacitating injury, to incapacitating injury, to fatal) of crash severity. The ordered regression models have been widely used to consider the ordinal nature of crash data mainly. However, an underlying assumption of ordered regression models is that the estimated parameters across crash severity
levels are constant. This assumption is referred to as “the proportional odds or parallel regression” assumption. In this study, the authors initially fit ordered regression models and test this possible assumption by using the Brant test (Brant, 1990). It is found that the variable associated with the presence of ASCS does not violate the assumption. However, ordered regression models cannot capture unobserved heterogeneity across observations. Thus, the models may result in incorrect estimates (Washington et al., 2020). The random-parameter ordered regression model enables the parameters to vary across observations and has been explored by previous studies (Dabbour et al., 2017; Jalayer et al., 2018; Khattak et al., 2019a). However, in previous studies related to crash severity outcome modeling, random-parameter ordered regression models have not been integrated with observed heterogeneity.

3.1 Random-parameter ordered regression model

The random-parameter ordered regression models (Greene, 2003) are implemented in this paper. The ordered regression model is used to study the following latent process:

\[ y_i^* = X_i \beta_i + \varepsilon_i, \quad i = 1, \ldots, n \]  
\[ \beta_i \sim g(\beta_i | \theta) \]

where \( y_i^* \) is a latent variable for the observation (i.e., crash) \( i \); \( X_i \) is a vector of the explanatory variables; \( \beta_i \) is a vector of the coefficients; \( \varepsilon_i \) is the error term; \( n \) is the total number of observations.
In Eq. (1) and (2), $\beta_i$ is allowed to be different for each observation $i$ rather than fixed for all observations. The distribution $g(\beta_i | \theta)$ is specified to enable $\beta_i$ vary across observations, where $\theta$ is a vector of the mean and variance of the random distribution.

$\beta_i$ can be written as $\beta_i = \beta + L \omega_i$, where $\beta$ is the vector of the mean of the coefficients. Each coefficient $\beta_{ki}$ can be expressed as $\beta_{ki} = \beta_k + \sigma_k \omega_{ki}$. $\beta_k$ is the $k^{th}$ element in $\beta_i$. $\omega_i$ is a vector of random variables that follow a random distribution. $\omega_{ki}$ is the $k^{th}$ element in $\omega_i$. $\omega_{ki}$ has a specific random distribution such as normal distribution and uniform distribution. $L$ is a diagonal matrix of the standard deviations of the coefficients, $\sigma_k$. The unobserved heterogeneity is represented by $\sigma_k$. $\beta_{ki}$ has a specific random distribution such as normal distribution and uniform distribution. For example, $\beta_{ki}$ follows a normal distribution with a mean of $\beta_k$ and a variance of $\sigma_k^2$ when $\omega_i \sim \mathcal{N}(0,1)$.

The probability of the crash severity level $j$ for the crash $i$, can be calculated as:

$$p(y_i = j) = P(\mu_{j-1} < y_i^* \leq \mu_j) = F(\mu_j - \mathbf{X}_i \beta_i) - F(\mu_{j-1} - \mathbf{X}_i \beta_i)$$  \hspace{1cm} (3)

where $y_i$ is an ordered categorical variable, $\mu_j$ is the $j^{th}$ threshold in the model, and $F$ is the standard normal Cumulative Distribution Function (CDF) for the ordered probit model or logistic CDF for the ordered logit model.
“KABCO” injury scale (K- fatal; A- incapacitating injury; B- non-incapacitating injury; C- possible injury; O- no injury) is usually used for classifying injuries. The crash severity levels provided by the South Carolina Department of Transportation (SCDOT) crash database include five categories: non-injury, possible injury, non-incapacitating injury, incapacitating injury, and fatal, which correlates to the KABCO injury scale. Since relatively fewer crashes (1.08% out of observations) are reported for incapacitating injury (i.e., A) and fatal (i.e., K) categories in this study, the two categories are combined with the non-incapacitating injury (i.e., B) category. The KAB represents a sum of K, A, and B injury crashes, which is typically evaluated in safety studies (National Research Council (US), 2010). The crash severity levels are coded as three categories in this paper: (1) non-injury (i.e., O), (2) possible injury (i.e., C), and (3) non-incapacitating injury, incapacitating injury and fatal combined (i.e., KAB).

The authors use the following model to express the response variable, \( y_i \), which is composed of three crash severity levels. It is expressed as

\[
y_i = 0 \quad \text{if} \quad y_i^* \leq \mu_0 \\
y_i = 1 \quad \text{if} \quad \mu_0 < y_i^* \leq \mu_1 \\
y_i = 2 \quad \text{if} \quad y_i^* > \mu_1
\]

where, \( y_i = 0 \) indicates that the crash is O (KABCO scale); \( y_i = 1 \) indicates that the crash is C (KABCO scale); \( y_i = 2 \) indicates that the crash is K, A or B (KABCO scale); \( \mu_0 \) and \( \mu_1 \) represent different thresholds for three crash severity levels; \( \mu_0 \) is 0 here for non-injury. Here, only one threshold (i.e., \( \mu_1 \)) needs to be estimated.
3.2 Random-parameter ordered regression model with observed heterogeneity

The random-parameter ordered regression model with observed heterogeneity can accommodate observed heterogeneity by allowing parameter variations to be captured by available explanatory variables. This model is also referred to as a hierarchical model (Greene and Hensher, 2010; Sarrias, 2016).

The hierarchical model is used to represent the multilevel-structure of the crash data. In the study corridors considered in this research, ASCS is usually deployed at several signalized intersections along corridors, thus the hierarchical structure exists inherently in the crash data. As shown in Figure 1, each crash can be associated with one specific intersection that belongs to one specific corridor. The crash data structure can be viewed as a two-level hierarchy with Level 1 being an individual crash, and Level 2 being the intersection and corridor that include the individual crash. The intersection and corridor are considered at the same level in the model. Also, the two-level hierarchy model is considered to avoid excessive complexity of the model development. The ASCS effect on the crash severity that exists in the hierarchical structure can be estimated by implementing the hierarchical model.

