Accident Analysis and Prevention 41 (2009) 829–838

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Markov switching multinomial logit model: An application to accident-injury severities

Nataliya V. Malyshkina*, Fred L. Mannering

School of Civil Engineering, 550 Stadium Mall Drive, Purdue University, West Lafayette, IN 47907, United States

ARTICLE INFO

Article history:
Received 28 December 2008
Received in revised form 12 April 2009
Accepted 14 April 2009

Keywords:
Accident-injury severity
Multinomial logit
Markov switching
Bayesian
MCMC

ABSTRACT

In this study, two-state Markov switching multinomial logit models are proposed for statistical modeling of accident-injury severities. These models assume Markov switching over time between two unobserved states of roadway safety as a means of accounting for potential unobserved heterogeneity. The states are distinct in the sense that in different states accident-severity outcomes are generated by separate multinomial logit processes. To demonstrate the applicability of the approach, two-state Markov switching multinomial logit models are estimated for severity outcomes of accidents occurring on Indiana roads over a four-year time period. Bayesian inference methods and Markov Chain Monte Carlo (MCMC) simulations are used for model estimation. The estimated Markov switching models result in a superior statistical fit relative to the standard (single-state) multinomial logit models for a number of roadway classes and accident types. It is found that the more frequent state of roadway safety is correlated with better weather conditions and that the less frequent state is correlated with adverse weather conditions.

1. Introduction

In the past there have been a large number of studies that have focused on modeling accident-severity outcomes. Common accident-severity modeling approaches include multinomial logit models, nested logit models, mixed logit models and ordered probit models (O'Donnell and Connor, 1996; Shankar and Mannering, 1996; Shankar et al., 1996; Duncan et al., 1998; Chang and Mannering, 1999; Carson and Mannering, 2001; Khattak, 2001; Khattak et al., 2002; Kockelman and Kweon, 2002; Lee and Mannering, 2002; Abdel-Aty, 2003; Kweon and Kockelman, 2003; Ulfarsson and Mannering, 2004; Yamamoto and Shankar, 2004; Khorashadi et al., 2005; Eluru and Bhat, 2007; Savolainen and Mannering, 2007; Milton et al., 2008). All these modeling approaches involve nonlinear regression analysis of the observed accident-injury severity outcomes on various accident characteristics and related factors (such as roadway and driver characteristics, environmental factors, and so on). However, all of the approaches to date have assumed a single-state of safety. That is, for a given period of time, the parameter estimates are assumed to be constant or to vary over observations but not over time.

In an earlier paper, Malyshkina et al. (2009) developed two-state Markov switching count-data models of accident frequencies. In the current study we extend this approach to accident-injury severity by considering two-state Markov switching multinomial logit models for predicting accident-severity outcomes.[1] These models assume that there are two unobserved states[2] of roadway safety with regard to injury severity, that roadway entities (roadway segments) can switch between these states over time, and that the switching process is Markovian. The two states intend to account for possible unobserved heterogeneity effects in roadway safety, which may be caused by various unpredictable, unidentified or unobservable risk factors that influence accident-severity. Because the risk factors can interact and change, roadway entities can switch between the two states over time. The two-state Markov switching multinomial logit model assumes separate multinomial logit processes for accident-severity in the two states.

2. Model specification

Markov switching models are parametric and can be fully specified by $f(\mathbf{Y}|\mathbf{\Theta}, \mathcal{M})$, which is the conditional probability distribution of the vector of all observations $\mathbf{Y}$, given the vector of all parameters $\mathbf{\Theta}$ of model $\mathcal{M}$. First, let us consider $\mathbf{Y}$. Let N_t be the number of accidents observed during time period t, where $t = 1, 2, \ldots, T$

* Corresponding author.
E-mail addresses: nmalyshkina@alumni.purdue.edu (N.V. Malyshkina), flm@ecn.purdue.edu (F.L. Mannering).

0001-4575/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2009.04.006

[1] The approach we present could also be adapted for different severity models such as the mixed logit and generalized forms of ordered probability models that overcome the inherent limitations of standard ordered probability models in their application to accident-severity analysis (Eluru et al., 2008; Milton et al., 2008).

[2] Two states are used for simplicity and demonstration of method. One could easily consider three or more states in this process.

and T is the total number of time periods. Let there be I discrete outcomes observed for accident-severity (for example, $I = 3$ and these outcomes are fatality, injury and property damage only). Let us introduce accident-severity outcome dummies $\delta_{t,n}^{(i)}$ that are equal to unity if the ith severity outcome is observed in the nth accident that occurs during time period t, and to zero otherwise. Here $i = 1, 2, \ldots, I$, $n = 1, 2, \ldots, N_t$ and $t = 1, 2, \ldots, T$. Then, our observations are the accident-severity outcomes, and the vector of all observations $\mathbf{Y} = \{\delta_{t,n}^{(i)}\}$ includes all outcomes observed in all accidents that occur during all time periods. Second, let us consider model specification variable $\mathcal{M}$. It is $\mathcal{M} = \{M, \mathbf{x}_{t,n}\}$ and includes the model's name M (for example, M = "multinomial logit") and the vector $\mathbf{x}_{t,n}$ of all accident-characteristic variables (weather and environment conditions, vehicle and driver characteristics, roadway and pavement properties, and so on).

To define the data probability distribution $f(\mathbf{Y}|\boldsymbol{\Theta}, \mathcal{M})$, we first introduce an unobserved (latent) state variable s_t, which determines the state of all roadway entities during time period t. At each t, the state variable s_t can assume only two values: $s_t = 0$ corresponds to one state and $s_t = 1$ corresponds to the other state ($t = 1, 2, \ldots, T$). The state variable s_t is assumed to follow a stationary two-state Markov chain process in time,[3] which can be specified by time-independent transition probabilities as

$$P(s_{t+1} = 1|s_t = 0) = p_{0\to 1}, \quad P(s_{t+1} = 0|s_t = 1) = p_{1\to 0}. \tag{1}$$

Here, for example, $P(s_{t+1} = 1|s_t = 0)$ is the conditional probability of $s_{t+1} = 1$ at time $t + 1$, given that $s_t = 0$ at time t. Transition probabilities $p_{0\to 1}$ and $p_{1\to 0}$ are unknown parameters to be estimated from accident-severity data. The stationary unconditional probabilities of states $s_t = 0$ and $s_t = 1$ are $\bar{p}_0 = p_{1\to 0}/(p_{0\to 1} + p_{1\to 0})$ and $\bar{p}_1 = p_{0\to 1}/(p_{0\to 1} + p_{1\to 0})$, respectively.[4] Without loss of generality, we assume that (on average) state $s_t = 0$ occurs more or equally frequently than state $s_t = 1$. Therefore, $\bar{p}_0 \geq \bar{p}_1$, and we obtain restriction[5]

$$p_{0\to 1} \leq p_{1\to 0}. \tag{2}$$

We refer to states $s_t = 0$ and $s_t = 1$ as "more frequent" and "less frequent" states, respectively.

Next, a two-state Markov switching multinomial logit (MSMNL) model assumes multinomial logit (MNL) data-generating processes (McFadden, 1981) for accident-severity in each of the two states. With this, the probability of the ith severity outcome observed in the nth accident during time period t is

$$P_{t,n}^{(i)} = \begin{cases} \dfrac{\exp(\boldsymbol{\beta}'_{(0),i}\mathbf{x}_{t,n})}{\sum_{j=1}^{I}\exp(\boldsymbol{\beta}'_{(0),j}\mathbf{x}_{t,n})} & \text{if } s_t = 0, \\[2ex] \dfrac{\exp(\boldsymbol{\beta}'_{(1),i}\mathbf{x}_{t,n})}{\sum_{j=1}^{I}\exp(\boldsymbol{\beta}'_{(1),j}\mathbf{x}_{t,n})} & \text{if } s_t = 1, \end{cases} \tag{3}$$
$$i = 1, 2, \ldots, I, \quad n = 1, 2, \ldots, N_t, \quad t = 1, 2, \ldots, T,$$

Here prime means transpose (so $\boldsymbol{\beta}'_{(0),i}$ is the transpose of $\boldsymbol{\beta}_{(0),i}$). Parameter vectors $\boldsymbol{\beta}_{(0),i}$ and $\boldsymbol{\beta}_{(1),i}$ are unknown estimable parameters of the two standard multinomial logit probability mass functions (Washington et al., 2003) in the two states, $s_t = 0$ and $s_t = 1$, respectively. We set the first component of $\mathbf{x}_{t,n}$ to unity, and, therefore, the first components of vectors $\boldsymbol{\beta}_{(0),i}$ and $\boldsymbol{\beta}_{(1),i}$ are the intercepts in the two states. In addition, without loss of generality, we set all β-parameters for the last severity outcome to zero,[6] $\boldsymbol{\beta}_{(0),I} = \boldsymbol{\beta}_{(1),I} = \mathbf{0}$.

