A driving behavior model evaluation for UBI

Zhishuo Liu, Qianhui Shen and Jingmiao Ma
School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

Received 25 August 2017
Revised 8 September 2017
Accepted 9 September 2017

Abstract

Purpose – This paper aims to provide a driving behavior scoring model to decide the personalized automobile premium for each driver.

Design/methodology/approach – Driving behavior scoring model.

Findings – The driving behavior scoring model could effectively reflect the risk level of driver's safe driving.

Originality/value – A driving behavior scoring model for UBI.

Keywords Automobile insurance premium, Driving behavior evaluation, Improved EW-AHP method, On-board diagnostics, Usage-based insurance

Paper type Technical paper

1. Introduction

Usage-based insurance, which also means pay-as-you-drive and pay-how-you-drive, is that the automobile insurance premium was decided by the actual driving time, place, driving habits and the driving behaviors (Dimitros *et al.*, 2016). The driving behavior data were transferred to remote server-based telematics, and then the automobile insurance company redesigns the insurance product based on the data (Qiao, 2015). The driver with safe driving behavior should get the premium discount. So the insurance was determined based on usage not only can let the policyholders pay a more reasonable premium, and it also can reduce the claims cost for the insurance company.

This type of insurance was first proposed by Lves in the Cincinnati automobile club. He put forward that the car premium can be replaced by gasoline tax in 1925. A paper published in 1930 point out that the use condition of the car, road situation, the traffic flow density, laws and regulations, driver information, speed, weather and other information can be taking into account in automobile insurance pricing (Greenberg, 2008). But because of the technical limited, it cannot be achieved in that time. Then, in 1968, Vickrey put forward several insurance pricing models based on mileage (Vickrey, 1968). However, in that time, the theory and technology about usage-based insurance are immature.

Today, monitoring devices can provide data about mileage, total driving time, location, safe driving, seat belt use, turn signal use, vehicle speed, sudden braking and sudden acceleration or deceleration. On-board diagnosis (OBD) system. The main function of the OBD system is to

The research is supported by National Science and Technology Support Program in China (No.2014BAH24F00).

International Journal of Crowd Science
Vol. 1 No. 3, 2017
pp. 223-236
Emerald Publishing Limited
2398-7294
DOI 10.1108/IJCS-08-2017-0022

diagnosis the automobile fault and detect the engine combustion effect. The OBD in the automobile can monitoring the engine state, emission control system, fuel system, exhaust gas recirculation system, etc. and the malfunction indicator lamp or the check engine warning light will turn on when something wrong in these systems. And the fault message can be read in the form of a fault code by the serviceman through the standard diagnostic instrument and diagnostic interface. According to the fault code, the maintenance person can determine the nature and location of the fault quickly and accurately. The OBD box can read the basic vehicle information such as engine speed, water temperature fuel consumption, etc. through the diagnostic interface and uploading the vehicle information to the driver's smartphone and the insurance company computers after filtering private information such as driving trajectory during data uploading. Finally, only accurate data needed for the research are read.

We build a driving behavior scoring model of UBI to score the drivers. And the higher the score, the lower the risky probability. The data were gathered from OBD. We discuss the theory and application status in Section 2 and pricing mode of UBI in Section 3. In Section 4, we choose five important driving behavior factors and design an evaluation indices system of driving behavior and use an improved entropy weight analytic hierarchy process (EW-AHP) method to determine the index weight. Finally, the driving behavior scoring model for UBI is established in Section 5 and its validity is verified by field experiments in Section 6. At last, the summary of the review, limitations and future research directions are outlined in Section 7.

2. Usage-based insurance development status

In recent years, UBI has become a research hotspot at home and abroad. Daniela *et al.* calculate the important degree of each premium factors (15 factors include person, car and the usage of the car) and pricing the insurance in consideration of market, customers individual and sales channels (Daniela, 2016). Sinisa *et al.* gives an overview of the system architecture of one of the telematics systems offered and used on the market, as well as the data model used in the billing process (Sinisa *et al.*, 2015). Dimitrios reviewed the existing literature on UBI schemes and pointed out that UBI provides a strong motivation for drivers to improve their driving behavior and reduce their degree of exposure by receiving their driving behavior data (Dimitros *et al.*, 2016). Peng jiangqin *et al.* put forward an intelligent UBI system with cloud, dig data, ubiquitous communication based on carrier-cloud-client, and they design a premium pricing model considering quick acceleration, hard deceleration and swerve maneuvers (Peng *et al.*, 2016).

At the same time, some insurance companies have also launched related UBI products, of which the USA, Britain and Germany insurance companies are more prominent (Litmant, 2011a). There are some companies which have applied driving behavior to personalized premium successfully. Progressive is one of the largest automobile insurance company in America. When the policyholders take part in their UBI automobile insurance plan, the car will be installed an OBD box called Snapshot. The OBD box will gather the drivers' driving behavior data include speed, driving time, hard deceleration, acceleration and total duration of equipment. (Progressive Insurance, 2014). And the Metronome is an OBD box installed by Metromile. The box not only gathers the driving behavior data and transfer the data to Metromile but also can help to find the lost car. In UK, the UBI products attract more and more customers. Some insurers made tentative attempts and got good results, such as Insure the box, Allianz and Aviva. Provinzial in Germany developed an item named MeiCopolot can offer up to 10 per cent discount for the drivers who has a good driving behavior (Qiao, 2015).