![Figure 1 Hierarchical Structure of Crash Data](image)

Notation:
- \( m_e \): the number of crashes associated with Intersection \( e \) and Corridor 1
- \( m_f \): the number of crashes associated with Intersection \( f \) and Corridor 2
- \( s \): the number of corridors
- \( d_1 \): the number of intersections on Corridor 1
- \( d_2 \): the number of intersections on Corridor 2
- \( d_s \): the number of intersections on Corridor \( s \)
In the crash level (Level 1 in the hierarchical model), \( y^*_i \) is used to study the latent process as shown below:

\[
y^*_i = X_i \beta_i + \varepsilon_i, \quad i = 1, ..., n
\]  

(7)

where, \( X_i \) is a vector of the crash-level explanatory variables for the \( i^{th} \) observation; \( \beta_i \) is a vector of the coefficients; \( \varepsilon_i \) is the error term; \( n \) is the total number of observations.

In the intersection/corridor level (Level 2 in the hierarchical model), \( \beta_i \) is specified by Eq. (8).

The specification of Eq. (8) allows the coefficients to vary with different intersections and corridors.

\[
\beta_i = \beta + \Pi s_i + L \omega_i
\]

(8)

where, \( \beta \) is a vector of the mean of coefficients; \( \omega_i \) is a vector of random variables that follow random distributions; \( L \) is a diagonal matrix of the standard deviations of the coefficients; \( s_i \) is a vector of intersection/corridor-level explanatory variables; \( \Pi \) is a matrix of coefficients of the intersection and corridor related variables. Then, the expectation of coefficients is

\[
E(\beta_i) = \beta + \Pi s_i.
\]

The expectation of coefficients is a function of the intersection/corridor-level variables, \( s_i \).

More specifically, in Eq. (8), two components, \( \Pi s_i \) and \( L \omega_i \) are introduced to allow the coefficients to vary with different levels. First, \( \Pi s_i \) is a linear function depending on the intersection/corridor related variables, \( s_i \). The primary purpose of using \( \Pi s_i \) is to capture the
observed heterogeneity across different intersections and corridors. It is expected that the varying intersections and corridor features (e.g., number of legs at an intersection, number of through/left/right lanes at an intersection, speed limit difference between a major street and a minor street at an intersection, and signalized intersection distance on a corridor) may lead to different crash severity. This specification of $\Pi_{is}$ is the same as reported in (Huang et al., 2008).

Second, $L_{oi}$ represents random effects, which capture both the intersection/corridor-level variability (Level 2) and the crash-level variability (Level 1). The primary purpose of using $L_{oi}$ is to capture the unobserved heterogeneity in the crash data. These random effects are assumed not only to vary across various intersection/corridors (Level 2) but also to vary for the crashes (Level 1) within the same intersection/corridor. This specification of random effects is different from the (Huang et al., 2008) study, in which it is assumed that the random effects only vary across different crashes (Level 2 in the model of the (Huang et al., 2008) study), whereas they are kept the same for all the driver-vehicle units (Level 1 in the model of the (Huang et al., 2008) study) within the same crash.

The conditional mean of the parameters (Sarrias, 2016), conditional on the specific data of each crash is estimated by Simulated Maximum likelihood (SML) procedure, which is expressed as:

$$\hat{E}(\beta_i \mid \text{data}_i) = \sum_{r=1}^{R} \left( \frac{\hat{P}(y_i \mid X_i, \beta_{ir})}{\sum_{r=1}^{R} \hat{P}(y_i \mid X_i, \beta_{ir})} \right) \hat{\beta}_{ir}$$

(9)

where, $\hat{\beta}_{ir} = \hat{\beta} + \hat{\Pi}_{is} + \hat{L}_{oi}$; $\hat{P}(y_i \mid X_i, \beta_{ir})$ is the estimated simulated probability for a crash $i$ evaluated at the $r^{th}$ draw of $\beta_i$; data$_i$ represents the explanatory variables; $R$ is the total number
of draws in the SML procedure. The random draws are generated by a Halton random number
generator with a standard uniform distribution, $U(0,1)$. Detail Halton draws procedure can be
found in (Sarrias, 2016).

3.3 Model implementation and estimation

The random-parameter ordered regression models are estimated through the SML procedure
described in (Sarrias, 2016). R software is used to perform SML procedure to obtain model estimation results using the “Rchoice” library (Sarrias, 2016). 300 Halton draws are used in the SML procedure, which is in line with a previous study (Khattak et al., 2019a). “Rchoice” library provides some options of distributions for random parameters, such as normal distribution and uniform distribution. Different distributions of random parameters are implemented and tested in the models. Eventually, the uniform distribution is used since it provides a better model fit. The signs of the coefficient of predictors are of particular interest. In the model estimation results in this paper, a positive sign of the coefficient of predictors is associated with higher crash severity (i.e., C and KAB), whereas a negative sign of the coefficient is associated with lower crash severity (i.e., O).

To evaluate the effect of the explanatory variables on the probability of crash severity, especially on the intermediate level (i.e., C), marginal effects for the three crash severity levels (i.e., O, C, and KAB) are computed. The marginal effect of explanatory variables indicates the change of the probability of crash severity associated with a one-unit change in the continuous variables or change from “0” to “1” in the indicator variables. It should also be noted that marginal effects are estimated at the sample mean of the explanatory variables using the expectation of
parameters when computed for random parameters. Marginal effects for the three crash severity levels are computed (Greene, 2003; Washington et al., 2020) as follows:

\[
\frac{\partial p(y_i = 0)}{\partial x} = -\phi(-X\beta)\beta
\]

(10)

\[
\frac{\partial p(y_i = 1)}{\partial x} = \phi(-X\beta)\beta - \phi(\mu_i - X\beta)\beta
\]

(11)

\[
\frac{\partial p(y_i = 2)}{\partial x} = \phi(\mu_i - X\beta)\beta
\]

(12)

where, \( \phi \) is the standard normal Probability Density Function (PDF) for the ordered probit model or logistic PDF for the ordered logit model; \( x \) is the explanatory variable.