If accident events are assumed to be independent, the data probability distribution function is

$$f(\mathbf{Y}|\boldsymbol{\Theta}, \mathcal{M}) = \prod_{t=1}^{T}\prod_{n=1}^{N_t}\prod_{i=1}^{I}[P_{t,n}^{(i)}]^{\delta_{t,n}^{(i)}}. \tag{4}$$

Here, because the state variables s_t are unobservable, the vector of all estimable parameters $\boldsymbol{\Theta}$ must include all states, in addition to model parameters ($\beta - s$) and transition probabilities. Thus, $\boldsymbol{\Theta} = [\boldsymbol{\beta}'_{(0)}, \boldsymbol{\beta}'_{(1)}, p_{0\to 1}, p_{1\to 0}, \mathbf{S}']'$, where vector $\mathbf{S} = [s_1, s_2, \ldots, s_T]'$ has length T and contains all state values. Eqs. (1)–(4) define the two-state Markov switching multinomial logit (MSMNL) model considered here. The likelihood function is defined as $L(\boldsymbol{\Theta}) = f(\mathbf{Y}|\boldsymbol{\Theta}, \mathcal{M})$ (Robert, 2001).[7]

3. Model estimation methods

Statistical estimation of Markov switching models is complicated by unobservability of the state variables s_t.[8] As a result, the traditional maximum likelihood estimation (MLE) procedure is of very limited use for Markov switching models. Instead, a Bayesian inference approach is used. Given a model $\mathcal{M}$ with data probability distribution function $f(\mathbf{Y}|\boldsymbol{\Theta}, \mathcal{M})$, the Bayes formula is

$$f(\boldsymbol{\Theta}|\mathbf{Y}, \mathcal{M}) = \frac{f(\mathbf{Y}, \boldsymbol{\Theta}|\mathcal{M})}{f(\mathbf{Y}|\mathcal{M})} = \frac{f(\mathbf{Y}|\boldsymbol{\Theta}, \mathcal{M})\pi(\boldsymbol{\Theta}|\mathcal{M})}{\int f(\mathbf{Y}, \boldsymbol{\Theta}|\mathcal{M})\,d\boldsymbol{\Theta}}. \tag{5}$$

Here $f(\boldsymbol{\Theta}|\mathbf{Y}, \mathcal{M})$ is the posterior probability distribution of model parameters $\boldsymbol{\Theta}$ conditional on the observed data $\mathbf{Y}$ and model $\mathcal{M}$. Function $f(\mathbf{Y}, \boldsymbol{\Theta}|\mathcal{M})$ is the joint probability distribution of $\mathbf{Y}$ and $\boldsymbol{\Theta}$ given model $\mathcal{M}$. Function $f(\mathbf{Y}|\mathcal{M})$ is the marginal likelihood for model $\mathcal{M}$)—the probability distribution of data $\mathbf{Y}$ given model $\mathcal{M}$. Function $\pi(\boldsymbol{\Theta}|\mathcal{M})$ is the prior probability distribution of parameters that reflects prior knowledge about $\boldsymbol{\Theta}$. The intuition behind Eq. (5) is straightforward: given model $\mathcal{M}$, the posterior distribution accounts for both the observations $\mathbf{Y}$ and our prior knowledge of $\boldsymbol{\Theta}$.

In our study (and in most practical studies), the direct application of Eq. (5) is not feasible because the parameter vector $\boldsymbol{\Theta}$ contains too many components, making integration over $\boldsymbol{\Theta}$ in Eq. (5) extremely difficult. However, the posterior distribution $f(\boldsymbol{\Theta}|\mathbf{Y}, \mathcal{M})$ in Eq. (5) is known up to its normalization constant, $f(\boldsymbol{\Theta}|\mathbf{Y}, \mathcal{M}) \propto f(\mathbf{Y}|\boldsymbol{\Theta}, \mathcal{M})\pi(\boldsymbol{\Theta}|\mathcal{M})$. As a result, we use Markov Chain Monte Carlo (MCMC) simulations, which provide a convenient and practical computational methodology for sampling from a probability distribution known up to a constant (the posterior distribution in our case). Given a large enough posterior sample of parameter vector $\boldsymbol{\Theta}$, any posterior expectation and variance can be found and Bayesian inference can be readily applied. The reader is referred to Malyshkina et al. (2009) where the choice of the prior distribution $\pi(\boldsymbol{\Theta}|\mathcal{M})$ and the MCMC simulation algorithm is described in the context of Markov switching count-data (negative binomial) models.[9] Although, in this study we estimate a two-state

[3] Markov property means that the probability distribution of s_{t+1} depends only on the value s_t at time t, but not on the previous history $s_{t-1}, s_{t-2}, \ldots$. Stationarity of $\{s_t\}$ is in the statistical sense.

[4] These can be found from stationarity conditions $\bar{p}_0 = (1 - p_{0\to 1})\bar{p}_0 + p_{1\to 0}\bar{p}_1$, $\bar{p}_1 = p_{0\to 1}\bar{p}_0 + (1 - p_{1\to 0})\bar{p}_1$ and $\bar{p}_0 + \bar{p}_1 = 1$.

[5] Without any loss of generality, restriction (2) is introduced for the purpose of avoiding the problem of state label switching $0 \leftrightarrow 1$. This problem would otherwise arise because of the symmetry of Eqs. (1)–(4) under the label switching.

[6] This can be done because $\mathbf{x}_{t,n}$ are assumed to be independent of the outcome i.

[7] In general, a likelihood function can be defined as $L(\boldsymbol{\Theta}) \propto f(\mathbf{Y}|\boldsymbol{\Theta}, \mathcal{M})$.

[8] Below we will have 208 time periods ($T = 208$). In this case, there are 2^{208} possible combinations for value of vector $\mathbf{S} = [s_1, s_2, \ldots, s_T]'$.

[9] Our priors for $\beta - s$, $p_{0\to 1}$ and $p_{1\to 0}$ are flat or nearly flat, while the prior for the states $\mathbf{S}$ reflects the Markov process property, specified by Eq. (1).

Markov switching multinomial logit model for accident-severity outcomes and Malyshkina et al. (2009) estimated a two-state Markov switching negative binomial model for accident frequencies, this difference is not essential for Bayesian–MCMC model estimation methods. In fact, the main difference is in the data probability distribution (multinomial logit as opposed to negative binomial). So we used the same numerical MCMC code, written in the MATLAB programming language, discussed and proven in previous work (Malyshkina, 2008; Malyshkina et al., 2009). We tested our code on artificial data sets of accident-severity outcomes. The test procedure included a generation of artificial data with a known model. Then these data were used to estimate the underlying model by means of our simulation code. With this procedure we found that the MSMNL models, used to generate the artificial data, were reproduced successfully with our estimation code.

For comparison of different models we use a formal Bayesian approach. Let there be two models $\mathcal{M}_1$ and $\mathcal{M}_2$ with parameter vectors $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$, respectively. Assuming that we have equal preferences of these models, their prior probabilities are $\pi(\mathcal{M}_1) = \pi(\mathcal{M}_2) = 1/2$. In this case, the ratio of the models' posterior probabilities, $P(\mathcal{M}_1|\mathbf{Y})$ and $P(\mathcal{M}_2|\mathbf{Y})$, is equal to the Bayes factor. The later is defined as the ratio of the models' marginal likelihoods (see Kass and Raftery, 1995). Thus, we have

$$\frac{P(\mathcal{M}_2|\mathbf{Y})}{P(\mathcal{M}_1|\mathbf{Y})} = \frac{f(\mathcal{M}_2,\mathbf{Y})/f(\mathbf{Y})}{f(\mathcal{M}_1,\mathbf{Y})/f(\mathbf{Y})} = \frac{f(\mathbf{Y}|\mathcal{M}_2)\pi(\mathcal{M}_2)}{f(\mathbf{Y}|\mathcal{M}_1)\pi(\mathcal{M}_1)} = \frac{f(\mathbf{Y}|\mathcal{M}_2)}{f(\mathbf{Y}|\mathcal{M}_1)}, \tag{6}$$

where $f(\mathcal{M}_1,\mathbf{Y})$ and $f(\mathcal{M}_2,\mathbf{Y})$ are the joint distributions of the models and the data, $f(\mathbf{Y})$ is the unconditional distribution of the data. As in Malyshkina et al. (2009), to calculate the marginal likelihoods $f(\mathbf{Y}|\mathcal{M}_1)$ and $f(\mathbf{Y}|\mathcal{M}_2)$, we use the harmonic mean formula $f(\mathbf{Y}|\mathcal{M})^{-1} = E[f(\mathbf{Y}|\mathbf{\Theta},\mathcal{M})^{-1}|\mathbf{Y}]$, where $E(\ldots|\mathbf{Y})$ means posterior expectation calculated by using the posterior distribution. If the ratio in Eq. (6) is larger than one, then model $\mathcal{M}_2$ is favored, if the ratio is less than one, then model $\mathcal{M}_1$ is favored. An advantage of the use of Bayes factors is that it has an inherent penalty for including too many parameters in the model and guards against overfitting.