In domestic, the current UBI products are still immature in the market at present. Such as Launch Tech cooperate with Ping An automobile insurance and Baidu, have pushed out the metadata of Golo's Baidu map based on OBD, but the user experience is not good. The OBD always related to an smartphone app. The OBD box collects driving behavior data and then uploads the data to cloud platform, and after cloud platform processing data, the data will be transferred to the driver's cell phone. But the market demand of OBD is low because different OBD products have the same function, difference degree of the product is low and product quality is uneven. Sun Lijuan gives several advices to the insurers based on IOV (Sun, 2017).

3. Usage-based insurance pricing models

UBI mainly has three pricing models:

(1) *Per-mile Premiums model*: The vehicle insurance is sold by the vehicle-mile rather than the vehicle-year. This converts the unit of exposure from the vehicle year to vehicle-mile. The insurance company pricing per kilometers. You should pay the insurance based on the kilometers you drive. And the mileage data were gathered on odometer.

(2) *GPS-based pricing*: GPS can collect the driving time, location, speeding, mileage data on the car. This system uses GPS transponders installed in vehicles to price insurance based on time and location. The drivers should pay a higher premium when he drives in the congested road and rush hour.

(3) *Mileage rate factors (MRF) model*: The insurance company takes mileage as a rating factor. The drivers should estimate their total mileage in the next year and report the number to the insurance company. The insurers will offer some discounts to whose real mileage lower than the estimated mileage and surcharge to those traveled higher mileages. But drivers have a difficulty of predicting their annual mileage and tend to underestimate the mileage as well. (Litmant, 2011a, 2011b).

Generally, driving behavior rating factors (DBRF) can be seen a kind of MRF. It takes the driving behaviors as factors to determine whether to offer the discount or surcharge while pricing the insurance. Insurers can gather the drivers risky driving behavior performance, such as speed, region, mileage, the time of day, roadway types, hard deceleration, acceleration, swerve maneuvers etc. For DBRF model, the rate adjustment coefficients can be determined by a linkage model between the rate adjustment coefficients and the driving behavior score based on the evaluation of driving behavior. First, the scores are determined by driving behaviors. Then discount rate is proposed for each score grade.

While the relationship of various driving behavior factors to the risk of accident has been discussed by many researchers, literature about individual rating of UBI based on driving behaviors is sparse. Many researchers discussed the influence of mileage, speed and some temporal-spatial driving behavior activity such as time of day, roadway types, hard deceleration, acceleration and swerve maneuvers on road accidents (Yanagihara *et al.*, 2015; Campbell, 2003; Litmant, 2011a, 2011b; Maclean *et al.*, 2003; Martain, 2002; Traffic management bureau of the public security ministry, 2014; Ferreira and Minike, 2012; Boucher, 2013; Staplin *et al.*, 2008; Langford *et al.*, 2008; Paefgen *et al.*, 2014; Davis *et al.*, 2006; Elvik *et al.*, 2004; Jun *et al.*, 2007; Klauer *et al.*, 2009; Russell Henk *et al.*, 2010).

As far as UBI pricing is concerned, Ferreira (Ferreira and Minike, 2012) used the generalized linear model to compute pure premium per mile in which mileage is used in conjunction with those traditional rating factors. That study just considered mileage and ignored other drive behavior factors. Chenghui Han *et al.* built a pricing model based on GLM combined the static premium with dynamic premium. The model considering car model, driving zone, driving kilometers (Han, 2015). Consequently, this study established a risky driving behavior scoring model for the UBI pricing based on more driving behavior factors including mileage, speed, driving time of one day, hard deceleration, acceleration and swerve maneuvers.

4. Driving behavior scoring indices system and weight determination

4.1 Analyzing impact factors of driving safety

There are many factors that influence driving behavior and safety. Some scholars have summed up the 53 influencing factors (Klauer *et al.*, 2009), and the number of influence factors are as many as 18 by expert scoring and questionnaire investigation (Li and Li, 2015). For the purpose of this research, in this paper, 5 most typical impact factors including mileage, driving time, traffic flow, speeding, traffic violations, hard deceleration, acceleration and swerve maneuvers were selected to analyze the impact to driving safety.

4.1.1 Mileage. The relationship between the distance run by a vehicle and its influence on the risk of accident has been discussed by many researchers. Some of them consider that this relationship is proportional (Bordoff and Noel, 2008), whereas others argue that it is not proportional (Langford *et al.*, 2008; Litmant, 2005). The Texas Mileage Study published by Progressive Insurance found a linear relationship between mileage and insurance claims (Progressive Insurance, 2005). Boucher has used a generalized Poisson regression model to fit the relationship between mileage and risk and found that the correlation between mileage and risk is not linear (Boucher, 2013). Litmen found out that the mileage is positively correlated with accidents and insurance claims, and he thinks that those drivers who use more the car have fewer accidents per unit of distance than those who use less the car (Litmant, 2005). In general, vehicle mileage is positively related to driving risk. The greater the mileage traveled by the same vehicle during the insurance year, the greater the likelihood of an accident. In China, traffic accident loss can increase by 1 per cent every increase of mileage 5.41 per cent (Zhang and Duan, 2012).