The Variance Inflation Factor (VIF) is used to check for potential Multi-Collinearity (MC). Commonly a VIF of 10 has been used by many researchers as a rule of thumb to indicate severe MC issues (O’brien, 2007). The best fit models are selected based on a comparison of the Akaike Information Criteria (AIC) (Burnham and Anderson, 2004), with the model with the lowest AIC value deemed the best fit model. Also, the likelihood ratio test (Washington et al., 2020) is used to select a model with better goodness of fit of the model. Three classification model performance metrics: accuracy, precision, and recall, which are widely used for evaluating a classification model, are used for evaluating the performance of the random-parameter ordered regression model (i.e., ordered probit or logit) with observed heterogeneity. A training dataset with 80% of the sample and a test dataset with 20% of the sample are obtained. The training dataset and test dataset are randomly sampled. During the sampling, both datasets are ensured to have similar percentages of data points by category (i.e., by crash severity outcomes). The sampling procedure is repeated 30 times (Rahman et al., 2019; Xie et al., 2019). For each time,
the model is developed using the training dataset, and then the model is evaluated using the test dataset. The accuracy, precision, and recall are computed for the test dataset using Eq. (13) to Eq. (15) (Sammut Claude and Webb, 2010; Ting, 2010).

\[
\text{Accuracy} = \frac{CT}{N} \quad (13)
\]

\[
\text{Precision} = \frac{TP_t}{(TP_t + FP_t)} \quad (14)
\]

\[
\text{Recall} = \frac{TP_t}{(TP_t + FN_t)} \quad (15)
\]

where, \(CT\) is the total number of correctly classified instances for all classes; \(N\) is the total number of instances for all classes; \(t\) is the class label (i.e., O, C, or KAB); \(TP_t\) is true positive for the class label \(t\); \(FP_t\) is false positive for the class label \(t\); \(FN_t\) is false negative for the class label \(t\).

Overall precision and recall are evaluated by computing the micro-average values of precision and recall (Van Asch, 2013), which are derived by Eq. (16) and Eq. (17).

\[
\text{Precision}_{\text{Micro-average}} = \frac{\sum_{t=1}^{T} TP_t}{(\sum_{t=1}^{T} TP_t + \sum_{t=1}^{T} FP_t)} \quad (16)
\]

\[
\text{Recall}_{\text{Micro-average}} = \frac{\sum_{t=1}^{T} TP_t}{(\sum_{t=1}^{T} TP_t + \sum_{t=1}^{T} FN_t)} \quad (17)
\]

where, \(T\) is the total number of classes, including three classes (i.e., O, C, and KAB) in this paper.

### 4 Data description

Initially, the authors obtained crash data from 13 corridors that have installed ASCS. Original crash data have before period and after period data. The authors only include corridors that have
at least a two-year after period crash data for this study. As listed in Table 1, there is a total of six corridors (one in Lexington, one in Greenville, two in Charleston, one in Pawleys Island, and one in Summerville) with a total of 65 intersections where ASCS has been deployed. In total, 6,570 crashes are analyzed in this study. Only one type of ASCS is investigated in this study. For this specific type of ASCS, there are three main components, including server, local traffic controller, and vehicle detection. The server is responsible for processing data and calculating updated signal timings. The local traffic controller is responsible for gathering detection data and executing the commands received from the server. The interaction between the server and the local traffic controller is performed every few seconds to ensure signal timings are always up-to-date. The primary objective of the algorithm of this type of ASCS is to minimize traffic delays at the intersections while ensuring reasonable progression bandwidth of the corridor. This type of ASCS optimizes the cycle length, splits, and offsets based on real-time traffic conditions. By handling conflicting traffic movements and establishing dynamic coordination between intersections in real-time, ASCS can smooth traffic flow and reduce traffic congestion, thus potentially yielding safety benefits for signalized intersections.

Table 1 Corridor Information

<table>
<thead>
<tr>
<th>Location</th>
<th>Corridor Name</th>
<th>Number of Signals</th>
<th>Installation Date</th>
<th>Number of Crashes at Intersections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greenville</td>
<td>Roper Mt Rd/Garlington Rd</td>
<td>5</td>
<td>November 2016</td>
<td>193</td>
</tr>
<tr>
<td>Charleston</td>
<td>SC 642</td>
<td>18</td>
<td>June 2015</td>
<td>931</td>
</tr>
</tbody>
</table>
According to SCDOT, the intersection-related crashes are those that occurred within 0.05 miles of the center of the intersection. Using the threshold of 0.05 miles, the intersection-related crashes are identified. Crash data are collected for all study corridors from 2011 to 2018. Crash data of six months after the installation of ASCS are removed from the analysis, which eliminates the effect of acclimation to ASCS of drivers. As shown in Table 2, around 79% of the crashes are O (i.e., no injury), and a small proportion (around 6%) of the crash is KAB (i.e., fatal, incapacitating injury, and non-incapacitating injury combined).

### Table 2 Frequency (and percentage) of Crash Severity

<table>
<thead>
<tr>
<th>Crash Severity</th>
<th>Frequency (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O*</td>
<td>5182 (78.9%)</td>
</tr>
<tr>
<td>C*</td>
<td>1005 (15.3%)</td>
</tr>
<tr>
<td>KAB*</td>
<td>383 (5.8%)</td>
</tr>
</tbody>
</table>
In order to properly analyze the crash data, the authors collect information from SCDOT regarding whether any other possible safety improvements, in addition to the ASCS, have been implemented at intersections. Flashing Yellow Arrow (FYA) was installed on some signalized intersections before or after the ASCS was installed. Although only 3.9% (256 out of 6,570) of the observations have FYA, FYA may affect the crash severity outcome. Therefore, the authors initially consider FYA as one of the explanatory variables of the model. A categorical variable is considered to distinguish the effects of different numbers of FYA at the intersections on the crash severity outcomes. It is found that the categorical variable is not significant and adding the categorical variable increases AIC of the model. Thus, the FYA variable is taken out of the model since it cannot provide useful information. Left-turn lanes were modified at one intersection in one of study corridors after the ASCS was installed. To exclude the effect of such improvements that may impact safety, the crashes that occurred at this intersection are not included in the analysis. An additional signal phase was added to one signal after the ASCS was installed; thus, the crashes located at this intersection are not included in the analysis as well.