To evaluate the performance of model $\{\mathcal{M},\mathbf{\Theta}\}$ in fitting the observed data $\mathbf{Y}$, we carry out the Pearson's χ^2 goodness-of-fit test (Maher and Summersgill, 1996; Cowan, 1998; Wood, 2002; Press et al., 2007). We perform this test by Monte Carlo simulations to find the distribution of the Pearson's χ^2quantity, which measures the discrepancy between the observations and the model predictions (Cowan, 1998). This distribution is then used to find the goodness-of-fit p-value, which is the probability that χ^2 exceeds the observed value of χ^2 under the hypothesis that the model is true (the observed value of χ^2 is calculated by using the observed data $\mathbf{Y}$). For additional details, please see Malyshkina (2008) and Malyshkina et al. (2009).

4. Estimation results

The severity outcome of an accident is determined by the injury level sustained by the most injured individual (if any) involved into the accident. In this study we consider three accident-severity outcomes: "fatality", "injury" and "PDO (property damage only)", which we number as $i = 1, 2, 3$, respectively ($I = 3$). We use data from 811,720 accidents that were observed in Indiana in 2003–2006. As in Malyshkina et al. (2009), we use weekly time periods, $t = 1, 2, 3, \ldots, T = 208$ in total.[10] Thus, the state s_t can change every week. To increase the predictive power of our models, we consider accidents separately for each combination of accident type (1-vehicle and 2-vehicle) and roadway class (interstate highways, US routes, state routes, county roads, other roads). We do not consider accidents with more than two vehicles involved.[11] Thus, in total, there are 10 roadway-class-accident-type combinations that we considered. For each roadway-class-accident-type combination the following three types of accident-severity models are estimated

- First, we estimate a standard multinomial logit (MNL) model without Markov switching by maximum likelihood estimation (MLE).[12] We refer to this model as "MNL-by-MLE".
- Second, we estimate the same standard multinomial logit model by the Bayesian inference approach and the MCMC simulations. We refer to this model as "MNL-by-MCMC". As one expects, the estimated MNL-by-MCMC model turned out to be very similar to the corresponding MNL-by-MLE model (estimated for the same roadway-class-accident-type combination).
- Third, we estimate a two-state Markov switching multinomial logit (MSMNL) model by the Bayesian–MCMC methods. In order to make comparison of explanatory variable effects in different models straightforward, in the MSMNL model we use only those explanatory variables that enter the corresponding standard MNL model.[13] To obtain the final MSMNL model reported here, we also consecutively construct and use 60%, 85% and 95% Bayesian credible intervals for evaluation of the statistical significance of each β-parameter. As a result, in the final model some components of $\boldsymbol{\beta}_{(0)}$ and $\boldsymbol{\beta}_{(1)}$ are restricted to zero or restricted to be the same in the two states.[14] We refer to this final model as "MSMNL".

Note that the two states, and thus the MSMNL models, do not have to exist for every roadway-class-accident-type combination. For example, they will not exist if all estimated model parameters turn out to be statistically the same in the two states, $\boldsymbol{\beta}_{(0)} = \boldsymbol{\beta}_{(1)}$, (which suggests the two states are identical and the MSMNL models reduce to the corresponding standard MNL models). Also, the two states will not exist if all estimated state variables s_t turn out to be close to zero, resulting in $p_{0\to1} \ll p_{1\to0}$ [compare to Eq. (2)], then the less frequent state $s_t = 1$ is not realized and the process stays in state $s_t = 0$.

Turning to the estimation results, the findings show that two states of roadway safety and the appropriate MSMNL models exist for severity outcomes of 1-vehicle accidents occurring on all roadway classes (interstate highways, US routes, state routes, county roads, other roads), and for severity outcomes of 2-vehicle accidents occurring on "other roads". We did not find two states in the cases of 2-vehicle accidents on interstate highways, US routes, state routes and county roads (in these cases all estimated state variables s_t were found to be close to zero). The model estimation results for severity

[10] A week is from Sunday to Saturday, there are 208 full weeks in the 2003–2006 time interval.

[11] Among 811,720 accidents 241,011 (29.7%) are 1-vehicle, 525,035 (64.7%) are 2-vehicle, and only 45,674 (5.6%) are accidents with more than two vehicles involved.

[12] To obtain parsimonious standard models, estimated by MLE, we choose the explanatory variables and their dummies by using the Akaike Information Criterion (AIC) and the 5% statistical significance level for the two-tailed t-test. Minimization of $AIC = 2K - 2LL$, were K is the number of free continuous model parameters and LL is the log-likelihood, ensures an optimal choice of explanatory variables in a model and avoids overfitting (Tsay, 2002; Washington et al., 2003). For details on variable selection, see Malyshkina (2006).

[13] A formal Bayesian approach to model variable selection is based on evaluation of model's marginal likelihood and the Bayes factor (6). Unfortunately, because MCMC simulations are computationally expensive, evaluation of marginal likelihoods for a large number of trial models is not feasible in our study.

[14] A β-parameter is restricted to zero if it is statistically insignificant. A β-parameter is restricted to be the same in the two states if the difference of its values in the two states is statistically insignificant. A $(1-a)$ credible interval is chosen in such way that the posterior probabilities of being below and above it are both equal to $a/2$ (we use significance levels a = 40%, 15%, 5%).

Table 1

Estimation results for multinomial logit models of severity outcomes of one-vehicle accidents on Indiana interstate highways.

Variable	MNL-by-MLE [a]		MNL-by-MCMC [b]		MSMNL [c]			
	Fatality	Injury	Fatality	Injury	State $s=0$		State $s=1$	
					Fatality	Injury	Fatality	Injury
Intercept (constant term)	$-11.9_{-13.7}^{-10.1}$	$-3.69_{-3.84}^{-3.53}$	$-12.4_{-14.5}^{-10.6}$	$-3.72_{-3.88}^{-3.56}$	$-12.2_{-14.4}^{-10.5}$	$-3.98_{-4.17}^{-3.79}$	$-12.2_{-14.4}^{-10.5}$	$-3.22_{-3.45}^{-2.98}$
Summer season, June–August (1 if yes, 0 if no)	$.235_{.142}^{.329}$	$.235_{.142}^{.329}$	$.237_{.143}^{.329}$	$.237_{.143}^{.329}$	$.176_{.0551}^{.293}$	$.176_{.0551}^{.293}$	$.176_{.0551}^{.293}$	$.615_{.282}^{.959}$
Thursday (1 if yes, 0 if no)	$-.798_{-1.48}^{-.115}$	–	$-.853_{-1.59}^{-.206}$	–	$-.872_{-1.61}^{-.225}$	–	$-.872_{-1.61}^{-.225}$	–
Construction at the accident location (1 if yes, 0 if no)	$-.418_{-.623}^{-.213}$	$-.418_{-.623}^{-.213}$	$-.425_{-.632}^{-.224}$	$-.425_{-.632}^{-.224}$	$-.566_{-.822}^{-.319}$	$-.566_{-.822}^{-.319}$	$-.566_{-.822}^{-.319}$	–
Daylight or street lights are lit up if dark (1 if yes, 0 if no)	$-.392_{-.748}^{-.0368}$	$.137_{.0501}^{.224}$	$-.387_{-.740}^{-.0301}$	$.143_{.0568}^{.230}$	$-.378_{-.729}^{-.0236}$	$.139_{.0522}^{.226}$	$-.378_{-.729}^{-.0236}$	$.139_{.0522}^{.226}$
Precipitation: rain/freezing rain/snow/sleet/hail (1 if yes, 0 if no)	$-1.38_{-1.92}^{-.830}$	$-.361_{-.457}^{-.264}$	$-1.41_{-1.99}^{-.884}$	$-.363_{-.460}^{-.267}$	$-1.54_{-2.10}^{-1.03}$	$-.563_{-.729}^{-.404}$	$-1.54_{-2.10}^{-1.03}$	–
Roadway surface is covered by snow/slush (1 if yes, 0 if no)	$-1.28_{-2.46}^{-.0917}$	$-.432_{-.583}^{-.280}$	$-1.43_{-2.84}^{-.328}$	$-.438_{-.590}^{-.288}$	$-.0515_{-.671}^{-.361}$	$-.0515_{-.671}^{-.361}$	$-.0515_{-.671}^{-.361}$	$-.0515_{-.671}^{-.361}$
Roadway median is drivable (1 if yes, 0 if no)	$.571_{.213}^{.929}$	–	$.577_{.223}^{.939}$	–	$.566_{.211}^{.930}$	–	$.566_{.211}^{.930}$	–
Roadway is at curve (1 if yes, 0 if no)	$.114_{.0165}^{.212}$	$.114_{.0165}^{.212}$	$.116_{.0186}^{.213}$	$.116_{.0186}^{.213}$	–	–	–	–
Primary cause of the accident is driver-related (1 if yes, 0 if no)	$4.24_{3.18}^{5.30}$	$1.53_{1.43}^{1.64}$	$4.39_{3.39}^{5.64}$	$1.54_{1.43}^{1.64}$	$4.48_{3.48}^{5.73}$	$2.00_{1.84}^{2.18}$	$4.48_{3.48}^{5.73}$	$.715_{.468}^{.946}$
Help arrived in 20 min or less after the crash (1 if yes, 0 if no)	$.790_{.693}^{.887}$	$.790_{.693}^{.887}$	$.790_{.691}^{.891}$	$.790_{.691}^{.891}$	$.785_{.684}^{.886}$	$.785_{.684}^{.886}$	$.785_{.684}^{.886}$	$.785_{.684}^{.886}$
The vehicle at fault is a motorcycle (1 if yes, 0 if no)	$3.88_{3.17}^{4.59}$	$2.74_{2.36}^{3.12}$	$3.87_{3.13}^{4.57}$	$2.75_{2.37}^{3.15}$	$4.61_{3.74}^{5.49}$	$3.23_{2.70}^{3.83}$	–	$1.39_{.326}^{2.49}$
Age of the vehicle at fault (in years)	$.0285_{.0201}^{.0370}$	$.0285_{.0201}^{.0370}$	$.0286_{.0201}^{.0370}$	$.0286_{.0201}^{.0370}$	–	$.0286_{.0200}^{.0371}$	–	$.0286_{.0200}^{.0371}$
Number of occupants in the vehicle at fault	$.366_{.269}^{.463}$	$.123_{.0859}^{.159}$	$.367_{.264}^{.465}$	$.123_{.0861}^{.159}$	$.366_{.263}^{.464}$	$.124_{.0874}^{.161}$	$.366_{.263}^{.464}$	$.124_{.0874}^{.161}$
Roadway traveled by the vehicle at fault is multi-lane and divided two-way (1 if yes, 0 if no)	$2.60_{1.20}^{4.00}$	–	$2.86_{1.56}^{4.63}$	–	$2.86_{1.56}^{4.66}$	–	$2.86_{1.56}^{4.66}$	–
At least one of the vehicles involved was on fire (1 if yes, 0 if no)	$1.24^{2.12}$	$-.345_{-.665}^{-.0257}$	$1.18_{.206}^{2.02}$	$-.345_{-.669}^{-.0335}$	$1.66_{.621}^{2.56}$	$-.332_{-.659}^{-.0198}$	–	$-.332_{-.659}^{-.0198}$
Gender of the driver at fault (1 if female, 0 if male)	–	$.328_{.246}^{.410}$	–	$.331_{.248}^{.413}$	–	$.224_{.107}^{.338}$	–	$.479_{.328}^{.637}$
Probability of severity outcome [$P_{t,n}^{(i)}$ given by Eq. (3)], averaged over all values of explanatory variables $\mathbf{x}_{t,n}$	.00756	.179	.00724	.176	.00733	.174	.00672	.192
Markov transition probability of jump $0 \to 1$ ($p_{0\to1}$)	–		–		$.151_{.0704}^{.254}$			
Markov transition probability of jump $1 \to 0$ ($p_{1\to0}$)	–		–		$.330_{.164}^{.532}$			
Unconditional probabilities of states 0 and 1 ($\bar{p}_0$ and $\bar{p}_1$)	–		–		$.683_{.540}^{.814}$ and $.317_{.186}^{.460}$			
Total number of free model parameters ($\beta - s$)	25		25		28			
Posterior average of the log-likelihood (LL)	–		$-8486.78_{-8494.61}^{-8480.82}$		$-8396.78_{-8416.57}^{-8379.21}$			
Max(LL): estimated max. log-likelihood (LL) for MLE; maximum observed value of LL for Bayesian–MCMC	−8465.79 (MLE)		−8476.37 (observed)		−8358.97 (observed)			
Logarithm of marginal likelihood of data ($\ln[f(\mathbf{Y}\|\mathcal{M})]$)	–		$-8498.46_{-8499.21}^{-8494.22}$		$-8437.07_{-8440.02}^{-8424.77}$			
Goodness-of-fit p-value	–		0.255		0.222			
Maximum of the potential scale reduction factors (PSRF) [d]	–		1.00302		1.00060			
Multivariate potential scale reduction factor (MPSRF) [d]	–		1.00325		1.00067			
Number of available observations	accidents = fatalities + injuries + PDOs : 19, 094 = 143 + 3, 369 + 15, 582							

