Transportation Research Part B 67 (2014) 109–128

Contents lists available at ScienceDirect

Transportation Research Part B

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

# The analysis of vehicle crash injury-severity data: A Markov switching approach with road-segment heterogeneity

Yingge Xiong [a,*], Justin L. Tobias [b,1], Fred L. Mannering [a,2]

[a] Purdue University, School of Civil Engineering, 550 Stadium Mall Drive, West Lafayette, IN 47907-2051, United States
[b] Purdue University, Department of Economics, 403 W. State Street, West Lafayette, IN 47907-2056, United States

ARTICLE INFO

*Article history:*
Received 30 September 2013
Received in revised form 19 April 2014
Accepted 23 April 2014

*Keywords:*
Markov switching models
Random parameter models
Ordered response model
Data augmentation
Markov Chain Monte Carlo simulation

ABSTRACT

Time-constant assumptions in discrete-response heterogeneity models can often be violated. To address this, a time-varying heterogeneity approach to model unobserved heterogeneity in ordered response data is considered. A Markov switching random parameters structure (which accounts for heterogeneity across observations) is proposed to accommodate both time-varying and time-constant (cross-sectional) unobserved heterogeneity in an ordered discrete-response probability model. A data augmented Markov Chain Monte Carlo algorithm for non-linear model estimation is developed to facilitate model estimation. The performance of the cross-sectional heterogeneity model and time-varying heterogeneity model are examined with vehicle crash-injury severity data. The time-varying heterogeneity model (Markov switching random parameters ordered probit) is found to provide the best overall model fit. Two roadway safety states are shown to exist and roadway segments transition between these two states according to Markov transition probabilities. The results demonstrate considerable promise for Markov switching models in a wide variety of applications.



## 1. Introduction

Crash injury-severity data are typically reported as discrete data (no-injury, injury, fatality), and these data have been studied with a variety of ordered and unordered discrete outcome models. Because unobserved heterogeneity is likely to exist among the population of crash-involved road users (such as differences in risk-taking behavior and physiological factors), crash-involved vehicles (such as variations in tire wear, braking systems, and crash worthiness), and related roadway/environmental factors (such as differences in road surface conditions, precipitation and fog affecting visibility), allowing for the effects of explanatory variables on crash-injury severities to vary across observations is now a critical methodological consideration in transportation safety analysis (Savolainen et al., 2011; Mannering and Bhat, 2014).

Random parameters models (which account for heterogeneity across observations), such as the mixed logit model and others, have been widely used to address unobserved heterogeneity for discrete outcome data in general (McFadden and Train, 2000) and crash-severity data in particular (Milton et al., 2008;Castro et al., 2013). These random parameters models are well suited to capture cross-sectional heterogeneity, which is really unobserved time-constant heterogeneity because it

* Corresponding author. Tel.: +1 765 430 7898; fax: +1 765 494 0395.
*E-mail addresses:* xiong0@purdue.edu (Y. Xiong), jltobias@purdue.edu (J.L. Tobias), flm@ecn.purdue.edu (F.L. Mannering).
[1] Tel.: +1 765 494 8570; fax: +1 765 496 1778.
[2] Tel.: +1 765 496 7913; fax: +1 765 494 0395.

http://dx.doi.org/10.1016/j.trb.2014.04.007
0191-2615/

is often approximated from the perspective of the short-term time horizon of cross-sectional data. In fact, random parameters and finite mixture (latent class) methods have emerged as the two dominant approaches for characterizing this type of heterogeneity (Bhat, 1999;Srinivasan, 2002; Hensher and Greene, 2003;Green and Hensher, 2003; Depaire et al., 2008;Shen, 2009; Eluru et al., 2012;Xiong and Mannering, 2013).[3]

In addition to time-constant heterogeneity, the possibility of time-varying unobserved heterogeneity can also be of concern in crash-injury data. Potentially unobserved factors such as weather conditions and traffic congestion can vary from period-to-period and substantially impact the severity of roadway accidents. For discrete-outcome data gathered over time, the time-constant heterogeneity assumption can therefore lead to biased parameter estimates (attenuation bias, Hirose, 2012) and erroneous inferences when variability over time is present.

Hidden-state models such as change-point and Markov switching models have become increasingly popular in capturing unobserved heterogeneity across time periods by assuming a data-generation process that is governed by hidden regime transitions. The change-point model is formulated to allow for a number of latent discrete-states that are typically specified to evolve according to a discrete-time, discrete-state Markov process with constrained transition probabilities (Chib, 1998). As examples of this approach, Elliott and Shope (2003) used a Bayesian change-point linear model to investigate graduated driver licensing effects that take into account the uncertainty when these effects begin and end, and whether or not a "rebound" in crash rates occurs afterward. Kobayashi et al. (2012) presented a multiple-stage change-point model to tackle selection bias in monitoring data for predicting the deterioration progress of infrastructure systems.

