ELSEVIER

Accident Analysis and Prevention 40 (2008) 1023–1032

ACCIDENT ANALYSIS & PREVENTION

www.elsevier.com/locate/aap

Willingness to use safety belt and levels of injury in car accidents

Matthieu de Lapparent *

Institut National de Recherche sur les Transports et leur Securite, Departement Economie et Sociologie des Transports, 2, avenue du General Malleret-Joinville, 94114 Arcueil, France

Received 8 January 2006; received in revised form 19 November 2007; accepted 23 November 2007

Abstract

In this article, we develop a bivariate ordered Probit model to analyze the decision to fasten the safety belt in a car and the resulting severity of accidents if it happens. The approach takes into account the fact that the decision to fasten the safety belt has a direct causal effect on the category of injury if an accident happens. Our application to a sample drawn from the database of French accident reports in 2003 for three populations of car users (drivers, front passengers, rear passengers) shows that fastening the safety belt is significantly related to a decrease in severe injuries but it shows also that these car users compensate partly for this safety benefit. Furthermore, it is observed that demographic characteristics of car users, as well as transport facilities, play important roles in decisions to fasten safety belts and in the eventual resulting accident injuries.

Keywords: Safety belt; Accident severity; Bivariate ordered probit; Risk compensation

1. Introduction

Reports of the European Commission (2001, 2003) distinguished three pillars of road safety analysis: the design of transportation networks and their characteristics (horizontal and vertical profiles of roads, pavements, restraint barriers, etc.), motorized and non-motorized vehicle standards (chassis, active and passive safety devices, engines, weights, etc.) and the behaviour of road users (alcohol impairment, fastening safety belts, wearing helmets, speed, etc.). These pillars interact together to produce accidents and their levels of severity. The purpose to the analysis of these interactions is to achieve sustainable transport, that is transport that fulfils the primary purpose of moving people and goods while simultaneously contributing to environmental, economic and social sustainability by avoiding the occurrence of too many or too severe accidents.

When the effects of wearing a safety belt on the level of accident severity are examined, the decision to fasten the safety belt is almost never explained but it is rather used as an exogenous variable or a control variable, among many others, to explain the differences in the distributions of accident severity between and within different populations of road users, see for instance Abdel-Aty (2003), Al-Ghamdi (2002), Huelke and Compton (1995), Kockelman and Kweon (2002), Srinivasan (2002), Wang and Kockelman (2005). Although there are many stylized facts as well as many theoretical and empirical studies about the effects of the different aspects of road safety decisions on the severity of accidents, there are few econometric models that make them endogenous simultaneously, to wit: O'Donnell and Connor (1996), Eluru and Bhat (2007).

The model to be presented in this article is a bivariate ordered Probit model.[1] It is defined as a hierarchical system of two equations, in which the safety belt choice is modelled simultaneously with injury severity. The relationship between these variables is recursive: the choice of the seat belt drives the injury severity but the injury severity does not enter the seat belt choice. The decision to fasten the safety belt is assumed to have a direct causal effect on the level of injury if an accident happens. The model structure has also an important additional dimension: it makes the error terms of the seat belt choice and the injury severity relationships correlated, which makes it possible to detect if people overcompensate for safety, among other things. However, as highlighted by Eluru and Bhat (2007), even if the use of ordered response models for analysis of accident severity is accepted, it should be noticed that such models do impose a

* Corresponding author. Tel.: +33 1 47 40 72 63.
E-mail address: lapparent@inrets.fr.

[1] The reader should note that ordered probit models that are defined for K outcomes amounts to estimating K standard binary probit models with the constraint that their coefficients, but not their intercepts, are equal.

0001-4575/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2007.11.005

monotonic restriction on the effect of exogenous variables, a potential limitation to be kept in mind.

As the contribution of Eluru and Bhat (2007) makes an important step towards the integration of traveler's behavior and resulting accident severity, it seems necessary to compare the model proposed in this article to theirs. The statistical model they have proposed relies on a joint binary logit-ordered logit structure with random coefficients and latent variables. The latter identify the source of unobserved heterogeneity and create correlation between the observed endogenous outcomes. These authors have also accounted for effects on injury severity of safety belt use differring according to accident configuration by interacting the safety belt variable with variables that describe the type of accident in the injury severity equation.

By contrast, our model, in addition to lacking such interaction terms, cannot identify sources of unobserved correlation between the endogenous variables. To that extent, our proposed approach neglects some refinements pertaining to the variability of the effects of some explanatory variables on the observed outcomes, having neither random parameters (model parameters stay the same accross individuals) nor latent (unobserved) explanatory variables.

In our more restrictive framework of analysis, each dimension of the problem is modeled conventionally, as we provide an average representation of the linkages between safety belt use and its observable determinants, as well as of accident severity and its observable determinants (including the link between safety belt use and accident severity). We make no room for deviations that might arise due to accident configurations or latent behavior factors such as driver levels of concentration, anticipations and reactions, etc.

The rest of the article is organized as follows. Section 2 of this article presents the statistical formulation of our model while Section 3 presents data drawn from the French accident report database of year 2003: descriptive statistics of the observed endogenous variables and the selected exogenous variables are discussed. Section 4 discusses the results of the estimated model. The last section draws conclusions about model limit, possible extensions of the approach, and future research tracks.

2. The model

It is postulated that a car user, either a driver or a passenger, is incited to think about fastening his/her safety belt for various reasons both due to education and because fastening the safety belt is mandatory and the law provides for sanctions. Nevertheless, some car users do ride without their safety belts fastened, either because they forget to do it them and/or because they consciously take the decision of not doing it (they disobey traffic laws).