The superscript and subscript numbers to the right of individual parameter estimates are 95% confidence/credible intervals.

[a] Standard (conventional) multinomial logit (MNL) model estimated by maximum likelihood estimation (MLE).

[b] Standard multinomial logit (MNL) model estimated by Markov Chain Monte Carlo (MCMC) simulations.

[c] Two-state Markov switching multinomial logit (MSMNL) model estimated by Markov Chain Monte Carlo (MCMC) simulations.

[d] PSRF/MPSRF are calculated separately/jointly for all continuous model parameters. PSRF and MPSRF are close to 1 for converged MCMC chains.

Table 2
Estimation results for multinomial logit models of severity outcomes of one-vehicle accidents on Indiana US routes.

Variable	MNL-by-MLE [a]		MNL-by-MCMC [b]		MSMNL [c]			
	Fatality	Injury	Fatality	Injury	State $s=0$		State $s=1$	
					Fatality	Injury	Fatality	Injury
Intercept (constant term)	$-6.51_{-8.03}^{-5.00}$	$-2.13_{-2.47}^{-1.79}$	$-6.62_{-8.14}^{-5.16}$	$-2.12_{-2.47}^{-1.78}$	$-5.72_{-6.92}^{-4.69}$	$-2.05_{-2.40}^{-1.71}$	$-5.72_{-6.92}^{-4.69}$	$-2.79_{-3.23}^{-2.37}$
Summer season, June–August (1 if yes, 0 if no)	$.514_{.134}^{.894}$	$.200_{.0947}^{.305}$	$.509_{.124}^{.883}$	$.200_{.0951}^{.305}$	$.190_{.0789}^{.300}$	$.190_{.0789}^{.300}$	$.190_{.0789}^{.300}$	–
Daylight or street lights are lit up if dark (1 if yes, 0 if no)	$-.498_{-.855}^{-.142}$	$.194_{.101}^{.287}$	$-.492_{-.848}^{-.136}$	$.203_{.110}^{.296}$	$-.493_{-.857}^{-.136}$	$.197_{.105}^{.290}$	–	$.197_{.105}^{.290}$
Snowing weather (1 if yes, 0 if no)	$-1.17_{-2.18}^{-.170}$	–	$-1.30_{-2.47}^{-.357}$	–	$-1.10_{-2.27}^{-.151}$	$.165_{.0115}^{.317}$	$-1.10_{-2.27}^{-.151}$	$.165_{.0115}^{.317}$
No roadway junction at the accident location (1 if yes, 0 if no)	$.701_{.149}^{1.25}$	$.217_{.0994}^{.335}$	$.727_{.199}^{1.31}$	$.213_{.0968}^{.331}$	$.787_{.259}^{1.36}$	$.214_{.0965}^{.332}$	$.787_{.259}^{1.36}$	$.214_{.0965}^{.332}$
Roadway is straight (1 if yes, 0 if no)	$-.741_{-1.10}^{-.383}$	$-.295_{-.399}^{-.191}$	$-.739_{-1.09}^{-.377}$	$-.296_{-.399}^{-.192}$	$-7.37_{-1.09}^{-.372}$	$-.294_{-.398}^{-.189}$	$-7.37_{-1.09}^{-.372}$	$-.294_{-.398}^{-.189}$
Primary cause of the accident is environment-related (1 if yes, 0 if no)	$-3.45_{-4.18}^{-2.72}$	$-1.89_{-1.99}^{-1.78}$	$-3.51_{-4.32}^{-2.81}$	$-1.89_{-2.00}^{-1.79}$	$-3.59_{-4.40}^{-2.89}$	$-2.09_{-2.24}^{-1.96}$	$-3.59_{-4.40}^{-2.89}$	$-.701_{-1.16}^{-.263}$
Help arrived in 10 min or less after the crash (1 if yes, 0 if no)	$.594_{.507}^{.681}$	$.594_{.507}^{.681}$	$.562_{.475}^{.650}$	$.562_{.475}^{.650}$	$.560_{.472}^{.648}$	$.560_{.472}^{.648}$	$.560_{.472}^{.648}$	$.560_{.472}^{.648}$
The vehicle at fault is a motorcycle (1 if yes, 0 if no)	$2.62_{1.78}^{3.47}$	$3.20_{2.86}^{3.55}$	$2.57_{1.65}^{3.38}$	$3.21_{2.87}^{3.56}$	$3.22_{2.88}^{3.58}$	$3.22_{2.88}^{3.58}$	$3.22_{2.88}^{3.58}$	$3.22_{2.88}^{3.58}$
Age of the vehicle at fault (in years)	$.0363_{.0283}^{.0444}$	$.0363_{.0283}^{.0444}$	$.0367_{.0287}^{.0448}$	$.0367_{.0287}^{.0448}$	–	$.0366_{.0285}^{.0447}$	–	$.0366_{.0285}^{.0447}$
Speed limit (used if known and the same for all vehicles involved)	$.0363_{.00950}^{.0631}$	$.0121_{.00640}^{.0178}$	$.0373_{.0117}^{.0643}$	$.0118_{.00616}^{.0176}$	$.0285_{.0104}^{.0495}$	$.0102_{.00635}^{.0178}$	–	$.0120_{.00635}^{.0178}$
Roadway traveled by the vehicle at fault is two-lane and one-way (1 if yes, 0 if no)	$-.216_{-.391}^{.0417}$	$-.216_{-.391}^{.0417}$	$-.223_{-.398}^{.0517}$	$-.223_{-.398}^{.0517}$	$-.224_{-.401}^{.0504}$	$-.224_{-.401}^{.0504}$	$-.224_{-.401}^{.0504}$	$-.224_{-.401}^{.0504}$
At least one of the vehicles involved was on fire (1 if yes, 0 if no)	$1.19_{.439}^{1.94}$	–	$1.13_{.315}^{1.85}$	–	$1.27_{.452}^{1.98}$	–	$1.27_{.452}^{1.98}$	–
Age of the driver at fault (in years)	$.0114_{.00150}^{.0213}$	–	$.0113_{.00137}^{.0211}$	–	$.0101_{.0000542}^{.0200}$	–	–	–
Weekday (Monday through Friday) (1 if yes, 0 if no)	–	$-.104_{-.196}^{.0116}$	–	$-.104_{-.196}^{.0124}$	–	$-.125_{-.227}^{.0242}$	–	–
Gender of the driver at fault (1 if female, 0 if male)	–	$.272_{.183}^{.362}$	–	$.276_{.186}^{.365}$	–	$.280_{.190}^{.369}$	–	$.280_{.190}^{.369}$
Probability of severity outcome [$P_{t,n}^{(i)}$ given by Eq. (3)], averaged over all values of explanatory variables $\mathbf{x}_{t,n}$	.00766	.183	.00747	.179	.00823	.183	.00218	.158
Markov transition probability of jump $0 \to 1$ ($p_{0\to 1}$)	–		–		$.0767_{.0269}^{.157}$			
Markov transition probability of jump $1 \to 0$ ($p_{1\to 0}$)	–		–		$.613_{.337}^{.864}$			
Unconditional probabilities of states 0 and 1 ($\bar{p}_0$ and $\bar{p}_1$)	–		–		$.887_{.770}^{.959}$ and $.113_{.0409}^{.230}$			
[5pt] Total number of free model parameters ($\beta - s$)	24		24		25			
Posterior average of the log-likelihood (LL)	–		$-7406.39_{-7414.03}^{-7400.61}$		$-7349.06_{-7364.47}^{-7335.46}$			
Max(LL): estimated max. log-likelihood (LL) for MLE; maximum observed value of LL for Bayesian–MCMC	−7384.05 (MLE)		−7396.37 (observed)		−7318.21 (observed)			
Logarithm of marginal likelihood of data ($\ln[f(\mathbf{Y}\|\mathcal{M})]$)	–		$-7417.98_{-7420.23}^{-7413.72}$		$-7377.49_{-7380.00}^{-7369.62}$			
Goodness-of-fit p-value	–		0.337		0.255			
Maximum of the potential scale reduction factors (PSRF) [d]	–		1.00319		1.00073			
Multivariate potential scale reduction factor (MPSRF) [d]	–		1.00376		1.00085			
Number of available observations	accidents = fatalities + injuries + PDOs : 17, 797 = 138 + 3, 184 + 14, 485							