At the same time, Paefgen discusses the correlation between mileage and driving risk. The discussed models combine mileage as a measure of the "extent" of exposure with several groups of situational variables that represent the "degree" of exposure, such as daytime, weekday, road type, and velocity (Paefgen *et al.*, 2014).

4.1.2 Fatigue driving. Studies have shown that fatigue driving is one of the important reasons of traffic fatalities. Drivers who drive more miles are more likely to violate HOS regulations, drive when drowsy (McCartt *et al.*, 2000) and be involved in crashes (Williams, 2001). Maclean's statistics show that 20 per cent of traffic accidents are related to fatigue driving (Maclean *et al.*, 2003). A research report of DOT has also pointed out that the commercial vehicle accident 20-40 per cent is due to fatigue driving. For the sleepy drivers, the cumulative number of lane departure events per minute was significantly higher. The experimental result shows that the probability of a lane departure event occurring is 0.35 (Hallvig *et al.*, 2014). Through the analysis of 182 cause of heavy truck driver died accidents, it shows that 31 per cent of accidents is related to driver fatigue (NO *et al.*, 2003). In terms of reasons of fatigue driving, drivers in sleep-related crashes are more likely to work night shifts, drive more often late at night, driving for longer time (Stuttsa *et al.*, 2003). Nighttime risk ranks at the top of the list for the youngest motorists on the road (Russell Henk *et al.*, 2010). The fatigue driving can be detected through monitor electroencephalographic activity and eye blink (or eye closure) duration (Lenné and Jacobs, 2016).

4.1.3 Traffic flow. The magnitude of traffic flow directly influences the degree of mental tension and the rate of traffic accidents, which is one of the main factors that influence the number of traffic accidents. Through long observation experiments, Martain finds that traffic flow is closely related to the traffic accident and the severity of the accident (NO *et al.*, 2003). The experimental data show that the severity of traffic accidents has a direct relationship with the traffic flow (Hou *et al.*, 2011). Lee *et al.*'s found suggest that crash likelihood during a peak period is higher than during an off-peak period. And Abdel-Aty and Pande had the same viewpoint (Abdel-Aty *et al.*, 2005).

Puala suggests that crash frequency, severity and type are expected to be affected by the changing flow conditions that occur when traffic starts to become congested (Paula and Wendy, 2010).

4.1.4 Speeding. Speeding has an impact on vehicle safety by affecting the driver's visual characteristics and vehicle stability. Joksch hold the opinion that speed is the most important determinant of crash severity (Joksch, 1993). Davis took two groups of experiment and used Bayesian relative risk regression to relate speed to crash risk and found crash risk clearly tended to increase as speed increased (Davis *et al.*, 2006). Matthew's research points out that the degree of crash risk increases exponentially with the increase of driving speed (Matthew and Climatology, 2004). Elvik found that the relationship between speed and accidents or accident victims can be represented by a set of power functions (Elvik *et al.*, 2004). Farmer has researched the safety effect of increase in the US state maximum speed limits and found out that when speed limits are raised there is a definite trend of increased fatality risk (Farmer, 2016). But Nilsson developed a power model of the relationship between speed and accidents (Nilsson, 2004).

OECD/ECMT estimated that a 5 per cent average increase in speed leads to approximately 10 per cent increase in total accidental injuries and 20 per cent increase in fatal accidents. Similarly, for a 5 per cent average decrease in speed, there are typically 10 per cent fewer accidental injuries and 20 per cent fewer fatal accidents (OECD/ECMT, 2006). In the first quarter of 2014, there occurred road accidents reaching 40,283 in the country which had resulted in 10,575 deaths and RMB 210m of direct property losses. The number of accidents caused by high-speed driving accounted for 5.5 per cent of the total (Traffic Management Bureau of the Public Security Ministry, 2014).

4.1.5 Hard deceleration, acceleration and swerve maneuvers. Hard deceleration, acceleration and swerve maneuvers means the sudden acceleration/braking/turning behavior. These behaviors would affect the vehicle technical condition, so the vehicle is prone to occur security risks. A study carried out by the American Progress company shows that the driver's driving cost of high-risk driving behaviors with hard deceleration, acceleration and swerve maneuvers are about 2.5 times than the low risk's (Progressive Insurance, 2014). Jun carried out an empirical investigation to determine if drivers with a crash experience have driven differently in terms of speed, time of day and roadway types, hard deceleration. He found that crash-involved drivers had usually traveled longer mileage, normally traveled at higher speeds than non-crash drivers and frequently engaged in hard deceleration events (Jun *et al.*, 2007). Klauer classify the drivers into two groups: one is the drivers of safe driving and the other one is unsafe. The latter one has a higher frequency of hard deceleration, acceleration, and swerve maneuvers. They found that the accident rate of the latter is 7.4 times the former. The results of the analysis indicate that during baseline driving, unsafe drivers turned their vehicles at greater than 0.30 g, decelerated greater than 0.30 g and swerved greater than 3 ft/s significantly more frequently than either the moderately safe or safe drivers (Klauer *et al.*, 2009).