In the present approach, the model assumes that individual has a willingness to fasten his/her safety belt y_1^*. It is modeled as the weighted sum of explanatory variables x_1 plus an error term ε_1:

$$y_1^* = x_1'\delta_1 + \varepsilon_1. \tag{1}$$

The vector of weights δ_1 of the vector of the exogenous variables x_1 (x_1' is the transpose of x_1) measure the willingness to fasten the safety belt. The error term ε_1 summarizes all the other factors that are not observable and/or measurable and that play roles on the willingness to fasten the safety belt. What is observed when an accident happens is not the willingness to fasten the safety belt but wether the safety belt was effectively fastened or not. It may be hypothetized that individuals fasten their belt if, and only if, this willingness crosses a threshold $\bar{y}$. The observed dichotomous variable y_1 is related to y_1^* with the following observation rule:

$$y_1 = \begin{cases} 0 & \text{if } y_1^* \le \bar{y} \\ 1 & \text{if } y_1^* > \bar{y} \end{cases}. \tag{2}$$

The level of injury is a latent variable y_2^*. It is modeled as the weighted sum of explanatory variables stacked in a vector x_2, plus the contribution of the safety belt if fastened, and an error term ε_2:

$$y_2^* = x_2'\delta_2 + \alpha \mathbb{I}\left(y_1^* > 0\right) + \varepsilon_2. \tag{3}$$

$\mathbb{I}(a)$ is a boolean operator that takes the value 1 if a is true or the value 0 if a is false.

A safety belt is a harness designed to hold the occupant of a car or other vehicle in place if a collision happens or, more commonly, if it stops suddenly. Safety belts are designed to distribute the force of a crash over the strongest areas of the body: the bones of your hips, chest and shoulders. Safety belts are intended to reduce injuries by stopping the wearer from hitting hard interior elements of the vehicle or from being thrown out from the vehicle. Cars safety belts also prevent rear-seat passengers from crashing into those in the front seats. Fastening the safety belt should cause the latent level of injury to diminish, therefore α is expected to be negative. It models how much is diminishing the latent level of injury when fastening the safety belt. The vector of weights δ_2 models the total effects of the vector of the exogenous variables x_2 on the level of injury, and ε_2 models all the factor that are unobservable and/or measurable and that play roles on the level of injury. When an accident happens, one does not observe the "true" level of injury but rather a membership of one of some ordered categories of injury. They are usually defined from the lowest latent level of injury (no injury, material damages only) to the highest latent level of injury (fatal injury). The variable y_2 that is observed for K classes of injury is related to the latent level of injury y_2^* by the following rule of observation:

$$y_2 = \begin{cases} 1 & \text{if } \beta_0 < y_2^* \le \beta_1 \\ 2 & \text{if } \beta_1 < y_2^* \le \beta_2 \\ \vdots & \\ K & \text{if } \beta_{K-1} < y_2^* \le \beta_K \end{cases}, \tag{4}$$

where $\beta_0 \equiv -\infty$ and $\beta_K \equiv +\infty$ by convention.

The approach assumes that the error terms of Eqs. (1) and (3) are normally distributed:

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \rightsquigarrow \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Omega = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_2^2 \end{bmatrix}\right), \sigma_{1,2} = \sigma_{2,1}. \quad (5)$$

The error variables of the two equations are likely to be correlated because the observed outcomes might be explained by common unobservable factors. When the sign of the covariance is positive, it means that the car user take more risk on the road when his/her safety belt is fastened (risk compensation). When the sign of the covariance is negative, it means that the car user takes less risk on the road when his/her safety belt is fastened. The sign and the magnitude of the associated correlation model either an increase of risk taking on the road or a decrease of risk taking on the road.

By definition, the covariance matrix Ω is symmetric and positive definite. It admits a Cholesky decomposition Γ defined by $\Gamma\Gamma' = \Omega$ such that

$$\Gamma = \begin{bmatrix} \sigma_1 & 0 \\ \dfrac{\sigma_{2,1}}{\sigma_1} & \sigma_2\sqrt{1-\rho^2} \end{bmatrix}, \rho = \frac{\sigma_{2,1}}{\sigma_2\sigma_1} \in [0, 1]. \quad (6)$$

The Cholesky decomposition is applied to a vector of standardized normal random variables η_1 and η_2 to obtain the same distribution as the vector of error terms ε_1 and ε_2:

$$\Gamma \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}, \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \rightsquigarrow \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right). \quad (7)$$

Such a decomposition is very useful when having to write explicitly the likelihood function for an observation and then to set possible additional constraints on the parameters of the model to identify it.

The probability of observing $y_1 = m$, $y_2 = j$ when sampling an observation with explanatory variables x is equal to

$$\forall m \in \{0, 1\}, \forall j \in \{1, \cdots, J\}, Pr(y_1 = m, y_2 = j|x, \gamma) = Pr(y_2{=}j, y_1{=}0|x, \gamma)^{\mathbb{I}(y_1=0)} Pr(y_2{=}j, y_1{=}1|x, \gamma)^{\mathbb{I}(y_1=1)}, \quad (8)$$

where $\gamma = \left(\delta_1', \delta_2', \alpha, \sigma_1^2, \sigma_2^2, \sigma_{2,1}\right)'$ is the vector of unknown parameters. The likelihood for an observation is then defined as

$$\ell(y_1, y_2|x, \gamma) = \prod_{j=1}^{J} Pr(y_2 = j, y_1 = 0|x, \gamma)^{\mathbb{I}(y_1=0)\mathbb{I}(y_2=j)} \times \prod_{j=1}^{J} Pr(y_2 = j, y_1 = 1|x, \gamma)^{\mathbb{I}(y_1=1)\mathbb{I}(y_2=j)}, \quad (9)$$

where

$$Pr(y_2 = j, y_1 = 0|x, \gamma) = \Phi\left(\frac{\bar{y} - x_1'\delta_1}{\sigma_1}\right) \int_{-\infty}^{\bar{y}-x_1'\delta_1/\sigma_1} \Phi \times \left(\frac{\beta_j - x_2'\delta_2 - (\sigma_{2,1}/\sigma_1)\eta_1}{\sigma_2\sqrt{1-\rho^2}}\right) \varphi(\eta_1)\,\mathrm{d}\eta_1 - \Phi\left(\frac{\bar{y} - x_1'\delta_1}{\sigma_1}\right) \times \int_{-\infty}^{\bar{y}-x_1'\delta_1/\sigma_1} \Phi\left(\frac{\beta_{j-1} - x_2'\delta_2 - (\sigma_{2,1}/\sigma_1)\eta_1}{\sigma_2\sqrt{1-\rho^2}}\right) \varphi(\eta_1)\,\mathrm{d}\eta_1 \quad (10)$$