In contrast, Markov switching models (Hamilton, 1989, 1990) restrict the number of hidden states (two hidden states in most cases) and the process can only move back and forth among these states. Since Chib (1996) proposed a Bayesian Markov Chain Monte Carlo (MCMC) method to sample the hidden states, Markov switching models have become popular in time series data and count data applications (Malyshkina et al., 2009;Malyshkina and Mannering, 2010). With regard to discrete outcome models, Dueker (1999) was among the first to consider Markov-switching heterogeneity in a dynamic ordered probit framework in his study of discrete changes in the bank prime lending rate.[4] In the crash-severity field, Malyshkina and Mannering (2009) hypothesized the presence of two states of highway safety resulting from distinct combinations of driving-environment conditions, driver reactions, and other factors that not only vary across observations but also interact and change over time (resulting in roadway segments changing from one state to another over time, under specified probabilistic conditions). Their subsequent empirical analysis, using a Markov-switching multinomial logit model of crash severity, showed the presence of two states with the less-safe state (the state likely to result in the more severe crash injuries) being highly correlated with adverse climatic conditions relating to precipitation, low temperatures, snow, and visibility. It is likely that the presence (or combined presence) of these less than ideal weather-related conditions (many of which are typically not available for empirical analyses) results in drivers reacting in different ways as they struggle to estimate the change in crash probabilities, and develop strategies to modify their speeds and car-following distances accordingly.[5]

However, previous empirical studies such as Malyshkina and Mannering (2009), only accounted time-varying heterogeneity and not cross-sectional heterogeneity. As such, in each time-dependent state, observations from all road-way segments would be constrained to share the same set of parameters which could be an issue if there is unobserved heterogeneity across observations. Past exceptions of this constraint include the work of Scott et al. (2005) in their Bayesian application of hidden Markov models to longitudinal data for comparing the effectiveness of two antipsychotic medications for schizophrenia. Their model structure allows latent health states and longitudinal effects to be simultaneously estimated. To predict the success of spawning in fish, Holan et al. (2009) derived a Markov switching model with generalized auto-regressive conditional heteroscedastic dynamics in order to provide an eigenvalue predictor to a Bayesian hierarchical linear regression. In other work, Geweke and Amisano (2011) developed a hierarchical Markov normal mixture model for asset return prediction in financial markets. The mixture components were designed to be non-Gaussian and be themselves mixtures of normal distributions, so that they can capture serial correlation as well as conditional heteroscedasticity in financial-return data. Park (2012) developed a change-point panel-data regression model based on reversible jump MCMC methods. The model is formulated to deal with political economy debates by allowing regime-specific parameters which depended on time and observation-specific parameters varied by individuals. The most recent research that has dealt with time-varying unobserved cross-sectional heterogeneity is that of Hirose (2012). There, the author developed a model that allowed observation-specific intercepts to vary within each state based on observation-specific Markov transition probabilities, and validated it through an application of militarized interstate dispute. All of these Markov switching panel modeling

[3] The random parameters approach accommodates individual unobserved heterogeneity by placing a distributional assumption on the parameters of interest, while finite mixture approaches adopt a semiparametric way to impose a discrete distribution assumption which is represented by a specified number of mass points. In order to precisely depict a "true" distribution of unobserved heterogeneity, which may have possible skewness, kurtosis and multimodality, a hybrid formulation that incorporates a random parameters structure with a finite mixture model that can account for group-specific as well as within-group variation has also been undertaken by a number of researchers (Verbeke and Lesaffre, 1996;Allenby et al., 1998; Lenk and DeSarbo, 2000;Yau et al., 2003; Bujosa et al., 2010; Greene and Hensher, 2013; Xiong and Mannering, 2013).

[4] Advances in estimation methods have enabled further exploration of the Markov models. For example, Scott (2002) demonstrated that a forward–backward recursive Gibbs sampler mixes more rapidly than direct Gibbs sampler which was initially used for hidden Markov models. Based on this approach, Chopin (2007) developed a Monte Carlo algorithm for estimating a sequentially ordered hidden Markov models.

[5] Mannering and Bhat (2014) provide further examples of why two states may exist in highway safety analysis, including the increase in the variance of driver speeds under adverse versus normal weather conditions and the possible presence and variation in drivers' risk compensating behavior in response to changes in roadway conditions.