4.2 Establishing the indices system

Based on the five factors influencing driving safety and the data from OBD, in this paper, ten variables are selected as the evaluation indices including the monthly total mileage, peak time on weekday, night and weekend time, the time rate (80-120 km/h, >120 km/h), times of violations, hard deceleration, acceleration and swerve maneuvers. The multi-level driving behavior evaluation indices system is set up based on these indices, as is shown in Table I. The data of these variables we have chosen can be collected from the combination of OBD.

4.3 Determining the weights

It is very important to reasonably determine the weight of each index in the progress of establishing the scoring model. In this paper, an improved EW-AHP method is used to calculate the weights of each index.

4.3.1 Improved entropy weight analytic hierarchy process. The traditional EW-AHP is a simple combination to get weight of the bottom indices which is got through AHP and entropy weight method. There will be an imbalance comprehensive weight because of greater difference of values obtained by the two methods. To avoid the above shortcomings, this paper adopts improved EW-AHP method which combines both intermediate calculation process to get the final weights of the indices. Not only the data itself of the method can be reflected but also the method meets the actual application demands (Guo *et al.*, 2014).

4.3.2 The method of weight calculation based on entropy weight analytic hierarchy process. First of all, suppose that the number of upper level criterion and sub criterion are m and n. Each upper criterion consists of following variables: $\boldsymbol{n_1, n_2, \ldots, n_m}$. Through the judgment matrix of AHP method, the weights of upper level criteria and sub criteria are achieved, respectively, as $B = \{\beta_1, \beta_2, \ldots, \beta_n\}$ & $= \{\gamma_1, \gamma_2, \ldots, \gamma_n\}$.

Next, suppose that the weight of each index by EW as $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

Then, integrating the sub criteria weight D and EW A to get the comprehensive weights in sub criteria level as $\tau = (\tau_1, \tau_2, [\ldots] \tau_n)$, and:

$$\tau_i = \frac{\alpha_i \gamma_i}{\sum_{i=1}^{n} \alpha_i \gamma_i} \qquad (\mathrm{i} = 1, 2 \ldots, \mathrm{n}) \tag{1}$$

According the correspondence of sub criteria and upper criteria, the comprehensive weight is carried out again as $T = \{\tau_{11}, \tau_{12}, \ldots, \tau_{1n_1}, \tau_{21}, \tau_{22}, \ldots, \tau_{2n_2}, \quad \tau_{m1}, \tau_{m2}, \ldots, \tau_{mn_m}\}$. And normalization processing each sub level criterion and getting the weight as $\mu = \{w_{11}, w_{12}, \ldots, w_{1n_1}, w_{21}, w_{22}, \ldots, w_{2n_1}, w_{m1}, w_{m2}, \ldots, w_{mn_n}\}$ and:

$$w_{ij} = \frac{\tau_{ij}}{\sum_{i=1}^{n} \tau_{ij}} \tag{2}$$

($\mathrm{i} = 1, 2, \ldots \mathrm{n}; \mathrm{j} = 1, 2, \ldots \mathrm{m}; \mathrm{k} = n_1, n_2, \ldots, n_m$)

Multiplying the weight B (upper level) and μ (corresponding comprehensive weight) to get the new weight as $\mu' = \{w'_{11}, w'_{12}, \ldots, w'_{1n_1}, w'_{21}, w'_{22}, \ldots, w'_{2n_1}, w'_{m1}, w'_{m2}, \ldots, w'_{mn_n}\}$, and:

Target level	First class	Second class	Variable description
Driving behaviors scoring	Mileage and time	Monthly total mileage	Continuous
		Weekday peak time	Continuous
		Night driving time	Continuous
		Weekend driving time	Continuous
	Speeding time rate	80-120 km/h	Continuous
		>120 km/h	Continuous
	Different driving condition times	Sudden acceleration times	Continuous
		Sudden braking times	Continuous
		Turning times	Continuous
		Violations times	Continuous

Table I.
Behavior evaluation indices system

$$w'_{ij} = \beta_i w_{ij} \tag{3}$$

[i = 1, 2,. . ., n; j = 1,2,. . ., k; k ∈ ($n_1, n_2, \ldots, n_m$)]
Reformulating the μ into $\mu = \{w_1, w_2, \ldots, w_n\}$ and normalizing processing to get w = $\{w_1, w_2, \ldots, w_n\}$, and:

$$w_i' = \frac{w_i'}{\sum_{i=1}^{n} w_i'} \tag{4}$$

4.3.3 The results of weight. A questionnaire on importance of driving behavior indices has been issued to experts in the transport filed and insurance companies to get the evaluation of the importance of each factor of driving behavior. On this basis, through formula (1) to (4), improved EW-AHP method is used to calculate the weight of each index. At the same time, AHP and EW were used to calculate the indices. Compare with the three methods the results are shown in Table II.