and

$$Pr(y_2 = j, y_1 = 1|x, \gamma) = \left(1 - \Phi\left(\frac{\bar{y} - x_1'\delta_1}{\sigma_1}\right)\right) \int_{\bar{y}-x_1'\delta_1/\sigma_1}^{+\infty} \Phi \times \left(\frac{\beta_j - x_2'\delta_2 - \alpha - (\sigma_{2,1}/\sigma_1)\eta_1}{\sigma_2\sqrt{1-\rho^2}}\right) \varphi(\eta_1)\,\mathrm{d}\eta_1 - \left(1 - \Phi\left(\frac{\bar{y} - x_1'\delta_1}{\sigma_1}\right)\right) \int_{\bar{y}-x_1'\delta_1/\sigma_1}^{+\infty} \Phi \times \left(\frac{\beta_{j-1} - x_2'\delta_2 - \alpha - (\sigma_{2,1}/\sigma_1)\eta_1}{\sigma_2\sqrt{1-\rho^2}}\right) \varphi(\eta_1)\,\mathrm{d}\eta_1. \quad (11)$$

Φ stands for the cumulative distribution function of the standardized normal distribution. The log-likelihood function is defined as the logarithm of the likelihood function presented in Eq. (9):

$$\ln \ell(y_1, y_2|x, \gamma) = \sum_{j=1}^{J} \mathbb{I}(y_2 = j)\,\mathbb{I}(y_1 = 0) \ln(Pr(y_2 = j, y_1 = 0|x, \gamma)) + \sum_{j=1}^{J} \mathbb{I}(y_2 = j)\,\mathbb{I}(y_1 = 1) \ln(Pr(y_2 = j, y_1 = 1|x, \gamma)). \quad (12)$$

There are additional constraints to set on the parameters of the model to make unique the mathematical relation between the log-likelihood function and the vector of parameters to estimate. It must be ensured that, for given values of the parameters to estimate, the value of the log-likelihood function is unique. Conversely, it must be also ensured that there is only one set of values for the parameters that lead to a given value of the log-likelihood function. In the present model, only the parameters $\rho, \alpha/\sigma_2, \delta_2/\sigma_2, \forall j \in \{1, \cdots, J\}, \beta_j/\sigma_2$, and $\bar{y}/\sigma_1, \delta_1/\sigma_1$ can be estimated. Additional identification problems arise if there are intercept terms in the explanatory variables x_1 and x_2. If there is an intercept term in x_1 then only the difference (scaled by σ_1) of $\bar{y}$ with it can be estimated. If there is an intercept term in x_2 then only the differences (scaled by σ_2) of all β_j with it can be estimated. Except for the correlation coefficient of the random error terms, the numerical values of the estimates have no interpretation. Only matter their signs and their levels of statistical significancy. One can however use these estimates to compute marginal effects of the explanatory variables on the probabilities of the observable outcomes and/or elasticities of these probabilities with respect to the explanatory variables, which have numerical values that can be interpreted.

Even parametrized in terms of estimable coefficients, say a vector $\theta = \left(\rho, \alpha/\sigma_2, \delta_2'/\sigma_2, \tilde{\beta}_1/\sigma_2, \cdots, \tilde{\beta}_{J-1}/\sigma_2, \tilde{y}/\sigma_1, \delta_1'/\sigma_1\right)'$ where $\tilde{\beta}_1, \cdots, \tilde{\beta}_{J-1}, \tilde{y}$ models the intercept terms of the model that can be estimated, the log-likelihood function does not have any closed form solution. Monte-Carlo integration techniques, our choice here, or numerical integration techniques can be used.

The two integrals to approximate are the distributions of y_2 conditional to the outcomes of y_1. The probability that y_2 takes the value j conditional to the fact that y_1 takes the value 0 is approximated by a Riemann expansion of the integral in Eq. (10):

$$\begin{aligned} &Pr^R(y_2 = j|y_1 = 0, x, \zeta, \theta) \\ &= \Phi\left(\frac{\bar{y} - x_1'\delta_1}{\sigma_1}\right) \times \frac{1}{R}\sum_{r=1}^{R}\left(\Phi\left(\frac{\beta_j - x_2'\delta_2 - (\sigma_{2,1}/\sigma_1)\zeta_1^r}{\sigma_2\sqrt{1-\rho^2}}\right)\right. \\ &\quad \left. - \Phi\left(\frac{\beta_{j-1} - x_2'\delta_2 - (\sigma_{2,1}/\sigma_1)\zeta_1^r}{\sigma_2\sqrt{1-\rho^2}}\right)\right), \end{aligned} \tag{13}$$

where, for each draw $r = 1, \cdots, R$, ζ_1^r is drawn in a truncated normal distribution:

$$\zeta_1^r = \Phi^{-1}\left(u^r \Phi\left(\frac{\bar{y} - x_1'\delta_1}{\sigma_1}\right)\right), \tag{14}$$

and $u^r \rightsquigarrow \mathcal{U}_{[0,1]}$ is drawn in a [0, 1] uniform distribution. The probability that y_2 takes the value j conditional to the fact that y_1 takes the value 1 is approximated by a Riemann approximation of the integral in Eq. (11):

$$\begin{aligned} &Pr^R(y_2 = j|y_1 = 1, x, \zeta, \theta) \\ &= \left(1 - \Phi\left(\frac{\bar{y} - x_1'\delta_1}{\sigma_1}\right)\right) \\ &\quad \times \frac{1}{R}\sum_{r=1}^{R}\left(\Phi\left(\frac{\beta_j - x_2'\delta_2 - \alpha - (\sigma_{2,1}/\sigma_1)\xi_1^r}{\sigma_2\sqrt{1-\rho^2}}\right)\right. \\ &\quad \left. - \Phi\left(\frac{\beta_{j-1} - x_2'\delta_2 - \alpha - (\sigma_{2,1}/\sigma_1)\zeta_1^r}{\sigma_2\sqrt{1-\rho^2}}\right)\right), \end{aligned} \tag{15}$$