The superscript and subscript numbers to the right of individual parameter estimates are 95% confidence/credible intervals.

[a] Standard (conventional) multinomial logit (MNL) model estimated by maximum likelihood estimation (MLE).

[b] Standard multinomial logit (MNL) model estimated by Markov Chain Monte Carlo (MCMC) simulations.

[c] Two-state Markov switching multinomial logit (MSMNL) model estimated by Markov Chain Monte Carlo (MCMC) simulations.

[d] PSRF/MPSRF are calculated separately/jointly for all continuous model parameters. PSRF and MPSRF are close to 1 for converged MCMC chains.

Table 3
Estimation results for multinomial logit models of severity outcomes of one-vehicle accidents on Indiana state routes.

Variable	MNL-by-MLE [a]		MNL-by-MCMC [b]		MSMNL [c]			
					State $s=0$		State $s=1$	
	Fatality	Injury	Fatality	Injury	Fatality	Injury	Fatality	Injury
Intercept (constant term)	$-3.98^{-3.66}_{-4.30}$	$-1.67^{-1.53}_{-1.80}$	$-4.03^{-3.71}_{-4.36}$	$-1.71^{-1.58}_{-1.85}$	$-3.44^{-3.10}_{-3.79}$	$-1.68^{-1.54}_{-1.81}$	$-4.96^{-4.15}_{-5.96}$	$-1.68^{-1.54}_{-1.81}$
Summer season, June–August (1 if yes, 0 if no)	$.232^{.307}_{.156}$	$.232^{.307}_{.156}$	$.232^{.307}_{.157}$	$.232^{.307}_{.157}$	$.238^{.314}_{.163}$	$.238^{.314}_{.163}$	$.238^{.314}_{.163}$	$.238^{.314}_{.163}$
Roadway type (1 if urban, 0 if rural)	$-.390^{-.302}_{-.478}$	$-.390^{-.302}_{-.478}$	$-.395^{-.306}_{-.483}$	$-.395^{-.306}_{-.483}$	–	$-.385^{-.296}_{-.474}$	$-2.05^{-.954}_{-3.62}$	$-3.85^{-.296}_{-.474}$
Daylight or street lights are lit up if dark (1 if yes, 0 if no)	$-.646^{-.408}_{-.884}$	$.193^{.261}_{.125}$	$-.641^{-.404}_{-.879}$	$.199^{.267}_{.132}$	$-.689^{-.448}_{-.931}$	–	$-.689^{-.448}_{-.931}$	$.277^{.378}_{.177}$
Precipitation: rain/freezing rain/snow/sleet/hail (1 if yes, 0 if no)	$-.854^{.466}_{-1.24}$	–	$-.868^{-.494}_{-1.27}$	–	$-.829^{-.448}_{-1.24}$	–	$-.829^{-.448}_{-1.24}$	–
Roadway median is drivable (1 if yes, 0 if no)	$-.583^{-.225}_{-.940}$	–	$-.596^{-.250}_{-.964}$	–	$-.589^{-.241}_{-.960}$	–	$-.589^{-.241}_{-.960}$	–
Roadway is straight (1 if yes, 0 if no)	$-.284^{-.214}_{-.353}$	$-.284^{-.214}_{-.353}$	$-.283^{-.214}_{-.352}$	$-.283^{-.214}_{-.352}$	$-.117^{-.0184}_{-.214}$	$-.117^{-.0184}_{-.214}$	$-.117^{-.0184}_{-.214}$	$-.465^{-.360}_{-.573}$
Primary cause of the accident is environment-related (1 if yes, 0 if no)	$-4.23^{-3.59}_{-4.86}$	$-1.83^{-1.76}_{-1.91}$	$-4.28^{-3.67}_{-4.97}$	$-1.84^{-1.76}_{-1.91}$	$-4.40^{-3.79}_{-5.10}$	$-2.30^{-2.16}_{-2.44}$	$-4.40^{-3.79}_{-5.10}$	$-1.41^{-1.26}_{-1.55}$
Help arrived in 20 min or less after the crash (1 if yes, 0 if no)	$.840^{.917}_{.762}$	$.840^{.917}_{.762}$	$.863^{.945}_{.781}$	$.863^{.945}_{.781}$	–	$.861^{.944}_{.778}$	$1.64^{2.64}_{.856}$	$.861^{.944}_{.778}$
The vehicle at fault is a motorcycle (1 if yes, 0 if no)	$3.10^{3.31}_{2.89}$	$3.10^{3.31}_{2.89}$	$3.10^{3.31}_{2.89}$	$3.10^{3.31}_{2.89}$	$3.37^{3.66}_{3.09}$	$3.37^{3.66}_{3.09}$	$3.37^{3.66}_{3.09}$	$2.82^{3.19}_{2.47}$
Number of occupants in the vehicle at fault	$.0557^{.0850}_{.0265}$	$.0557^{.0850}_{.0265}$	$.0565^{.0858}_{.0276}$	$.0565^{.0858}_{.0276}$	$.0942^{.138}_{.0528}$	$.0942^{.138}_{.0528}$	$.0942^{.138}_{.0528}$	–
At least one of the vehicles involved was on fire (1 if yes, 0 if no)	$1.90^{2.45}_{1.33}$	$.456^{.780}_{.133}$	$1.87^{2.42}_{1.28}$	$.447^{.768}_{.124}$	$1.87^{2.43}_{1.28}$	$.461^{.782}_{.137}$	$1.87^{2.43}_{1.28}$	$.461^{.782}_{.137}$
Age of the driver at fault (in years)	$14.6^{21.4}_{7.80}\times 10^{-3}$	$-2.80^{-.800}_{-4.70}\times 10^{-3}$	$14.5^{21.3}_{7.67}\times 10^{-3}$	$-2.71^{-.723}_{-4.69}\times 10^{-3}$	$14.5^{21.4}_{7.63}\times 10^{-3}$	$-2.46^{-.469}_{-4.44}\times 10^{-3}$	$14.5^{21.4}_{7.63}\times 10^{-3}$	$-2.46^{-.469}_{-4.44}\times 10^{-3}$
Gender of the driver at fault (1 if female, 0 if male)	$-.496^{-.211}_{-.780}$	$.279^{.344}_{.214}$	$-.505^{-.225}_{-.794}$	$.278^{.343}_{.213}$	$-.473^{-.192}_{-.764}$	$.283^{.348}_{.218}$	$-.473^{-.192}_{-.764}$	$.283^{.348}_{.218}$
Age of the vehicle at fault (in years)	–	$.0334^{.0392}_{.0276}$	–	$.0335^{.0393}_{.0277}$	–	$.0332^{.0390}_{.0274}$	–	$.0332^{.0390}_{.0274}$
license state of the vehicle at fault is a U.S. state except Indiana and its neighboring states (IL, KY, OH, MI)" indicator variable	–	$-.449^{-.217}_{-.681}$	–	$-.444^{-.217}_{-.679}$	–	$-.436^{-.208}_{-.671}$	–	$-.436^{-.208}_{-.671}$
Probability of severity outcome [$P^{(i)}_{t,n}$ given by Eq. (3)], averaged over all values of explanatory variables $\mathbf{x}_{t,n}$	.00912	.182	.00890	.179	.00951	.180	.00804	.179
Markov transition probability of jump $0\rightarrow 1$ ($p_{0\rightarrow 1}$)	–		–		$.335^{.465}_{.216}$			
Markov transition probability of jump $1\rightarrow 0$ ($p_{1\rightarrow 0}$)	–		–		$.450^{.610}_{.313}$			
Unconditional probabilities of states 0 and 1 ($\bar{p}_0$ and $\bar{p}_1$)	–		–		$.574^{.681}_{.504}$ and $.426^{.496}_{.319}$			
Total number of free model parameters ($\beta - s$)	22		22		28			
Posterior average of the log-likelihood (LL)	–		$-13867.40^{-13861.92}_{-13874.73}$		$-13781.76^{-13765.02}_{-13800.89}$			
Max(LL): estimated max. log-likelihood (LL) for MLE; maximum observed value of LL for Bayesian–MCMC	−13846.60 (MLE)		−13858.00 (observed)		−13745.61 (observed)			
Logarithm of marginal likelihood of data ($\ln[f(\mathbf{Y}\vert\mathcal{M})]$)	–		$-13877.89^{-13874.24}_{-13880.38}$		$-13820.20^{-13808.85}_{-13821.73}$			
Goodness-of-fit p-value	–		0.515		0.445			
Maximum of the potential scale reduction factors (PSRF) [d]	–		1.00027		1.00029			
Multivariate potential scale reduction factor (MPSRF) [d]	–		1.00041		1.00045			
Number of available observations	accidents = fatalities + injuries + PDOs : 33, 528 = 302 + 6, 018 + 27, 208							