From the results above, we found that in AHP method's calculation, the speed accounts for a large proportion. It is entirely subjective. In EW method's calculation, the time rate (80-120 km/h) weight is larger than the weight of time rate (>120 km/h). Obviously, it is unrealistic. The EW-AHP method taking not only subjective into consideration but also the importance of the other indices. So at last, we choose the EW-AHP methods calculation.
Finally, based on the weights calculated by the improved EW-AHP method, it is converted into a percentile system to obtain the scores of the second-class in the driving behavior scoring model.

5. Building the driving behavior evaluation indices system

To apply the model in practice, on the basis of determining the indices and weights, we need to set a number of options for each index. The driver's score is evaluated according to the options, and the scores of the indices are accumulated as the final scores.

Determining the option and its value. First, the sample data of each factor collected have been analyzed statistically. On the basis of last step and expert opinions, Alternative answers of driving behavior scoring model as well as scores are determined. The details are shown in Table III.

Indices	AHP weight	EW weight	EW-AHP weight	Scores
Monthly total mileage $\boldsymbol{\mu_1}$	0.0273	0.1039	0.0765	8
weekday peak time $\boldsymbol{\mu_2}$	0.0162	0.076	0.0462	5
Night driving time $\boldsymbol{\mu_3}$	0.0639	0.0682	0.1044	10
Weekend driving time $\boldsymbol{\mu_4}$	0.0098	0.1459	0.0401	4
Time rate (80-120 km/h) $\boldsymbol{\mu_5}$	0.1024	0.1207	0.1543	15
Time rate (>120 km/h) $\boldsymbol{\mu_6}$	0.512	0.1016	0.2601	26
Sudden acceleration times $\boldsymbol{\mu_7}$	0.0441	0.0879	0.0749	7
Sudden braking times $\boldsymbol{\mu_8}$	0.0282	0.0595	0.0381	4
Turning times $\boldsymbol{\mu_9}$	0.0764	0.0913	0.0804	8
Violations times $\boldsymbol{\mu_{10}}$	0.1196	0.1451	0.125	13
Total	1.0	1.0	1.0	100

Table II. The weight calculation results

Scoring Indices	Scores	Options	Values	Scores
1. Mileage and time	27	–	–	–
Monthly total mileage	8	≤100 km	8	–
		(100, 500) km	7	–
		(500, 1,000) km	5	–
		(1,000, 1,500) km	4	–
		(1,500, 2,000) km	3	–
		(2,000, 2,500) km	2	–
		≥2,500 km	1	–
Weekday peak time	5	≤5 h	5	–
		(5, 10) h	4	–
		(10, 20) h	3	–
		(20, 30) h	2	–
		≥30 h	1	–
Night driving time	10	≤1 h	10	–
		(1, 2) h	9	–
		(2, 4) h	8	–
		(4, 6) h	7	–
		(6, 10) h	6	–
		(10, 15) h	5	–
		(15, 20) h	4	–
		(20, 25) h	3	–
		≥25 h	1	–
Weekend driving time	4	(0, 5) h	4	–
		(5, 10) h	3	–
		(10, 15) h	2	–
		(15, 20) h	1	–
		≥20 h	0	–
2. Speeding time rate	41	–	–	–
80-120 km/h	15	0	15	–
		(0, 0.5%)	14	–
		(0.5, 1%)	13	–
		(1%, 3%)	11	–
		(3%, 6%)	9	–
		(6%, 9%)	7	–
		(9%, 12%)	5	–
		(12%, 15%)	3	–
		≥15%	1	–
>120 km/h	26	0	26	–
		(0, 0.1%)	24	–
		(0.1%, 0.5%)	22	–
		(0.5%, 1%)	20	–
		(1%, 2%)	17	–
		(2%, 3%)	14	–
		(3%, 4%)	10	–
		(4%, 5%)	6	–
		≥5%	1	–
3. Different driving condition times	32	–	–	–
Sudden acceleration times	7	0 times	7	–
		(1, 5) times	6	–
		(5, 10) times	5	–
		(10, 20) times	4	–
		(20, 30) times	3	–

(continued)

Table III.
Driving behaviors scoring model

Scoring Indices	Scores	Options	Values	Scores
		≥30 times	1	–
Sudden braking times	4	≤5 times	4	–
		(5, 10) times	3.5	–
		(10, 20) times	3	–
		(20, 35) times	2	–
		(35, 50) times	1	–
		≥50 times	0	–
Turning times	8	0 times	8	–
		(1, 5) times	7	–
		(5, 10) times	6	–
		(10, 20) times	5	–
		(20, 30) times	4	–
		(30, 40) times	3	–
		(40, 50) times	2	–
		≥50 times	0	–
Violations times	13	0 times	13	–
		1 times	10	–
		(2, 3) times	6	–
		(3, 5) times	2	–
		≥5 times	0	–

Table III.