where, for each draw $r = 1, \cdots, R$, ζ_1^r is drawn in a truncated normal distribution:

$$\zeta_1^r = \Phi^{-1}\left(u^r\left(1 - \Phi\left(\frac{\bar{y} - x_1'\delta_1}{\sigma_1}\right)\right)\right), \tag{16}$$

and $u^r \rightsquigarrow \mathcal{U}_{[0,1]}$ is drawn in a [0, 1] uniform distribution. The simulated log-likelihood function replaces the likelihood function in Eq. (9). It is defined as

$$\begin{aligned} &\ln \ell^R(y_1, y_2|x, \zeta, \theta) \\ &= \sum_{j=1}^{J}\mathbb{I}(y_2 = j)\,\mathbb{I}(y_1 = 0)\ln\left(Pr^R(y_2 = j, y_1 = 0|x, \zeta, \theta)\right) \\ &\quad + \sum_{j=1}^{J}\mathbb{I}(y_2 = j)\,\mathbb{I}(y_1 = 1)\ln\left(Pr^R(y_2 = j, y_1 = 1|x, \zeta, \theta)\right). \end{aligned} \tag{17}$$

The maximum simulated likelihood estimator is defined as the solution of the program of maximization of the simulated log-likelihood function with respect to the vector of unknown parameters to estimate:

$$\hat{\theta}_{\text{msl}} = \text{argmax}_\theta\left(\sum_{i=1}^{n}\ln \ell^R\left(y_{i,1}, y_{i,2}|x_i, \zeta_i, \theta\right)\right). \tag{18}$$

The maximum simulated likelihood estimator (e.g. Lee, 1992; Gourieroux and Montfort, 1997) $\hat{\theta}_{\text{msl}}$ is asymptotically equivalent to the classical maximum likelihood estimator (see for instance Newey and McFadden, 1994) as long as $n \to +\infty$, $R \to +\infty$, and $\sqrt{n}/R \to 0$. The latter condition implies that the number of simulations required for the approximation of probabilities should be greater than the square root of the number of observations in the sample. At finite distance, that is when n is finite, an acceptable minimal contribution of the approximations of probabilities to the bias of the estimator is obtained for a large number R of simulations.

3. Data

The statistical data source is the 2003 French database of accidents reported by police authorities (namely BAAC for Bulletin d'Analyse des Accidents Corporels). The reports give details about collisions or incidents that may or may not lead to injury, happening on a public or private road and involving at least one moving vehicle. There are four tables in the database: the vehicle file, the accident file, the individual file, and the location file. The individual file is split between a file concerning drivers and passengers of motorized or non-motorized vehicles, and a file concerning pedestrians. Observation is disaggregated at the level of the accident and its consequences on all involved individuals and their vehicles.

Many car crashes with property damage only and car crashes with property damage and minor injury are not reported in the sample, causing a bias: one overestimates the severity of accidents. It is likely that these crashes are not reported because safety belts were used thereby lessening the need to request a police report. For instance, car accidents with only material damages are underrepresented in the database of accident reports, due to the fact that, in this case, only insurance companies are involved and not the police. But the fact that material damages are the only consequences of a crash may be also due to the fact that a safety belt was fastened. Using only the BAAC database therefore tends to underestimate the effect of fastening a safety belt on the severity of an accident. It lends to underweigh accidents with material damage only and accident with minor injury. The BAAC database should be matched with crash data of insurance companies, hospitals and clinics, and surveys data about car safety in order to correct sample selection biases. One must however be aware that this bias cannot be fully corrected because there are drivers that would not report minor accidents to insurance companies either, hospital records have only injuries, and survey data have also specific sampling problems.

As proposed by Evans and Frick (1988), Huelke and Compton (1995), car users have to be distinguished according to their

Table 1
Descriptive statistics: endogenous variables, driver

Frequencies of observed endogenous variables (% in brackets), drivers			
Safety belt fastened	Yes	No	Total
No injury	46,938 (59.08)	723 (20.81)	47,661 (57.48)
Light injury	26,099 (32.85)	1586 (45.64)	27,684 (33.38)
Severe injury	4,854 (6.11)	653 (18.79)	5,507 (6.64)
Fatal injury	1,857 (1.96)	513 (14.76)	2,070 (2.50)
Total	79,448 (95.81)	3474 (4.19)	82,922 (100)

Table 2
Descriptive statistics: endogenous variables, front-seat passenger

Frequencies of observed endogenous variables (% in brackets), front-seat passengers			
Safety belt fastened	Yes	No	Total
No injury	4,773 (32.50)	78 (8.26)	4,851 (31.04)
Light injury	7,995 (54.44)	592 (62.71)	8,586 (54.93)
Severe injury	1,486 (10.12)	173 (18.33)	1,659 (10.62)
Fatal injury	432 (2.94)	101 (10.70)	533 (3.41)
Total	14,685 (93.96)	944 (6.04)	15,629 (100)

positions in their cars before their accidents happened. There are at least three possible positions to consider: driver, front-seat passenger, rear-seat passenger. In the sample used for the application, 77.08% are drivers, 14.53% are front passengers and 8.39% are rear passengers. Three samples are built, one for each category of car user. The probability distributions of their injuries resulting from car accidents are presented in Tables 1–3 . They report the frequencies of observed levels of injury for the 3 types of car occupants (car driver, front-seat passenger, rear-seat passenger). For each of these tables, there are three columns: the first column reports the observed distribution of accident severity given that the car occupant fastened his/her seat belt; the second column reports the observed distribution of accident severity given that the car occupant did not fasten his/her seat belt; the third column reports the marginal distribution of accident severity.