Table 3 shows a summary of descriptive statistics of the geometric features of intersections and speed limit data. The difference in the geometric features of intersections and speed limit between before period and after period is very small.
### Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Before Period</th>
<th></th>
<th>After Period</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.*</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>Number of legs at intersections</td>
<td>3.82</td>
<td>0.38</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Number of through lane(s) on major roads</td>
<td>5.37</td>
<td>1.44</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Number of the exclusive right-turn lane(s) on major roads</td>
<td>1.20</td>
<td>0.80</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Number of the exclusive left-turn lane(s) on major roads</td>
<td>2.28</td>
<td>0.91</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Number of through lane(s) on minor roads</td>
<td>2.16</td>
<td>1.21</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Number of the exclusive right-turn lane(s) on minor roads</td>
<td>1.02</td>
<td>0.70</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Number of the exclusive left-turn lane(s) on minor roads</td>
<td>1.81</td>
<td>0.89</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Number of access points on major roads</td>
<td>3.03</td>
<td>1.75</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Number of access points on minor roads</td>
<td>2.38</td>
<td>1.92</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Speed limit on major roads (mph)</td>
<td>42.64</td>
<td>5.00</td>
<td>25</td>
<td>55</td>
</tr>
<tr>
<td>Speed limit on minor roads (mph)</td>
<td>32.15</td>
<td>4.89</td>
<td>25</td>
<td>50</td>
</tr>
</tbody>
</table>

* S.D.-Standard deviation

The following crash attributes are provided by SCDOT: collision time, crash severity, Annual Average Daily Traffic (AADT), light condition, road surface condition, crash type, weather condition, work zone, first harmful event, and probable cause. The purpose of the inclusion of the first harmful event is to determine the involvement of pedestrians. The purpose of the inclusion of probable cause is to identify the careless (distracted) or aggressive drivers (i.e., aggressive operation of the vehicle or at excessive speed). The light (dawn, daylight, dusk or dark), and weather conditions (rain or not) are accounted for because those attributes may impact...
crash severity. Peak periods for each corridor analyzed in this study are identified by analyzing hourly average travel time data provided by the Iteris ClearGuide system (Iteris, 2020). The peak periods only exist on weekdays for the study corridors, and we found that the hourly average travel time does not vary much over the 24 h during weekends on our study corridors. That is why only weekday peak periods are considered in this paper, as shown in Table 4. Crash data available for South Carolina and provided by the SCDOT do not map crashes to traffic signal status (green, yellow, or red) (SCDOT, 2020a). Consequently, each crash cannot be associated with a specific signal phase from the available data. Due to this limitation, signal related parameters, such as signal status (green, yellow or red) and green/yellow/red time, could not be introduced into the model.

Table 4 Peak Periods for the Study Corridors

<table>
<thead>
<tr>
<th>Location</th>
<th>Corridor Name</th>
<th>Peak Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greenville</td>
<td>Roper Mt Rd/ Garlington Rd</td>
<td>7:00 - 10:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16:00 - 19:00</td>
</tr>
<tr>
<td>Charleston</td>
<td>SC 642</td>
<td>6:00 - 9:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15:30 - 18:30</td>
</tr>
<tr>
<td>Charleston</td>
<td>US 52</td>
<td>7:00 - 10:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14:00 - 18:00</td>
</tr>
<tr>
<td>Lexington</td>
<td>N Lake Drive</td>
<td>7:00 - 10:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15:00 - 19:00</td>
</tr>
<tr>
<td>Pawleys Island</td>
<td>US 17</td>
<td>11:00 - 15:00</td>
</tr>
<tr>
<td>Summerville</td>
<td>US 17A</td>
<td>11:00 - 14:00</td>
</tr>
</tbody>
</table>
The attributes from the crash database are converted to the response and explanatory variables. The response variable includes three crash severity levels - O, C, and KAB. The explanatory variables include light condition, ASCS presence, FYA presence, peak period, crash type (rear-end or angle), weather condition, careless driving, aggressive driving, the presence of pedestrians, and AADT. Also, the authors have collected area type (urban or not) and speed limit data from the SCDOT website (SCDOT, 2020b), and corridor geometric features (i.e., the average distance between signalized intersections) from Google Earth. The descriptive statistics for the response and the significant explanatory variables for both before period and after period are shown in Table 5. A Pearson correlation test between AADT and the peak period is conducted, and it is found that there is no correlation between AADT and the peak period in our study. A high traffic volume may be associated with a higher crash severity. An AADT threshold of 30,000 is used to identify relatively high traffic volume in this study based on a previous study (Fink et al., 2016). A threshold of 10 mph speed difference between a major road and a minor road at an intersection is used to divide the observations into two groups (one group for which the speed limit difference between a major road and a minor road is equal to or greater than 10 mph; another group for which the speed limit difference between a major road and a minor road is less than 10 mph) because, based on our analysis, the median speed limit difference between a major road and a minor road in the sample is about 10 mph. In this study, all explanatory variables are tested in terms of MC. It is found that the maximum value of VIF is 2.37. Thus, the MC issue should not be of concern for the variables considered in this study. The authors initially include the interaction variables into the model to account for the interaction between
ASCS and angle crash and the interaction between ASCS and rear-end crash in the model. However, the interaction variables are not significant and adding these interaction variables increases the AIC of the model. Therefore, the interaction variables are later excluded from the model.