The superscript and subscript numbers to the right of individual parameter estimates are 95% confidence/credible intervals.

[a] Standard (conventional) multinomial logit (MNL) model estimated by maximum likelihood estimation (MLE).

[b] Standard multinomial logit (MNL) model estimated by Markov Chain Monte Carlo (MCMC) simulations.

[c] Two-state Markov switching multinomial logit (MSMNL) model estimated by Markov Chain Monte Carlo (MCMC) simulations.

[d] PSRF/MPSRF are calculated separately/jointly for all continuous model parameters. PSRF and MPSRF are close to 1 for converged MCMC chains.

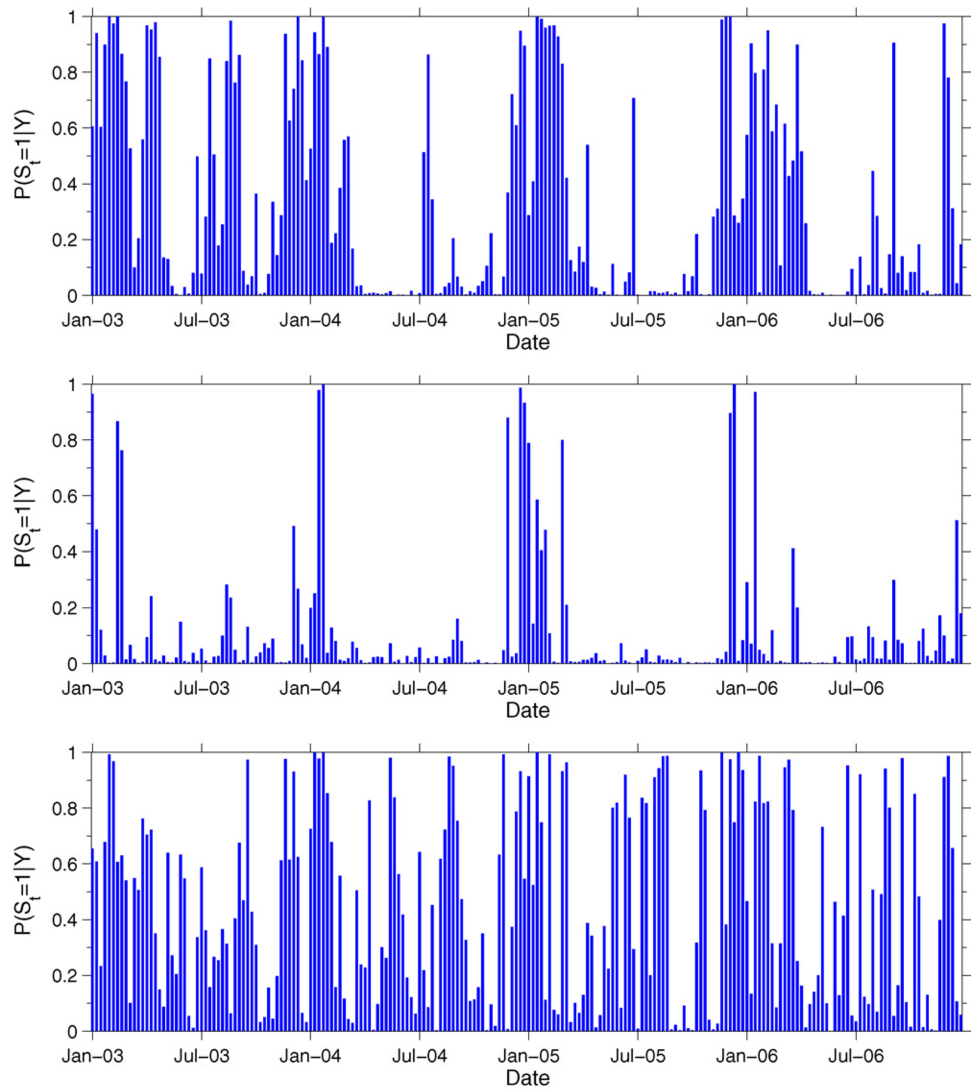

Fig. 1. Weekly posterior probabilities $P(s_t = 1|\mathbf{Y})$ for the MSMNL models estimated for severity of 1-vehicle accidents on interstate highways (top plot), US routes (middle plot) and state routes (bottom plot).

outcomes of 1-vehicle accidents occurring on interstate highways, US routes and state routes are given in Tables 1–3 . All continuous model parameters ($\beta - s$, $p_{0\to1}$ and $p_{1\to0}$) are given together with their 95% confidence intervals (if MLE) or 95% credible intervals (if Bayesian–MCMC), refer to the superscript and subscript numbers adjacent to parameter estimates in Tables 1–3.[15] Because of space limitations, in this paper we do not report estimation results for severity of 1-vehicle accidents on county roads and other roads, and for severity of 2-vehicle accidents [for the unreported results see Malyshkina (2008)].

The top, middle and bottom plots in Fig. 1 show weekly posterior probabilities $P(s_t = 1|\mathbf{Y})$ of the less frequent state $s_t = 1$ for the MSMNL models estimated for severity of 1-vehicle accidents occurring on interstate highways, US routes and state routes, respectively.[16]

We find that in all cases when the two states and Markov switching multinomial logit (MSMNL) models exist, these models are strongly favored by the empirical data over the corresponding standard multinomial logit (MNL) models. Indeed, from Tables 1–3 we see that the MSMNL models provide considerable improvements of the logarithm of the marginal likelihood (ranging from 40.5 to 61.4) as compared to the corresponding MNL models.[17] Thus, from Eq. (6) we find that, given the accident-severity data, the posterior probabilities of the MSMNL models are larger than the probabilities of the corresponding MNL models by factors ranging from $e^{40.5}$ to $e^{61.4}$. In the cases of 1-vehicle accidents on county roads, other roads and the case of 2-vehicle accidents on “other roads”, MSMNL models (not reported here) are also strongly favored by the empirical data over the corresponding MNL models (Malyshkina, 2008).

Let us now consider the maximum likelihood estimation (MLE) of the standard MNL models and a hypothetical MLE estimation of the MSMNL models. We find that, in this hypothetical case,

[15] Note that MLE assumes asymptotic normality of the estimates, resulting in confidence intervals being symmetric around the means (a 95% confidence interval is ±1.96 standard deviations around the mean). In contrast, Bayesian estimation does not require this assumption, and posterior distributions of parameters and Bayesian credible intervals are usually non-symmetric.