Comparison between different random prospects (here accident severity by position in the car) can sometimes be approached by examining their associated cumulative distribution functions. One expects that just about everyone will consider a prospect which has a lot of probability mass skewed towards higher returns (no/low injury) to be better than a prospect which has a lot of probability mass skewed towards low returns (severe/fatal injury). If such is the case, then we have what is called stochastic dominance (Rothschild and Stiglitz, 1970, 1971). Data show that when accident happens, the probability to have a lower level of injury than a predetermined one is always higher for drivers than for front-seat passengers and rear-seat passengers. The probability to have a lower level of injury than a predetermined one is always higher for rear-seat passengers than for front-seat passengers. Without regard to accident configuration, design of involved vehicles, and behaviour of involved road users sitting in the accidented car, data show that one should have minimized our probability of high injury level first as a driver, then as a rear-seat passenger, and finally as a front-seat passenger.

Although fastening the safety belt is mandatory, the rule is not always respected, as shown by our data. 4.19% of drivers, 6.04% of front-seat passengers, and 12.54% of rear-seat passengers, did not fasten their safety belts. When one distinguishes distributions of accident severity according to wether or not a safety belt was fastened, one obtains additional results. Firstly, whatever is the position of car occupant, it is observed that the probability to have lower injury when an accident happens is always larger when the safety belt is fastened. Secondly, when it is not fastened, observed data show that it is not possible to draw results (in terms of stochastic dominance) about the position in car which minimizes risk of severe injury when an accident happens.

The exogenous variables that are used to explain the levels of the latent variables of the model describe different types of factors. Some of them appear in both equations of the model and the others appear only in the equation of accident severity, as it is reported in Table 4 in the next section. The decision to fasten the safety belt is explained by the age and the gender of the car user as well as the time of day when the trip was carried out. In addition to the variable indicating whether or not a safety belt was fastened, the level of injury of the car user is also explained by his/her age and gender, by the time of the day, by the type of road on which the accident happened, by the geographical location of the accident, and by the type of collision. Table 5 reports descriptive statistics about the distributions of these characteristics accross the populations of drivers, front-seat passengers, and rear-seat passengers.

It is observed that front-seat passengers and rear seat passengers have relative parity in gender but the results is not true for drivers. 2/3 of them are men. As compared to accidented drivers, accidents involving front-seat passengers and/or rear-seat passengers happened more during weekends. 47.34% of the reported accidents of rear-seat passengers, 42.08% of those of front seat passengers, and 28.57% of those of drivers, happened during weekends. Morevoer, 73.85% of the reported accidents of rear-seat passengers, 64.07% of those of front seat passengers, and 41.83% of those of drivers, happened out of city. Descriptive statistics show that 15.64% of the reported accidents of rear-seat passengers, 22.92% of those of front seat passengers, and 42.10% of those of drivers, happened on local roads. One can also observe that, on average, rear-seat passengers involved

Table 3
Descriptive statistics: endogenous variables, rear-seat passenger

Frequencies of observed endogenous variables (% in brackets), rear-seat passengers			
Safety belt fastened	Yes	No	Total
No injury	3198 (40.48)	191 (16.86)	3389 (37.52)
Light injury	3849 (48.73)	584 (51.54)	4433 (49.08)
Severe injury	691 (8.75)	266 (23.48)	957 (10.60)
Fatal injury	161 (2.04)	92 (8.12)	253 (2.80)
Total	7899 (87.46)	1133 (12.54)	9032 (100)

Table 4
Estimates of the models, maximum simulated likelihood

Variable	Seat belt equation: Driver Estimate	Driver t-stat	Front-seat passenger Estimate	Front-seat passenger t-stat	Rear-seat passenger Estimate	Rear-seat passenger t-stat
Intercept	2.273 ***	98.954	1.988 ***	27.143	2.643 ***	18.740
D.V.: female gender	0.247 ***	10.391	0.313 ***	4.323	ns	–
D.V.: night trip	−0.187 ***	−13.660	−0.342 ***	−5.454	−0.421 ***	−6.954
D.V.: age < 12	–	–	−0.198 **	−2.023	0.501 ***	9.467
D.V.: age ∈ [12; 25]	0.092 *	1.846	ns	–	ns	–
D.V.: age ∈ [50; 70]	0.199 ***	4.523	0.268 ***	5.227	0.402 ***	3.192
D.V.: age ≥ 70	0.085 *	1.765	0.288 ***	4.539	0.742 ***	7.333

Variable	Injury severity equation: Driver Estimate	Driver t-stat	Front passenger Estimate	Front passenger t-stat	Rear passenger Estimate	Rear passenger t-stat
D.V.: female gender	ns	–	0.206 **	2.103	ns	–
D.V.: night trip	0.581 ***	29.122	0.098 ***	5.774	1.012 ***	2.632
D.V.: age < 12	–	–	ns	–	−0.115 **	−2.499
D.V.: age ∈ [12; 25]	0.240 ***	20.330	0.022 *	1.776	−0.045 *	−1.653
D.V.: age ∈ [50; 70]	0.194 ***	13.554	0.074 **	2.334	0.112 **	2.511
D.V.: age ≥ 70	ns	–	0.302 ***	9.771	0.203 ***	4.398
D.V.: highway	ns	–	ns	–	−0.224 **	−2.123
D.V.: national road	−0.085 ***	−2.689	0.099 *	1.947	ns	–
D.V.: secondary road	0.275 ***	12.222	0.101 ***	2.986	ns	–
D.V.: out of city	0.088 ***	10.009	0.218 ***	9.356	0.056 *	1.775
D.V.: medium city	ns	–	0.666 ***	7.840	0.492 ***	6.211
D.V.: large city	1.555 ***	33.090	0.210 **	2.520	0.100 **	1.901
D.V.: wet road	ns	–	0.099 *	1.777	ns	–
D.V.: at an intersection	0.314 ***	14.227	−0.090 **	−1.962	−0.298 ***	−2.790
D.V.: car–car frontal crash	0.757 ***	17.441	0.338 ***	7.004	ns	–
D.V.: car–car lateral crash	0.586 ***	32.422	−0.150 ***	−6.200	−0.119 **	−1.970
D.V.: car–car rear crash	−0.092 ***	−3.140	−0.189 ***	−7.904	−0.158 **	−2.109
D.V.: weekends	−0.118 ***	−5.339	ns	–	ns	–
Seat belt effect α/σ_2	−1.698 ***	−13.006	−1.529 ***	−8.095	−1.001 ***	−7.336
Bound β_1/σ_2	−0.907 ***	−6.003	−1.731 ***	−8.933	−1.282 ***	−8.001
Bound β_2/σ_2	0.303 *	1.935	ns	–	ns	–
Bound β_3/σ_2	1.111 ***	3.800	0.588 **	2.505	1.821 ***	3.574
Correlation ρ	0.280 ***	3.966	0.402 ***	4.522	0.207 *	1.910