Table 5 Summary of Descriptive Statistics of Response Variables and Significant Explanatory Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Level/Data Type</th>
<th>Before Period Frequency (Percentage)</th>
<th>After Period Frequency (Percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crash_Severity</td>
<td>Crash severity outcome</td>
<td>0 - O</td>
<td>3093 (77.7 %)</td>
<td>2089 (80.6 %)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 - C</td>
<td>631 (15.9 %)</td>
<td>374 (14.4 %)</td>
</tr>
<tr>
<td>Light</td>
<td>Dark (1 if dark, otherwise 0)</td>
<td>0</td>
<td>3067 (77.1 %)</td>
<td>2012 (77.7 %)</td>
</tr>
<tr>
<td></td>
<td>The presence of ASCS (1 if Yes, otherwise 0)</td>
<td>1</td>
<td>912 (22.9 %)</td>
<td>579 (22.3 %)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>3067 (77.1 %)</td>
<td>2012 (77.7 %)</td>
</tr>
<tr>
<td>ASCS</td>
<td>The presence of ASCS (1 if Yes, otherwise 0)</td>
<td>1</td>
<td>-</td>
<td>2591 (39.4 %)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>3979 (60.6 %)</td>
<td>-</td>
</tr>
<tr>
<td>Peak</td>
<td>Peak period (1 if peak period, otherwise 0)</td>
<td>1</td>
<td>1075 (27.0 %)</td>
<td>698 (26.9 %)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>2904 (73.0 %)</td>
<td>1893 (73.1 %)</td>
</tr>
<tr>
<td>Rear_end</td>
<td>Rear-end (1 if rear-end, otherwise 0)</td>
<td>1</td>
<td>2140 (53.8 %)</td>
<td>1335 (51.5 %)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1839 (46.2 %)</td>
<td>1256 (48.5 %)</td>
</tr>
<tr>
<td>Angle</td>
<td>Angle (1 if angle, otherwise 0)</td>
<td>1</td>
<td>1073 (27 %)</td>
<td>672 (25.9 %)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>2906 (73 %)</td>
<td>1919 (74.1 %)</td>
</tr>
<tr>
<td>Pedestrian</td>
<td>The presence of pedestrian (1 if pedestrian-involved, otherwise 0)</td>
<td>1</td>
<td>21 (0.5 %)</td>
<td>12 (0.5 %)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>3958 (99.5 %)</td>
<td>2579 (99.5 %)</td>
</tr>
<tr>
<td>Variable</td>
<td>Definition</td>
<td>Level/Data Type</td>
<td>Before Period Frequency (Percentage)</td>
<td>After Period Frequency (Percentage)</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>-------------------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>AADT_over_30k</td>
<td>AADT at a road on which a crash occurred (1 if greater than 30k, otherwise 0)</td>
<td>Numeric</td>
<td>3256 (81.8 %)</td>
<td>2272 (87.7 %)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>723 (18.2 %)</td>
<td>319 (12.3 %)</td>
</tr>
<tr>
<td></td>
<td>Speed limit difference between major roads and minor roads at an intersection (1 if equal to or greater than 10 mph, otherwise 0)</td>
<td></td>
<td>2894 (72.7 %)</td>
<td>1882 (70.3 %)</td>
</tr>
<tr>
<td>S_Difover10</td>
<td></td>
<td></td>
<td>1085 (27.3 %)</td>
<td>769 (29.7 %)</td>
</tr>
<tr>
<td>Speed Limit</td>
<td>Speed limit (mph)</td>
<td>Numeric</td>
<td>40.71 (6.20)</td>
<td>39.83 (6.33)</td>
</tr>
<tr>
<td>Signal Distance</td>
<td>Average signal distance on a corridor (miles)</td>
<td>Numeric</td>
<td>0.38 (0.12)</td>
<td>0.36 (0.12)</td>
</tr>
</tbody>
</table>

\* S.D.-Standard deviation

5 Results and analysis

Four random-parameter ordered regression models are firstly compared. The model estimation results on the association between the presence of ASCS and the crash severity and the relationship between other contributing factors and the crash severity are then presented.

5.1 Comparison of random-parameter ordered regression models

The following four models are estimated and compared:

- Random-parameter ordered probit model with observed heterogeneity (ROP)
- Random-parameter ordered logit model with observed heterogeneity (ROL)
Random-parameter ordered probit model (RP)

Random-parameter ordered logit model (RL)

In the model estimation results in Table 6, a positive sign of the coefficient of predictors is associated with higher crash severity, while a negative sign of the coefficient is associated with lower crash severity. From the negative sign of the coefficient (i.e., -0.113 for the RP model and -0.205 for the RL model) of the ASCS variable in Table 6, it shows that the presence of ASCS is associated with lower crash severity. Since the standard deviation associated with ASCS is found to be statistically insignificant in the RP, RL, ROP, and ROL models, ASCS is not considered as a random parameter for these models. Instead, ASCS is considered as a fixed parameter for the RP and RL models, and a varying parameter depending on intersection/corridor-level variables (i.e., S_Difover10 and Signal Distance) for the ROP and ROL models. Only the angle variable is considered as a random parameter with the mean (i.e., mean.Angle) and the standard deviation (i.e., S.D. Angle) in these models.

Table 6 Model Estimation Results

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>ROP</th>
<th>ROL</th>
<th>RP</th>
<th>RL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>p-value</td>
<td>Est.</td>
<td>p-value</td>
</tr>
<tr>
<td>Threshold, $\mu_1$</td>
<td>0.858</td>
<td>&lt; 0.001</td>
<td>1.675</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.554</td>
<td>&lt; 0.001</td>
<td>-2.822</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Pedestrian</td>
<td>2.078</td>
<td>&lt; 0.001</td>
<td>3.711</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>AADT_over_30k</td>
<td>0.235</td>
<td>&lt; 0.001</td>
<td>0.425</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Speed_Limit</td>
<td>0.009</td>
<td>0.008</td>
<td>0.019</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Light</td>
<td>0.324</td>
<td>&lt; 0.001</td>
<td>0.553</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
As indicated in the previous studies, the AIC difference between two competing models that is greater than 2 (Burnham and Anderson, 2004) or 2.5 (Hilbe, 2011) could be used as a threshold to distinguish different models. Based on the recommendation of these studies, the difference of AIC between two models greater than 2.5 is considered as the threshold to select the preferred models in this study. As indicated in Table 7, the ROP and ROL models are better than the RP and RL models in terms of AIC. As indicated in Table 7 (in the last two columns), in terms of AIC, there are no significant differences between the ROP and ROL models, as well as between the RP and RL models.