[16] Note that these posterior probabilities are equal to the posterior expectations of s_t, $P(s_t = 1|\mathbf{Y}) = 1 \times P(s_t = 1|\mathbf{Y}) + 0 \times P(s_t = 0|\mathbf{Y}) = E(s_t|\mathbf{Y})$.

[17] We use the harmonic mean formula to calculate the values and the 95% confidence intervals of the log-marginal-likelihoods given in Tables 1–3. The confidence intervals are calculated by bootstrap simulations. For details, see Malyshkina et al. (2009) or Malyshkina (2008).

Table 4
Correlations of the posterior probabilities $P(s_t = 1|\mathbf{Y})$ with each other and with weather-condition variables (for the MSMNL model).

	1-Vehicle, interstates	1-Vehicle, US routes	1-Vehicle, state routes	1-Vehicle, county roads	1-Vehicle, other roads	2-Vehicle, other roads
1-Vehicle, interstates	1	0.418	0.293	0.606	−0.013	−0.173
1-Vehicle, US routes	0.418	1	0.263	0.688	−0.070	−0.155
1-Vehicle, state routes	0.293	0.263	1	0.409	−0.047	−0.035
1-Vehicle, county roads	0.606	0.688	0.409	1	−0.022	−0.051
1-Vehicle, other roads	−0.013	−0.070	−0.047	−0.022	1	0.115
2-Vehicle, other roads	−0.173	−0.155	−0.035	−0.051	0.115	1
All year						
Precipitation (in.)	−0.139	−0.060	0.096	−0.037	0.067	0.146
Temperature (°F)	−0.606	−0.439	−0.234	−0.665	0.231	0.220
Snowfall (in.)	0.479	0.635	0.319	0.723	0.003	−0.100
> 0.0 (dummy)	0.695	0.412	0.382	0.695	−0.142	−0.131
> 0.1 (dummy)	0.532	0.585	0.328	0.847	−0.046	−0.161
Wind gust (mph)	0.108	0.100	0.087	0.206	0.164	0.051
Fog/frost (dummy)	0.093	0.164	0.193	0.167	0.047	0.119
Visibility distance (mile)	−0.228	−0.221	−0.172	−0.298	−0.019	−0.081
Winter (November–March)						
Precipitation (in.)	−0.134	−0.037	0.027	−0.053	0.065	0.356
Temperature (°F)	−0.595	−0.479	−0.397	−0.735	−0.008	0.236
Snowfall (in.)	0.439	0.592	0.375	0.645	0.157	−0.110
> 0.0 (dummy)	0.596	0.282	0.475	0.607	0.115	−0.142
> 0.1 (dummy)	0.445	0.518	0.370	0.789	0.112	−0.210
Wind gust (mph)	0.302	0.134	0.122	0.353	0.237	0.071
Frost (dummy)	0.537	0.544	0.440	0.716	0.052	−0.225
Visibility distance (mile)	−0.251	−.304	−0.249	−0.380	−0.155	−0.109
Summer (May–September)						
Precipitation (in.)	0.000	0.006	0.259	0.096	0.047	−0.063
Temperature (°F)	0.179	0.149	0.113	0.037	0.062	0.155
Snowfall (in.)	–	–	–	–	–	–
> 0.0 (dummy)	–	–	–	–	–	–
> 0.1 (dummy)	–	–	–	–	–	–
Wind gust (mph)	−0.126	−.009	0.164	0.029	0.121	0.034
Fog (dummy)	0.203	0.193	0.275	0.101	−0.076	−0.011
Visibility distance (mile)	−0.139	−0.124	−0.062	−0.009	0.077	−0.094

a classical statistics approach for model comparison, based on the MLE, would also favors the MSMNL models over the standard MNL models. For example, refer to line "max (LL)" in Table 1 given for the case of 1-vehicle accidents on interstate highways. The MLE gave the maximum log-likelihood value −8465.79 for the standard MNL model. The maximum log-likelihood value observed during our MCMC simulations for the MSMNL model is equal to −8358.97. A hypothetical MLE, at its convergence, would give a MSMNL log-likelihood value that would be even larger than this observed value. Therefore, if estimated by the MLE, the MSMNL model would provide large, at least 106.82 improvement in the maximum log-likelihood value over the corresponding MNL model. This improvement would come with only modest increase in the number of free continuous model parameters ($\beta - s$) that enter the likelihood function (refer to Table 1). Similar arguments hold for comparison of MSMNL and MNL models estimated for other roadway-class-accident-type combinations (see Tables 2 and 3).

To evaluate the goodness-of-fit for a model, we use the posterior (or MLE) estimates of all continuous model parameters ($\beta - s$, α, $p_{0\to1}$, $p_{1\to0}$) and generate 10^4 artificial data sets under the hypothesis that the model is true.[18] We find the distribution of χ^2 and calculate the goodness-of-fit p-value for the observed value of χ^2. For details, see Malyshkina (2008). The resulting p-values for our models are given in Tables 1–3. These p-values are around 20–50%. Therefore, all models fit the data well.

To get some additional insight into the estimation, consider Table 4. The first six rows of this table list time-correlation coefficients between posterior probabilities $P(s_t = 1|\mathbf{Y})$ for the six MSMNL models that exist and are estimated for six roadway-class-accident-type combinations (1-vehicle accidents on interstate highways, US routes, state routes, county roads, other roads, and 2-vehicle accidents on other roads).[19] We see that the states for 1-vehicle accidents on all high-speed roads (interstate highways, US routes, state routes and county roads) are correlated with each

[18] Note that the state values **S** are generated by using $p_{0\to1}$ and $p_{1\to0}$.

[19] Here and below we calculate weighted correlation coefficients. For variable $P(s_t = 1|\mathbf{Y}) \equiv E(s_t|\mathbf{Y})$ we use weights w_t inversely proportional to the posterior standard deviations of s_t. That is $w_t \propto \min\{1/\mathrm{std}(s_t|\mathbf{Y}), \mathrm{median}[1/\mathrm{std}(s_t|\mathbf{Y})]\}$.

other. The values of the corresponding correlation coefficients are positive and range from 0.263 to 0.688 (see Table 4). This result suggests an existence of common (unobservable) factors that can cause switching between states of roadway safety for 1-vehicle accidents on all high-speed roads.

The remaining rows of Table 4 show correlation coefficients between posterior probabilities $P(s_t = 1|\mathbf{Y})$ and weather-condition variables. These correlations were found by using daily and hourly historical weather data in Indiana, available at the Indiana State Climate Office at Purdue University (www.agry.purdue.edu/climate). For these correlations, the precipitation and snowfall amounts are daily amounts in inches averaged over the week and across Indiana weather observation stations.[20] The temperature variable is the mean daily air temperature (°F) averaged over the week and across the weather stations. The wind gust variable is the maximal instantaneous wind speed (mph) measured during the 10-min period just prior to the observational time. Wind gusts are measured every hour and averaged over the week and across the weather stations. The effect of fog/frost is captured by a dummy variable that is equal to one if and only if the difference between air and dewpoint temperatures does not exceed 5 °F (in this case frost can form if the dewpoint is below the freezing point 32 °F, and fog can form otherwise). The fog/frost dummies are calculated for every hour and are averaged over the week and across the weather stations. Finally, visibility distance variable is the harmonic mean of hourly visibility distances, which are measured in miles every hour and are averaged over the week and across the weather stations.[21]

From the results given in Table 4 we find that for 1-vehicle accidents on all high-speed roads (interstate highways, US routes, state routes and county roads), the less frequent state $s_t = 1$ is positively correlated with extreme temperatures (low during winter and high during summer), rain precipitations and snowfalls, strong wind gusts, fogs and frosts, low visibility distances. It is reasonable to expect that roadway safety varies by weather conditions, and that weather could be one of the key factors generating the two-state nature of roadway safety.

The results of Table 4 suggest that Markov switching for road safety on "other roads" is very different from switching on all other roadway classes. In particular, the states of roadway safety on "other roads" exhibit low correlation with states on roads of other classifications. In addition, only "other roads" exhibit Markov switching in the case of 2-vehicle accidents. Finally, states of roadway safety on "other roads" show little correlation with weather conditions. A possible explanation of these differences is that "other roads" are mostly located in urban areas and they have traffic moving at speeds lower that those on other roadway classifications (there may be other important characteristics that distinguish "other roads", such as roadway geometry, presence of at-grade intersections, and presence of pedestrians and bicyclists).

Now, we consider the estimation results for the stationary unconditional probabilities $\bar{p}_0$ and $\bar{p}_1$ of states $s_t = 0$ and $s_t = 1$ for MSMNL models (see Section 2). In the cases of 1-vehicle accidents on interstate highways, US routes and state routes these transition probabilities are listed in Tables 1–3. In the cases of 1-vehicle accidents on county roads and 1- and 2-vehicle accidents on "other roads", these values are available online, see Malyshkina (2008). We find that the ratio $\bar{p}_1/\bar{p}_0$ is approximately equal to 0.46, 0.13, 0.74, 0.25, 0.65 and 0.36 in the cases of 1-vehicle accidents on interstate highways, US routes, state routes, county roads, other roads, and 2-vehicle accidents on other roads, respectively. Thus for some roadway-class-accident-type combinations (for example, 1-vehicle accidents on US routes) the less frequent state $s_t = 1$ is quite rare, while for other combinations (for example, 1-vehicle accidents on state routes) state $s_t = 1$ is only slightly less frequent than state $s_t = 0$.