6. Verify the model by field experiments

A traffic communication company with a property insurance company in Chongqing launch a UBI field experiment to promote the installation of OBD products and collect data of customers' driving behavior. From October 2014 to April 2015, 165 users of OBD have been installed. The driving behavior data collected by OBD mainly contains monthly total mileage, weekday peak driving time, night driving time, weekend travel time, speeding time ratio (80-120 km/h, >120km/h), hard deceleration, acceleration, swerve maneuvers and the number of violations. And the insurance has the drivers' personal information including premium, vehicle model and claim frequency in the past year. Through the integration of customer personal data and driving behavior data, removing abnormal and invalid data, they get driving behavior data from 100 customers.

6.1 The results of driving behaviors scoring

First, import the customers' behavior data of 100 drivers into scoring model. Then, calculate scores of driving behavior of each customer. Finally, statistical analysis is done to accident cases during the experiment time to get a cross-reference table of driving behaviors scoring and history accident times, as is shown in Table IV (only listing 4 highest scores and the 6 lowest score).

6.2 A correlational analysis between driving behaviors scoring and the numbers of history accidents

Based on Table IV, a correlational analysis of Spearman is performed (Owning to the limitation of the scope, the table of analysis results will not show in this paper), as is shown in Table V.

Based on Table V, the correlation coefficient between the score and the number of accidents is -0.504, which means a negative correlation. The unrelated bilateral significance was $0 < 0.01$. So there was significant negative correlation between scores and accident times.

6.3 A correlational analysis between driving behaviors scoring and history accidents
Carry out a subtotal from Table IV and get the result in Table VI (Mean accident times = Total accident times/Number of people).

After the statistical analysis of the Table V. First, make y = Mean Accident Times and x = Scores. Then fit the relational model between them. The results are shown in Table VII and Figure 1.

From Table VII and Figure 1, the best fitting function is quadratic function model (R^2 = 0.998), followed by Logarithm model. From the F value, the best fitting is quadratic function (F = 693.415). The sig. value of the logarithm model and the quadratic model is 0, which shows that the model is more significant and has better predictive capability.

Here is the quadratic model between mean accident times and driving scores:

$$y = 2.713 - 0.713x + 0.058x^2 \tag{5}$$

Through the correlational analysis and model fitting, there is a significant correlation between the score and accident times; the higher the driving score is, the less the accident times. Therefore, the model can effectively reflect the driver's risk level of driving safety and has a great significance in practice.

7. Conclusion

UBI model is the trend of automobile insurance rates at home and abroad. In the DBRF pricing model, driver behavior evaluation is essential for the individual insurance

Table IV.
Driving behaviors scoring and accident times

Serial no.	Scores	Accident time
1	91	0
2	73	1
3	37	1
4	68	0
...	...	...
95	72	0
96	85	0
97	87	0
98	63	2
99	60.5	0
100	88.5	0

Table V.
Correlation coefficient

		Scores	Accident times
Spearman			
Correlation coefficient	1.000	−0.504**	
Scores	Sig. (double)		0.000
	N	100	100
Accident times	Correlation coefficient	−0.504**	1.000
	Sig. (double)	0.000	
	N	100	100

Note: **When the confidence (double measure) was 0.01, the correlation was significant rating for automobile. First of all, this paper summarized development status at home and abroad of UBI. And pointed out that one of the important data sources is from OBD. Next, we studied five principle driving behavior factors, and the driving behavior evaluation indices system was designed based on the five factors. After the indices weights were determined by the improved EW-AHP method, a DBRF system of UBI was established. Finally, the OBD driving behavior data of 100 customers is collected through field experiments. The score of each customer is calculated by scoring model and a statistical analysis is conducted. Final result shows that the DBRF can provide a basis for individual insurance rate so as to improve automobile insurance pricing model

Scores	No. of people	Total accident times	Mean accident times
1 (<40)	6	9	1.5
2 (40 ~ 60)	10	10	1
3 (60 ~ 70)	22	13	0.59
4 (70 ~ 80)	33	7	0.21
5 (80 ~ 90)	23	1	0.043
6 (>90)	6	0	0

Table VI. Subtotal scores of driving behaviors

	Model statistic					Parameter estimation		
Equation	R^2	F	df1	df2	Sig.	Constant	b1	b2
Linear	0.927	51.114	1	4	0.002	1.633	−0.307	
Logarithm	0.987	292.499	1	4	0.000	1.538	−0.894	
Quadratic	0.998	693.415	2	3	0.000	2.173	−0.713	0.058

Table VII. The fitting results of driving behaviors scoring and mean accident times

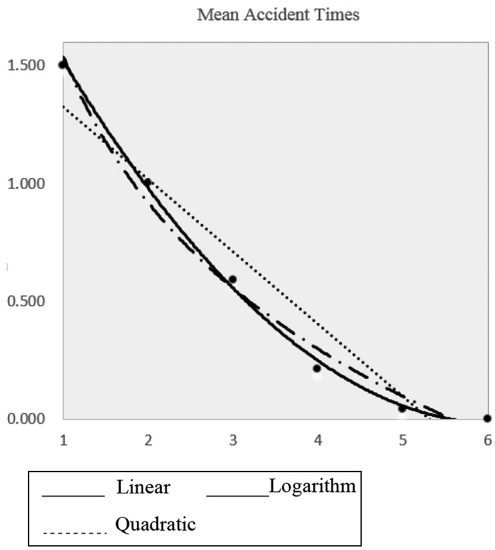

Figure 1. The fitting graph of driving behaviors scores and mean accident times

and to optimize the rate structure. At meanwhile, this will help to enrich insurance pricing theories of internet of vehicles.