Table 7 Model Comparison based on AIC Difference

<table>
<thead>
<tr>
<th>Model Comparisons</th>
<th>AIC</th>
<th>AIC</th>
<th>AIC</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8076.28</td>
<td>8078.20</td>
<td>8082.41</td>
<td>8084.15</td>
</tr>
</tbody>
</table>

a: ROP stands for random-parameter ordered probit model with observed heterogeneity; ROL stands for random-parameter ordered logit model with observed heterogeneity; RP stands for random-parameter ordered probit model; RL stands for random-parameter ordered logit model.

b: estimated threshold in Eq. (5) – (6)
c: mean of the coefficient
d: observed variable used to capture the observed heterogeneity of ASCS
e: standard deviation of the coefficient
In addition to using AIC, a likelihood ratio test (Washington et al., 2020) for comparing nested models (e.g., RL VS. ROL, RP VS. ROP, RL VS. ROP, RP VS. ROL) is conducted in this study to identify a superior model with better goodness of fit of the model, as shown in Table 8. The likelihood ratio Chi-squared statistics are statistically significant at a 0.05 significance level, suggesting that the ROP and ROL models are better than the RP and RL models in terms of the goodness of fit of the model. The likelihood ratio tests are not conducted for comparing non-nested models (e.g., ROP VS. ROL, RP VS. RL) as the likelihood ratio test does not apply to compare non-nested models. Based on both AIC (Table 7) and the likelihood ratio test (Table 8) findings, the ROP and ROL models are better than the RP and RL models in terms of both the AIC and goodness of fit of the model.

Table 8 Likelihood Ratio Test Results for Nested Models

<table>
<thead>
<tr>
<th>Model comparison</th>
<th>RL VS. ROL&lt;sup&gt;a&lt;/sup&gt;</th>
<th>RP VS. ROP&lt;sup&gt;a&lt;/sup&gt;</th>
<th>RL VS. ROP&lt;sup&gt;a&lt;/sup&gt;</th>
<th>RP VS. ROL&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference of degrees of freedom between two</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Model comparison</td>
<td>RL VS. ROL</td>
<td>RP VS. ROP&lt;sup&gt;a&lt;/sup&gt;</td>
<td>RL VS. ROP&lt;sup&gt;a&lt;/sup&gt;</td>
<td>RP VS. ROL&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>------------------</td>
<td>------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>competing models</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood ratio Chi-squared statistic</td>
<td>12.28</td>
<td>12.44</td>
<td>14.78</td>
<td>9.94</td>
</tr>
<tr>
<td>p-value</td>
<td>0.006&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.006&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.002&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.02&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Superior model (with better goodness of fit of the model)</td>
<td>ROL</td>
<td>ROP</td>
<td>ROP</td>
<td>ROL</td>
</tr>
</tbody>
</table>

<sup>a</sup>: ROP stands for random-parameter ordered probit model with observed heterogeneity; ROL stands for random-parameter ordered logit model with observed heterogeneity; RP stands for random-parameter ordered probit model; RL stands for random-parameter ordered logit model.

<sup>b</sup>: statistically significant at a 0.05 significance level.

Since the ROL and ROP models are compared as non-nested models, the likelihood ratio test (Washington et al., 2020) does not apply to the comparison of the ROL and ROP models. Alternatively, three metrics: precision, recall, and accuracy, which are widely used for evaluating a classification model, are used for evaluating the performance of the ROP and ROL models. A training dataset with 80% of the sample and a test dataset with 20% of the sample are obtained. The training dataset and test dataset are randomly sampled. During the sampling, both datasets are ensured to have similar percentages of data points by category (i.e., by crash severity outcomes). The sampling procedure is repeated 30 times (Rahman et al., 2019; Xie et al., 2019). For each time, the model is developed using the training dataset, and then the model is evaluated using the test dataset. The three metrics are evaluated for 30 times, and the results are averaged and presented in Table 9. The precision and recall are evaluated for each crash severity level (i.e., O, C, or KAB). Also, the overall precision and recall are evaluated by computing the micro-average values of the precision and recall, and the results are shown in Table 9. A t-test is
conducted to determine if the means of evaluated metrics for the ROP and ROL models are significantly different from each other. In terms of accuracy, overall precision, and overall recall, the ROP model outperforms the ROL model. The results of accuracy, overall precision, and overall recall for the ROP and ROL models are significantly different from each other at a 0.05 significance level.

Table 9 Classification Model Performance Metrics

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy For O</th>
<th>Random-Parameter Ordered Probit Model</th>
<th>For C KAB</th>
<th>Overall (Micro-average)</th>
<th>Recall For O</th>
<th>For C KAB</th>
<th>Overall (Micro-average)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROP</td>
<td>74.8%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>80.0%</td>
<td>19.1%&lt;sup&gt;b&lt;/sup&gt;</td>
<td>49.5%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>74.8%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>92.5%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>10.5%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>ROL</td>
<td>72.6%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>80.1%</td>
<td>17.8%&lt;sup&gt;b&lt;/sup&gt;</td>
<td>37.5%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>72.6%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>89.1%&lt;sup&gt;a&lt;/sup&gt;</td>
<td>13.5%&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

#: ROP stands for random-parameter ordered probit model with observed heterogeneity; ROL stands for random-parameter ordered logit model with observed heterogeneity

a: results of ROP and ROL are statistically different at a 0.05 significance level
b: results of ROP and ROL are statistically different at a 0.1 significance level

5.2 ASCS effects on crash severity

Since the ROP model is deemed as best based on the discussion in Subsection 5.1, only ROP model estimation results are discussed here. As shown in Table 6, observed heterogeneity of ASCS is estimated by two intersection/corridor-level variables (i.e., S_Difover10 and Signal Distance). Other variables related to intersection features such as the number of legs at an intersection and the number of through/left/right lanes at an intersection are attempted in the model, but these variables are not significant. Other variables related to corridor features such as average AADT on a corridor are tried in the model, but they are not significant.
The coefficient of the ASCS variable is a function of intersection/corridor-level variables (i.e., speed limit difference between a major road and a minor road at an intersection that is equal to or greater than 10 mph or S_Difover10, and average signal distance on a corridor or Signal Distance). Based on the estimation of the coefficient, the coefficient of the ASCS variable in the ROP model can be expressed as,

$$\beta_{ASCS,i} = -0.443 + 0.127(x_{S\_Difover10,i}) + 0.637(x_{Signal\_Distance,i})$$  \hspace{1cm} (18)$$

where, \(i\) is an observation ID (i.e., a specific crash). \(x_{S\_Difover10,i}\) is 1 if the speed limit difference between a major street and a minor street at an intersection is equal to or greater than 10 mph and otherwise is 0. \(x_{Signal\_Distance,i}\) is the average signal distance on a corridor.