Next, we set model parameters $(\beta - s)$ to their posterior means (or their MLE estimates), calculate the probabilities of fatality and injury outcomes by using Eq. (3) and average these probabilities over the values of the explanatory variables $\mathbf{x}_{t,n}$ for the observations in the data sample. We compare these probabilities across the two states of roadway safety, $s_t = 0$ and $s_t = 1$, for MSMNL models [refer to Tables 1–3 and to Malyshkina (2008)]. We find that in many cases these averaged probabilities of fatality and injury outcomes do not differ very significantly across the two states of roadway safety (the only significant differences are for fatality probabilities in the cases of 1-vehicle accidents on US routes, county roads and other roads). This means that in many cases states $s_t = 0$ and $s_t = 1$ are approximately equally dangerous as far as accident-severity is concerned. This has important implications in that, although the aggregate severity probabilities are close between the two states, the process generating these probabilities is quite different and counter measures to reduce severity will also be different.

Finally, let us discuss the effects of explanatory variables on accident severities. For brevity, we will focus on accident severities occurred on interstate highways and will consider only those variables that are significantly different between the two states. Table 1 shows that parameter estimates for summer season, construction at the accident location, precipitation, driver-related cause of the accident, at-fault motorcycle, fire, and at-fault female driver dummy variables are all significantly different between the two states. When summer season, construction at the accident location and precipitation are present, they lead to higher injury rates in state $s_t = 1$ as compared to state $s_t = 0$. This may be explained by drivers' over-confidence and underestimation of adverse weather/pavement conditions in state $s_t = 1$. For example, in state $s_t = 0$ at construction sites drivers may exercise caution (i.e., by reducing speed), leading to lower injury rates, but there is no such effect in state $s_t = 1$. When an accident is driver-related, a vehicle at-fault is a motorcycle and at least one of the vehicles involved is on fire, there are higher injury and fatality rates in state $s_t = 0$ as compared to the state $s_t = 1$. These differences are reasonable. For example, we expect that there is less motorcycle traffic during adverse weather conditions.

5. Conclusions

Markov-switching models of severity have the potential to capture unobserved heterogeneity in accident data that could relate to detailed weather conditions and other factors that may not be known to the analyst. In this study we found that Markov switching multinomial logit (MSMNL) models are valid for the severity of 1-vehicle accidents occurring on high-speed roads (interstate highways, US routes, state routes, county roads), but not for 2-vehicle accidents on high-speed roads. One of several possible explanations of this result is the fundamental difference between 1- and 2-vehicle accidents. It may be that multi-vehicle accidents tend to involve more occupants and this "number effect" could affect the severity likelihoods of the most severely injured person. Drivers' behavior might exhibit a multiple-state pattern. In particular, drivers might be overconfident and/or have difficulties in adjustments to adverse weather conditions. Also, the severity of a 2-vehicle accident may be influenced in unobserved ways with regard to the energy dissipation between the two vehicles. As far as slower-speed "other roads" are concerned, in this case both 1- and 2-vehicle accidents exhibit a two-state process for their injury-severity. Further studies are needed to understand these results. However, the

[20] Snowfall and precipitation amounts are weakly related with each other because snow density (g/cm^3) can vary by more than a factor of 10.

[21] The harmonic mean $\bar{d}$ of distances d_n is calculated as $\bar{d}^{-1} = (1/N)\sum_{n=1}^{N} d_n^{-1}$, assuming $d_n = 0.25$ miles if $d_n \le 0.25$ miles.

important result is that in many cases the two-state MSMNL models provide a superior statistical fit for accident-severity outcomes as compared to standard MNL models.

Finally, we found that in many cases states $s_t = 0$ and $s_t = 1$ are approximately equally dangerous as far as overall accident-severity is concerned. This result holds despite the fact that state $s_t = 1$ is correlated with adverse weather conditions. It seems that during adverse weather both the number of serious accidents (fatalities and injuries) and the number of minor accidents (PDOs) increase, so that their proportions stay approximately the same overall (however it should be noted that our results suggest that the processes generating these proportions are quite different as indicated by the different MNL models found in each of the two states). One possible explanation for this finding is that individuals drive more cautiously in adverse weather. Thus, although, using Indiana accident data, Malyshkina et al. (2009) found that the total number of accidents significantly increased during adverse weather in a two-state Markov switching process, it is possible that drivers' risk-compensating behavior in adverse weather may be sufficient to mitigate the effect of weather in terms of increasing injury severities.

References

Abdel-Aty, M., 2003. Analysis of driver injury severity levels at multiple locations using ordered probit models. Journal of Safety Research 34 (5), 597–603.
Carson, J., Mannering, F.L., 2001. The effect of ice warning signs on ice-accident frequencies and severities. Accident Analysis and Prevention 33 (1), 99–109.
Chang, L.-Y., Mannering, F.L., 1999. Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. Accident Analysis and Prevention 31 (5), 579–592.
Cowan, G., 1998. Statistical Data Analysis. Clarendon Press, Oxford Univ. Press, USA.
Duncan, C., Khattak, A., Council, F., 1998. Applying the ordered probit model to injury severity in truck-passenger car rear-end collisions. Transportation Research Record 1635, 63–71.
Eluru, N., Bhat, C., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. Accident Analysis and Prevention 39 (5), 1037–1049.
Eluru, N., Bhat, C., Hensher, D., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accident Analysis and Prevention 40 (3), 1033–1054.
Kass, R.E., Raftery, A.E., 1995. Bayes factors. Journal of the American Statistical Association 90 (430), 773–795.
Khattak, A., 2001. Injury severity in multi-vehicle rear-end crashes. Transportation Research Record 1746, 59–68.
Khattak, A., Pawlovich, D., Souleyrette, R., Hallmarkand, S., 2002. Factors related to more severe older driver traffic crash injuries. Journal of Transportation Engineering 128 (3), 243–249.
Khorashadi, A., Niemeier, D., Shankar, V., Mannering, F.L., 2005. Differences in rural and urban driver-injury severities in accidents involving large trucks: an exploratory analysis. Accident Analysis and Prevention 37 (5), 910–921.
Kockelman, K., Kweon, Y.-J., 2002. Driver injury severity: an application of ordered probit models. Accident Analysis and Prevention 34 (3), 313–321.
Kweon, Y.-J., Kockelman, K., 2003. Overall injury risk to different drivers: combining exposure, frequency, and severity models. Accident Analysis and Prevention 35 (4), 414–450.
Lee, J., Mannering, F.L., 2002. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. Accident Analysis and Prevention 34 (2), 149–161.
Maher, M.J., Summersgill, I., 1996. A comprehensive methodology for the fitting of predictive accident models. Accident Analysis and Prevention 28 (3), 281–296.
Malyshkina, N.V., 2006. Influence of Speed Limit on Roadway Safety in Indiana. MS Thesis. Purdue University. http://arxiv.org/abs/0803.3436.
Malyshkina, N.V., 2008. Markov Switching Models: An Application to Roadway Safety. PhD Thesis. Purdue University. http://arxiv.org/abs/0808.1448.
Malyshkina, N.V., Mannering, F.L., Tarko, A.P., 2009. Markov switching negative binomial models: an application to vehicle accident frequencies. Accident Analysis and Prevention 41(2), 217–226.
McFadden, D., 1981. Econometric models of probabilistic choice. In: Manski, C.F., McFadden, D. (Eds.), Structure Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge MA.
Milton, J., Shankar, V., Mannering, F.L., 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accident Analysis and Prevention 40 (1), 260–266.
O'Donnell, C., Connor, D., 1996. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. Accident Analysis and Prevention 28 (6), 739–753.
Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical Recipes 3rd edition: The Art of Scientific Computing. Cambridge Univ. Press, UK.
Robert, C.P., 2001. The Bayesian choice: from decision-theoretic foundations to computational implementation. Springer-Verlag, New York.
Savolainen, P., Mannering, F.L., 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Accident Analysis and Prevention 39 (5), 955–963.
Shankar, V., Mannering, F.L., 1996. An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. Journal of Safety Research 27 (3), 183–194.
Shankar, V., Mannering, F.L., Barfield, W., 1996. Statistical analysis of accident severity on rural freeways. Accident Analysis and Prevention 28 (3), 391–401.
Tsay, R.S., 2002. Analysis of Financial Time Series: Financial Econometrics. John Wiley & Sons, Inc.
Ulfarsson, G., Mannering, F.L., 2004. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. Accident Analysis and Prevention 36 (2), 135–147.
Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2003. Statistical and Econometric Methods for Transportation Data Analysis. Chapman & Hall/CRC.
Wood, G.R., 2002. Generalised linear accident models and goodness of fit testing. Accident Analysis and Prevention 34 (4), 417–427.
Yamamoto, T., Shankar, V., 2004. Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects. Accident Analysis and Prevention 36 (5), 869–876.