In general, we acknowledge the usual limitations that apply to case study research. The UBI is initial and used infrequently in domestic. The scale of data we used in the paper is not big enough. However, with the expansion of the UBI products, the data can be collected in a large scale. Although we choose five important factors to build the driving behavior model, there is a great deal of factors influencing the driver's driving behavior, such as gender, place, drunk driving and illegal driving. If you need a more accurate model, you can consider the other influence factors. And because of the fairness and multiple additional functions (such as help to find the loss car, make drivers driving behavior better) of the UBI, UBI will develop rapidly in China.

References

Abdel-Aty, M., Uddin, N. and Pande, A. (2005), "Split models for predicting multivehicle crashes during high-speed and low-speed operating conditions on freeways", *Transportation Research Record Journal of the Transportation Research Board*, Vol. 1908 No. 1, pp. 51-58.

Bordoff, J.E. and Noel, P.J. (2008), "Pay-as-you drive auto insurance: a simple way to reduce driving-related harms and increase equity", Hamilton Project: discussion papers.

Boucher, J. (2013), "Pay-as-you-drive insurance: the effect of the kilometers on the risk of accident", *Anales Del Instituto De Actuarios Españoles*, Vol. 19, pp. 135-154.

Campbell, K. (2003), "Detailed planning for research on making a significant improvement in highway safety: study 2-safety", *F-SHRP Web Document 2(NCHRP Project 20-58[2]): Contractor's Final Report*, Transportation Research Board of The National Academics.

Davis, G.A., Davuluri, S. and Pei, J.P. (2006), "A Case Control Study of Speed and Crash Risk".

Dimitros, I., Tselentis, G., Yannies, E. and Vlahogianni, I. (2016), "Innovative insurance schemes: pay as/how you drive", *Transportation Research Procedia, 2016*, Vol. 14, pp. 362-371.

Elvik, R., Christensen, P. and Amundsen, A. (2004), "Speed and Road Accidents: An Evaluation of the Power Model", TØI report 740/2004, Institute of Transport Economics TOI, Oslo.

Ferreira, J. and Minike, E. (2012), "Measuring per mile risk for pay-as-you-drive automobile", *Transportation Research Record Journal of the Transportation Research Board*, Vol. 2297 No. 2297, pp. 97-103.

Greenberg (2008), "Lowering the patent barrier to pay-as-you-drive insurance", Transportation Research Board Meeting.

Guo, J., Fu, X., Gao, X., *et al.* (2014), "An improved weight calculation method for multi objective decision making", *Journal of Xidian Univerisity(Natural Science Edition)*, Vol. 41 No. 6, pp. 118-125.

Hallvig, D., Anund, A., Fors, C., Kecklund, G. and Åkerstedt, T. (2014), "Real driving at night – predicting lane departures from physiological and subjective sleepiness", *Biological Psychology*, Vol. 101 No. 3, pp. 18-23.

Joksch, H.C. (1993), "Velocity change and fatality risk in a crash – a rule of thumb", *Accident; Analysis and Prevention*, Vol. 25 No. 1, pp. 103-104.

Jun, J., Ogle, J. and Guensler, R. (2007), "Relationships between crash involvement and temporal-spatial driving behavior activity patterns: use of data for vehicles with global positioning systems", *Transportation Research Record*, Vol. 2019, pp. 246-255.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D. and Ramsey, D.J. (2009), "Comparing Real-world Behaviors of Drivers with High versus Low Rates of Crashes and Near-crashes, National Highway Traffic Safety Administration", Report No. DOT HS 811 091.

Langford, J., Koppel, S., McCarthy, D. and Srinivasan, S. (2008), "In defence of the 'low-mileage bias", *Accident; Analysis and Prevention*, Vol. 40 No. 6, pp. 1996-1999.

Lenné, M.G. and Jacobs, E.E. (2016), "Predicting drowsiness-related driving events: a review of recent research methods and future opportunities", *Theoretical Issues in Ergonomics Science*, Vol. 17 No. 5, pp. 1-21.

Li, H. and Li, L. (2015), "Analysis of unsafe driving behavior based on questionnaire", *Automobile Applied Technology*, Vol. 2015 No. 2, pp. 145-148.

Litmant (2011a), *Pay-as-You-Drive Insurance: Recommendations for Implementation*, Victoria Transport Policy Institute.

Litmant (2011b), "Distance-Based Vehicle Insurance Feasibility, Benefits and Cost: Comprehensive Technical Report", Viloria Transport Policy Institute.

Litmant, T. (2005), "Efficient vehicles versus efficient transportation comparing transportation energy conservation strategies", *Transport Policy*, Vol. 12 No. 2, pp. 121-129.