Figures 2 to 4 show observed heterogeneity in terms of coefficient of the ASCS variable estimated by the ROP model, which represents hierarchical effects of ASCS on the crash severity. The hierarchical effects of ASCS on the crash severity represent the ASCS effect varied by intersection and corridor features. Based on Eq. (18), two linear functions are plotted in Figure 2. In Figure 2, a negative coefficient in the y-axis indicates that the presence of ASCS is associated with lower crash severity, whereas a positive coefficient indicates that the presence of ASCS is associated with higher crash severity. The following observations are derived from Figure 2:

- In Case 1, where speed limit difference between a major street and a minor street at an intersection (intersection feature) is equal to or greater than 10 mph, the coefficient of ASCS increases as the average signal distance on a corridor increases. When the average signal distance on a corridor (corridor feature) exceeds the threshold of 0.49 miles for
Case 1, the coefficient of ASCS becomes positive, suggesting that the presence of ASCS is associated with higher crash severity.

- In Case 2, where speed limit difference between a major street and a minor street at an intersection is less than 10 mph, the coefficient of ASCS increases as the average signal distance on a corridor increases. When the average signal distance on a corridor exceeds the threshold of 0.69 miles for Case 2, the coefficient of ASCS becomes positive, suggesting that the presence of ASCS is associated with higher crash severity.

- The threshold of average signal distance, as discussed earlier, for Case 2 is larger than that for Case 1, indicating that when the speed limit difference between a major street and a minor street at an intersection (i.e., less than 10 mph) is less likely to increase the crash severity, the larger signal distance on a corridor can be accepted to deploy the ASCS without increasing the probability of higher crash severity.

**Figure 2 Coefficient of ASCS VS. Average signal distance**
Figure 3 shows the kernel density of conditional means of the coefficient of the ASCS variable. The conditional means of the coefficient of the ASCS variable is calculated by using Eq. (9). It turns out that the majority (78%) of the conditional means (the unshaded portion in the figure) has negative signs, suggesting that the presence of ASCS is associated with lower crash severity for most of the observations. It is concluded that the presence of ASCS is associated with lower crash severity.

Figure 3 Kernel density of the conditional means for the coefficient of ASCS variable

Figure 4 shows 95% confidence intervals for the conditional means of the coefficient of the ASCS variable in the ROP model for observation IDs from 2600 to 2800 (total number of observations in the sample is 6,570 in this study). The ASCS effect on crash severity varies across different intersections and corridors. In contrast, some crashes have the same ASCS effect.
since they occurred at a similar intersection (same speed limit difference category) on the same corridor.

**Figure 4:** 95% confidence intervals for the conditional means of the coefficient of the ASCS variable in the ROP model for observation IDs from 2600 to 2800

*: Crashes that occurred at intersections where speed limit difference between major streets and minor streets is equal to or greater than 10 mph

#: Crashes that occurred at intersections where speed limit difference between major streets and minor streets is less than 10 mph

The marginal effects for the three crash severity levels are computed, as shown in Table 10. A positive sign of the value in the marginal effects table indicates an increase in the probability of a severity level for the ASCS variable, meaning that such a level is indeed likely to increase due to ASCS. However, a negative sign of the value in the marginal effects table indicates a decrease in
the probability of the severity level for the ASCS variable, meaning that such a level is likely to
decrease due to ASCS.

The marginal effects in Table 10 show that ASCS can reduce the probability of C and KAB for
the majority of intersections and corridors except for the N Lake Dr, SC 642, and US 17 with
speed limit difference between major roads and minor roads equal to or greater than 10 mph. The
marginal effects of ASCS vary in terms of intersection and corridor features. For example, for
US 17A with speed limit difference between major roads and minor roads less than 10 mph,
ASCS reduces the probability of C and KAB by 4.76 % and 2.13 %, respectively, while
increasing the probability of O by 6.89 %. Although the absolute value of the ASCS effect on the
KAB seems to be small, for the case of the small proportion of KAB (average is around 6 %) in
the studied intersections, ASCS is quite effective in reducing the probability of KAB for crashes
that occurred at intersections. The effectiveness of reducing KAB (marginal effect for KAB
divided by the proportion of KAB for the corresponding intersections) is computed in the last
column in Table 10. For example, the highest benefit is achieved for Garlinton Rd by 1.76
%/1.55 %=113.23 %.

Table 10 Marginal Effects of ASCS on Crash Severity Levels

<table>
<thead>
<tr>
<th>Corrid or Feature</th>
<th>Corridor Feature</th>
<th>Marginal Effect for</th>
<th>Marginal Effect for</th>
<th>Marginal Effect for</th>
<th>Proportion of KAB</th>
<th>Effectiveness of</th>
<th>Effectiveness of</th>
</tr>
</thead>
</table>

36
### Spe

<table>
<thead>
<tr>
<th>Speed Limit Difference Equal to or greater than 10 mph?</th>
<th>Average Signal Distance (miles)</th>
<th>O Level</th>
<th>C Level</th>
<th>KAB Level</th>
<th>Level reducing KAB Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>US17A</td>
<td>No</td>
<td>0.27</td>
<td>6.89%</td>
<td>-4.76%</td>
<td>-2.13%</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>0.27</td>
<td>3.87%</td>
<td>-2.61%</td>
<td>-1.25%</td>
</tr>
<tr>
<td>Garlinton Rd</td>
<td>No</td>
<td>0.36</td>
<td>5.58%</td>
<td>-3.82%</td>
<td>-1.76%</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>0.36</td>
<td>5.58%</td>
<td>-3.82%</td>
<td>-1.76%</td>
</tr>
<tr>
<td>N Lake</td>
<td>No</td>
<td>0.55</td>
<td>2.38%</td>
<td>-1.59%</td>
<td>-0.79%</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>0.55</td>
<td>-1.01%</td>
<td>0.66%</td>
<td>0.35%</td>
</tr>
<tr>
<td>SC 642</td>
<td>No</td>
<td>0.52</td>
<td>3.05%</td>
<td>-2.05%</td>
<td>-1.00%</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>0.52</td>
<td>-0.45%</td>
<td>0.30%</td>
<td>0.16%</td>
</tr>
<tr>
<td>US 52</td>
<td>No</td>
<td>0.31</td>
<td>6.31%</td>
<td>-4.35%</td>
<td>-1.97%</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>0.31</td>
<td>3.22%</td>
<td>-2.16%</td>
<td>-1.05%</td>
</tr>
<tr>
<td>US 17</td>
<td>Yes</td>
<td>0.61</td>
<td>-2.15%</td>
<td>1.39%</td>
<td>0.77%</td>
</tr>
</tbody>
</table>