Summary statistics	Driver	Front passenger	Rear passenger
Sample size	82922	15629	9032
Number of simulations	500	200	200
Final log-likelihood function	−84867.032	−19303.245	−12428.789
Final log-likelihood function with intercepts only	−196183.527	−37110.206	−25967.426
Likelihood Ratio Index, in %	56.741	47.984	52.137

D.V.: Dummy variable, ns: not significantly different from 0 at the level of 10%
* Significant at 10% risk level.
** Significant at 5% risk level.
*** Significant at 1% risk level.

in a car accident were younger than those involved as front-seat passengers and drivers.

4. Results

Because of the identification constraints on the parameters, an individual of reference and an accident of reference must be chosen. All the coefficients have to be interpreted as differences with respect to them. Our reference individual is a man older than 25 years old and younger than 50 years old. The reference Accident that is not a car collision happens on a local road, but not at an intersection, in a city less than 200,000 inhabitants during weekdays with dry weather conditions. There is also no car collision[2] and it happens during weekdays with dry weather conditions.

[2] In the database, the type of collision is distinguished for car–car collisions but not for other types of collisions. In the present approach, no car collision might indicate collision with a fixed object or a pedestrian, or no collision at all.

Table 5
Descriptive statistics: exogenous variables

	Frequencies of exogenous variables (% in brackets)		
Position of car occupant	Driver	Front passenger	Rear passenger
D.V.: female gender	26,220 (31.62)	8,230 (52.66)	4110 (45.50)
D.V.: night trip	25,150 (30.33)	5,794 (37.07)	3511 (38.87)
D.V.: age < 12	–	524 (3.35)	2811 (31.12)
D.V.: age ∈ [12; 25]	17,521 (21.13)	5,254 (33.62)	3598 (39.84)
D.V.: age ∈ [25; 50]	42,788 (51.60)	5,387 (34.47)	1448 (16.03)
D.V.: age ∈ [50; 70]	16,344 (19.71)	2,666 (17.06)	565 (6.26)
D.V.: age ≥ 70	6,269 (7.56)	1,797 (11.50)	610 (6.75)
D.V.: highway	7,886 (9.51)	2,176 (13.92)	1613 (17.86)
D.V.: national road	13,010 (15.69)	3,120 (19.96)	1765 (19.54)
D.V.: secondary road	27,157 (32.75)	6,752 (43.20)	4241 (46.96)
D.V.: non-dry road [a]	15,672 (18.90)	3,213 (20.56)	1897 (21.00)
D.V.: at an intersection [b]	27,331 (32.96)	4,498 (28.78)	2239 (24.79)
D.V.: out of city	34,686 (41.83)	10,014 (64.07)	6670 (73.85)
D.V.: medium city [c]	8,002 (9.65)	685 (4.38)	257 (2.85)
D.V.: large city [d]	9,005 (10.86)	891 (5.70)	254 (2.81)
D.V.: car–car frontal crash	10,000 (12.06)	2,362 (15.11)	1373 (15.20)
D.V.: car–car lateral crash	26,668 (32.16)	4,376 (28.00)	2177 (24.10)
D.V.: car–car rear crash	9,478 (11.43)	1,611 (10.31)	863 (9.55)
D.V.: weekends	23,691 (28.57)	6,577 (42.08)	4276 (47.34)

D.V.: Dummy Variable.
[a] A non-dry road is defined as a wet road or a road with snow or a road with mud.
[b] Cross-roads and roundabouts.
[c] A medium city is a city with a number of inhabitants between 200,000 and 1 million.
[d] A large city is a city with a number of inhabitants more than 1 million.

One must also stress that, as the seat belt variable is predicted in the injury model, if the seat belt model is not very good then this prediction isn't very good and the overall results are questionable.

As shown by the negative coefficients in the equations of accident severity, use of a safety belt results in a reduction of the level of injury whatever is the position of the car occupant. By wearing a belt when accident does happen, there is the best chance of reducing or avoiding injury if everyone is buckled up.

It has been also suggested that some car users who wear safety belts because of mandatory safety belt law may feel their increased protection allows them to take more risks on the road (Garbacz, 1992) although other studies refuse this claim (Thomas et al., 1989). In the present application, whatever is the position of the car occupant, the results show a positive correlation of the unobserved components of the equations. As it regards the driver, it means that there exists some risk compensation. When a change of variable occurs, there is a reaction in which the system responds in such a way as to reverse the direction of change. It tends to keep things constant. As compared to the situation where drivers would not compensate the benefit of safety belt use, the results show that they compensate it by behaving in a way that expose them to more severe injuries. As it regards the positive signs of the correlations for front-seat passengers and rear-seat passengers, they have no interpretation in terms of risk compensation: passengers cannot compensate for their safety, only the driver can compensate for safety by changing the driving behavior.