### 5.3 Effects of other contributing factors on crash severity

Note that, the coefficient of the angle variable is a random parameter that follows a random distribution. As shown in Table 6, the mean of the coefficient (mean.Angle) associated with angle is found to be positive and statistically significant in the ROP model. Its standard deviation (S.D.Angle) is also found to be statistically significant, implying the existence of unobserved heterogeneity across observations. The coefficient of the angle variable is estimated to follow a uniform distribution with a mean of 0.330 and a standard deviation of 1.030. It is found that all
observations have a positive coefficient associated with the angle crashes, suggesting an
association between angle crashes and higher crash severity.

As depicted by the marginal effects in **Table 11**, other contributing factors except for the peak
period are associated with higher crash severity levels (i.e., C and KAB) while less likely to be a
lower crash severity (i.e., O). Crashes involving pedestrians will lead to higher crash severity
levels and increase the probability of KAB by 57.69%. The presence of AADT over 30,000
vehicles/day results in a corresponding increase in the likelihood of C and KAB given the critical
role of high traffic volume in overall crashes at the signalized intersections. Not surprisingly, an
increase in the posted speed limit is associated with a greater likelihood of C and KAB. The
higher speed limit naturally results in a higher vehicle operational speed, with an increase in the
severity of crashes. The bad light condition (i.e., dark) of roadways is associated with a higher
likelihood of C and KAB. The peak period leads to lower crash severity (i.e., O). During peak
periods, the traffic volume is relatively higher compared to off-peak periods, which would
contribute to lower average speeds of the vehicles during peak periods, thus resulting in reduced
crash severity. The crash, which is either rear-end or angle crash, is associated with higher
probability of C and KAB.

**Table 11 Marginal Effects of other Contributing Factors**

<table>
<thead>
<tr>
<th>Other Contributing Factors</th>
<th>Marginal Effect for O</th>
<th>Marginal Effect for C</th>
<th>Marginal Effect for KAB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedestrian</td>
<td>-69.83 %</td>
<td>12.14 %</td>
<td>57.69 %</td>
</tr>
<tr>
<td>AADT_over_30k</td>
<td>-5.92 %</td>
<td>4.11 %</td>
<td>1.81 %</td>
</tr>
</tbody>
</table>
### Conclusions

This study has investigated the hierarchical effects of ASCS on the crash severity by developing random-parameter ordered regression models with observed heterogeneity, which accounts for both observed and unobserved heterogeneity. Crash data from six ASCS corridors with 65 signalized intersections are used to develop the models. Four different random-parameter ordered regression models (two ordered probit models, and two ordered logit models) are established and compared. It is found that the ROP and ROL models perform better than the RP and RL models in terms of the AIC and the goodness of fit of the model. The ROP model outperforms the ROL model in terms of classification model performance measures: accuracy, overall precision, and overall recall. This study is unique as it demonstrates the existence of the hierarchical effects of ASCS on the crash severity. The analyses reveal that the presence of ASCS is associated with lower crash severity. Speed limit difference between major streets and minor streets at an intersection (intersection feature) and average signal distance on a corridor (corridor feature) are found to be capable of capturing the hierarchical effects of ASCS on the crash severity. Other
variables related to intersection features such as the number of legs at an intersection and number of through/left/right lanes at an intersection and corridor features such as average AADT on a corridor are attempted in the model, but these variables are not statistically significant. Thus, these variables are not able to capture the hierarchical effects of ASCS on the crash severity in this study. In the future, variables related to zonal features such as population densities could be accounted for to capture the hierarchical effects of ASCS on the crash severity. Other contributing factors, such as annual average daily traffic, speed limit, lighting, crash type (rear-end, angle), pedestrian involvements, are associated with higher crash severity. The peak period leads to lower crash severity. Unobserved heterogeneity of the effect of angle crashes on crash severity is found to exist across the observations by using the uniform distribution to explicitly account for crash-specific variations in the effects of angle crashes.

To evaluate the performance of the ROP model on the high proportion of severe crash severity, further studies on investigating the applicability of the model in the case of the high proportion of severe crash severity may be carried out once a suitable case study is available.

Identifying the hierarchical effects of ASCS on the crash severity could help transportation agencies achieve higher safety benefits by selecting ASCS deployment sites by considering specific intersection and corridor features. The findings of this study have several practical implications for establishing ASCS implementation guidelines from the standpoint of safety. Two useful metrics, speed limit difference between a major street and a minor street at an intersection (intersection feature) and average signal distance on a corridor (corridor feature), could help transportation agencies to deploy ASCS appropriately. Two practical implications are
found: 1) when speed limit difference between major streets and minor streets at an intersection is equal to or greater than 10 mph, and the average signal distance on a corridor is less than the threshold of 0.49 miles, the ASCS is more likely associated with lower crash severity; and 2) when speed limit difference between major streets and minor streets at an intersection is less than 10 mph, and the average signal distance on a corridor is less than the threshold of 0.69 miles, the ASCS is associated with lower crash severity. This finding is related to the particular type of ASCS, and future studies may be conducted to include multiple types of ASCS.

Declarations of Interest: None.

Acknowledgments

The authors acknowledge the South Carolina Department of Transportation (SCDOT), which provided funding (SCDOT Research Project: SPR 737) for this research. The authors also sincerely acknowledge SCDOT SPR 737 Steering Committee for providing valuable guidance to the authors. The contents of this paper reflect the views of the authors who are responsible for the facts and the accuracy of the presented data. The contents do not reflect the official views of SCDOT or FHWA. This paper does not constitute a standard, specification, or regulation.

REFERENCE


https://doi.org/10.22004/ag.econ.241825