There are two complimentary effects that lead to higher probability of more severe injury when accident happens during night. There is a direct effect that isolates the role of the dummy variable "night" on the level of injury when accident happens. There is also an indirect effect that measures the role of the dummy variable "night" on the decision to fasten the safety belt, which in turn plays a direct role on the level of injury when accident happens. The following results are along with findings of Page et al. (2002). They are found for drivers, front-seat passengers and rear-seat passengers. Whatever the safety belt is fastened or not, the probability of more severe injury during night is higher than during daylight. It is also observed that the probability to fasten the safety belt decreases during night. Someone might say that car users are more likely to use seatbelt at night since they are more cautious in dark and high-speed conditions. However, the results show the contrary. On the one hand, there is a direct and significant positive effect of night on the probability of a more severe injury. On the other hand, there is a indirect effect that gives support to the first one and that passes by the increased probability of not fastening the safety belt and therefore not compensating the former increase of the probability of a more severe accident.

Injury severity is affected by age by a direct effect and an indirect effect. The direct effect isolates the role of the variable "age" on the level of injury when accident happens. the results show also that the probability to have more severe injury increases also as people get older. Older car users are physically weaker and they are more likely to be severely injured than younger car users. The indirect effect measures the role of the variable "age" on the decision to fasten the safety belt. The results show that the probability to fasten the safety belt increases significantly as people get older. Older car users anticipate that they are more exposed to severe injury than younger car users when accident happens. They are therefore more likely to fasten their safety belt to compensate partly their weaker physical conditions. These results are found for drivers, front-seat passengers, and rear-seat passengers. Cunningham et al. (2001) have drawn the same conclusions in their application.

The effects of gender vary with the position of the occupant. For rear-seat passengers, the results show that there is no difference in the probability of fastening the safety belt and in the probability of a level of injury. Distribution of accident severity for rear-seat passengers is not a matter of gender. As compared to men drivers, women drivers have higher probability to fasten their safety belts. But the results show that there is no difference in the direct effect of gender on the probability of a level of injury. But as women drivers have a higher probability of fastening their safety belts, and because fastening the safety belt diminishes significantly the probability of a severe accident, one can draw from the results the conclusion that women drivers have lower probability of injury than men drivers. For front-seat passengers, the results show a significant direct gender effect: women have higher probability of severe injury than men. They show also that women have a higher probability of fastening their safety belts. The indirect effect on accident severity is therefore negative. None of the results is surprising and all mathch those of Ulfarsson and Mannering (2004) or of Preusser et al. (1991). Women are more likely to be severely injured than men because of various physical differences that are favourable to men in case of an accident. One could also speculate different psychological patterns between men and women that give favour to women in terms of respect of the law and consciousness of road risk. Whatever reason is true, the results show that women are more exposed to severe injury in case of a car accident. The results show also that women tend to use the safety belt more often than men to ensure a partial compensation of their higher probability of severe injury.

As compared to an accident that happens at an intersection, the probability to be injured is significantly lower for rear-seat passengers and for front-seat passengers when it happens at an intersection. It is significantly higher for drivers. An additional investigation of the database shows that most of accidents that happened at intersections were cars coming from the left crashing in cars already engaged. Left-lateral car collision exposes driver to more severe injury than any other type of collision as shown by the positive coefficient of the dummy variable 'car–car lateral crash'. As such type of collision happened mostly at intersections, they were at locations where drivers were highly exposed to severe/fatal injuries. As this accident configuration dominates in crashes at intersections, it is not so surprising to find that drivers have higher probability of injury than if their accidents happened not at an intersection. Among results are also the decreasing probabilities of injury for front-seat and rear-seat passengers when their accidents happened at an intersection. Observatoire (2004) have shown, without regards to the types of car users, a lower level of accident severity at intersections as compared to other locations of accidents in France. As one distinguishes the types of car users, the results of the models show that probability of injury at intersections actually increases for drivers.

The type of road network plays significant roles on the distributions of accident severity of the different types of car users. The results show that accidents are less severe on highways for rear-seat passengers. They show also that accidents are significantly more severe on secondary roads (departmental road) and trunk roads (national roads) for front-seat passengers. They show finally that accidents are more severe on secondary roads and less severe on trunk roads for drivers. Each type of road network produces accidents of various levels of severity. Given these accidents and the distribution of their severity on the different types of network for the different positions of car occupant, the results state for drivers that the probability distribution of injury on trunk roads involves less risk than the probability distributions of injury on the other types of roads. The results state for front-seat passengers that the probability distributions of accident severity on local roads and highways involve less risk than the distributions of accident severity on secondary roads and trunk roads. The results state for rear-seat passengers that the probability distributions of accident severity on highways involves less risk than the distributions of accident severity on the other types of roads. As compared to a local road, the probability to be more severely injured increases for drivers and front-seat passengers when accidents happen on a secondary road. The probability to be more severely injured increases for front-seat passengers and it decreases for drivers when accidents happen on a trunk road (national road). The probability to be more severely injured decreases for rear-seat passengers when accidents happen on a highway.

Urban density plays also different roles on the levels of injury resulting from accidents. As compared to small cities, the probability of injury increases for every car occupants when accident happens outside of cities. When accident happens out of the city, one could think that travel speeds are higher because traffic is less dense. Although driver should be more cautious while driving faster, the kinetic energy generated by a crash or a collision is higher and therefore resulting injuries are likely to be more severe. The results state that accident severity is likely to be higher on rural roads as compared to urban roads. The results show also that the probability of injury increases for every car occupant as urban density becomes higher. As compared to small city environments, although travel speeds could be lower in more dense urban area, higher density of traffic flows of various types of road users generate more difficulty for drivers to control for the environment and to react quickly and appropriately to diminish severity when accident happens.

The results that focus on accidents that involve two cars and the way they collide together show that there are different consequences. As concerns drivers, probability of injury of car–car frontal and lateral collision is higher than probability of injury of any other type of collision. As concerns front-seat passengers, probability of injury of car–car frontal collision is higher than probability of injury of any other type of collision. As concerns rear-seat passengers and front-seat passengers, probability of injury of car–car lateral collision is lower than probability of injury of any other type of collision. For every types of car users, probability of injury of car–car rear collision is lower than probability of injury of any other type of collision.

The probability of injury when accident happens is not significantly influenced by road dampness. Edwards (1998) shows the same kind of result, but he states also that it is not systematic and it does vary according to different categories of hazardous weather conditions.

Probability distributions of accident severity for front-seat passengers and rear-seat passengers are not elastic to the day of the week. The fact that accidents happened during weekends does not play a significant role on understanding severity of their accidents. One could have thought that weekend travelling patterns as well as the levels of weekend traffic during weekends are different so that the risk of accident and its severity might have different levels as compared to working days, but the results undermine this argument for front-seat passengers and rear-seat passengers. It is however observed for drivers that probability of injury decreases when accidents happen during weekends.

5. Conclusions

There are very few situations where it would be better not to have the belt on. The use of a safety belt is very strongly related to a decrease in severe injuries and therefore tends to save lives. The results endorse an increase of safety belt enforcement and usage. There are several solutions: continuous education programs and advertisements, more severe penalties, safety belt reminder systems (visual and audible devices) that detect whether or not belts are in use in different seating positions and give out increasingly aggressive warning signals until the belts are used. However, the results about risk compensation show that safety belt enforcement and usage has to be complemented by additional safety policies to reduce risk compensation.

There are other factors affecting severity of accidents, not only use of a safety belt. Moreover, these factors play different roles according to the types of car users, their age and gender. Understanding the differences in the population of car users is crucial to prevent serious injuries. Road safety policies need therefore to take into account these differences to manage efficiently accident severity on roads. Policies address issues covering road engineering, signage, vehicle design, education of road users and enforcement of traffic safety measures.

Finally, it seems that the methodological approach is very promising for further investigation of the contribution of the behavior of road user to accidents. It should however be highlighted that there are numerous other variables that could be thought of as correlated to seat belt use but it could be difficult to add more and more simultaneous equations to this structure. Furthermore, one has to take care of the classical problems of simultaneity and endogeneity in estimation of systems of equations: if one assumes that other explanatory variables than seat belt use are set simultaneously with it and accident severity, and if one makes as if it is not the case, then it is well known that estimates are biased except if their error terms are not correlated (which is a strong a priori assumption). On top of that, as one does not consider endogeneity of these variables in the system, one cannot isolate their net effects on the other endogenous variables. A last remark is that even if the results presented in the application are consistent with intuition, it must be pointed out that the application is made by using data with an inherent sampling bias because of the way accident reports are collected. It forces us to have hindsight about the results.

References

Abdel-Aty, M.A., 2003. Analysis of driver injury severity levels at multiple locations using ordered probit models. J. Safe. Res. 34, 597–603.

Al-Ghamdi, S.A., 2002. Using logistic regression to estimate the influence of accident factors on accident severity. Accid. Anal. Prev. 36, 447–456.

Cunningham, C., Howard, D., Walsh, J., Coakley, D., O'Neill, D., 2001. The effects of age on accident severity and outcome in Irish road traffic accident patients. Irish Med. J. 94, 169–171.

Edwards, J.B., 1998. The relationship between road accident severity and recorded weather—exploratory analysis of road accidents. J. Safe. Res. 29, 249–262.

Eluru, N., Bhat, C., 2007. A joint econometric analysis of seat belt use and crash related injury severity. Accid. Anal. Prev. 39 (5), 1037–1049.

European Commission, 2001. White Paper- European Transport Policy for 2010: time to decide. Office for official publications of the European communities.

European Commission, 2003. Saving 20,000 lives on our roads: a shared responsibility. Office for official publications of the European communities.

Evans, L., Frick, M.C., 1988. Seating position in cars and fatality risk. Am. J. Public Health 78 (11), 1456–1458.

Garbacz, C., 1992. More evidence on the effectiveness of seat belt laws. Appl. Econ. 24, 313–315.

Gourieroux, C., Montfort, A., 1997. Simulation-based Econometric Methods. Oxford University Press.

Huelke, D.F., Compton, C.P., 1995. Effects of seat belts on injury severity of front and rear seat occupants in the same frontal crash. Accid. Anal. Prev. 27, 835–838.

Kockelman, K.M., Kweon, Y.J., 2002. Driver injury severity: an application of ordered Probit models. Accid. Anal. Prev. 34, 313–321.

Lee, L.-F., 1992. On the efficiency of methods of simulated moments and maximum simulated likelihood estimation of discrete response models. Economet. Theory 8, 518–552.

Newey, W., McFadden, D., 1994. Large sample estimation and hypothesis testing. In: Engle, R.F., McFadden, D. (Eds.), Handbook of Econometrics, vol. 4, pp. 2111–2245.

Observatoire National Interministriel de la Scurit Routire - ONISR, 2004. Les grandes donnes de l'accidentologie: caractristiques et causes des accidents de la route.

O'Donnell, C.J., Connor, D.H., 1996. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. Accid. Anal. Prev. 28, 739–753.

Page, Y., Cuny, S., Montel, M.C., 2002. Road accidents at night. Ingenieurs de l'automobile 757, 58–76.

Preusser, D., Williams, A., Lund, A., 1991. Characteristics of belted and unbelted drivers. Accid. Anal. Prev. 23 (6), 475–482.

Rothschild, M., Stiglitz, J.E., 1970. Increasing risk I: a definition. J. Econ. Theory 2 (3), 225–243.

Rothschild, M., Stiglitz, J.E., 1971. Increasing risk II: its economic consequences. J. Econ. Theory 3 (1), 66–84.

Srinivasan, K.K., 2002. Injury severity analysis with variable and correlated thresholds: ordered mixed logit formulation. Transportation Research Record, No. 1784, pp. 132–142.

Thomas, H., Fabry, J., Girard, R., 1989. The effect of the 1979 French seat-belt law on the nature and severity of injuries to front-seat occupants. Accid. Anal. Prev. 21, 589–594.

Ulfarsson, G., Mannering, F., 2004. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. Accid. Anal. Prev. 36 (2), 135–147.

Wang, X., Kockelman, K., 2005. Use of heteroscedastic ordered logit model to study severity of occupant injury: distinguishing effects of vehicle weight and type. Transportation Research Record 1908, pp. 195–204.