Mccartt, A.T., Rohrbaugh, J.W., Hammer, M.C. and Fuller, S.Z. (2000), "Factors associated with falling asleep at the wheel among long-distance truck drivers", *Accident Analysis & Prevention*, Vol. 32 No. 4, pp. 493-504.

Maclean, A.W., Davies, D.R.T. and Thiele, K. (2003), "The hazards and prevention of driving while be sleepy", *Sleep Medicine Reviews*, Vol. 7 No. 6, pp. 507-521.

Martain, J.L. (2002), "Relationship between crash rate and hourly traffic flow on interurban motorway", *Accident Analysis and Prevention*, Vol. 34 No. 5, pp. 619-629.

Matthew, T.F. and Climatology, D.F. (2004), "Chicago O'hare international airport July 1996-April 2002", *Bulletin of the American Meteorological Society*, Vol. 85 No. 4, pp. 515-517.

Nilsson, G. (2004), "Traffic safety dimensions and the power model to describe the effect of speed on safety", Bulletin 221, Lund Institute of Technology, Department of Technology and Society, Traffic Enginering, Lund.

No, O., Carrol, A., Multer, J., *et al.* (2003), "Research and Special Programs Administration: Commerial Transportation Operator Fatigue Management Reference", US Department of Transportation 2003.7.

OECD/ECMT (2006), "Speed management"

Paefgen, J., Staake, T. and Fleisch, E. (2014), "Multivariate exposure modeling of accident risk: insights from pay-as-you-drive insurance data", *Transportation Research Part A: Policy and Practice*, Vol. 61, pp. 27-40.

Paula, M. and Wendy, W. (2010), "The Relationship between Road Safety and Congestion on Motorways".

Peng, J., Liu, N., Zhao, H. and Yu, M. (2016), "Research on smart usage-based insurance system", *Computer Technology and Development*, Vol. 27 No. 1, pp. 142-146.

Progressive Insurance (2005), "Texas Mileage Study: Relationship between Annual Mileage and Insurance Losses, 2005".

Progressive Insurance (2014), "Linking Driving Behavior to Automobile Accidents and Insurance Rates: An Analysis of Five Billion Miles Driven", USA, 2014.7.

Russell Henk, P.E., Pezoldt, V. and Fette, B. (2010), "An analysis of fatal crashes under dark conditions in the US, 1999-2008", *Shedding Light on the Nighttime Driving Risk*, Texas Transportation Institution.

Sinisa, H., Perakovic, D., Forenbacher, I. and Mumdziev, M. (2015), "Telematics system in usage based motor insurance", *Procedia Engineering*, Vol. 100, pp. 816-825.

Staplin, L., Gish, K.W. and Joyce, J. (2008), "Low mileage bias' and related policy implications – a cautionary note", *Accident Analysis & Prevention*, Vol. 40 No. 3, pp. 1249-1252.

Stuttsa, J.C., Wilkinsb, J.W., Osbergc, J.S. and Vaughnb, B.V. (2003), "Driver risk factors for sleep-related crashes", *Accident Analysis and Prevention*, Vol. 35 No. 3, pp. 321-331.

Sun, L. (2017), "The status quo of the industry and some suggestions for operators", Guangdong communication technology.

Traffic Management Bureau of the Public Security Ministry (2014), "Annual Report on Road Traffic Accident in China in 2014".

Vickrey, W. (1968), "Automobile accidents, tort law, externalities, and insurance: an economist's critique", *Law & Contemporary Problems*, Vol. 33 No. 3, pp. 464-487.

Williams, K.M. (2001), "Accidents and hours-of-service violations among over-the-road drivers", *Transportation Quarterly*, Vol. 40 No. 1.

Yanagihara, M., Uno, N. and Nakamura, T. (2015), "Latent class analysis for driving behavior on merging section", *Transportation Research Procedia*, Vol. 6, pp. 259-271.

Further reading

Amamdan, R. (2014), "Developments in the pricing of auto insurance based upon driving behavior", Risk Management & Insurance Program.

Chenghai, H., Dao, Y. and Sujin, Z. (2016), "The Automobile Insurance Pricing Model Combining Static Premium with Dynamic Premium".

Clarles, M.F. (2016), "Relationship of Traffic Fatality Rates to Maximum State Speed Limits", Traffic Injury Prevention.

Laas, D., Schmeiser, H. and Wagner, J. (2016), "Empirical findings on motor insurance pricing in Germany, Austria and Switzerland", *The Geneva Papers on Risk and Insurance - Issues and Practice*, Vol. 41 No. 3, pp. 398-431.

Lee, F.S. and Hellinga, B. (2002), "Analysis of crash precursors on instrumented freeways", *Transportation Research Record Journal of the Transportation Research Board*, Vol. 1784 No. 1.

Lianzeng, Z. and Baige, D. (2012), "Research on effect of mileage on the net insurance premium", *Insurance Studies*, Vol. 6, pp. 29-38.

Mu, Q. (2015), "Development status and trend of UBI in big data context", *The Industrial Study*.

Shuzhan, H., Xiaorui, S., Yyulong, H., *et al.* (2011), "Study on the relationship between traffic accident severity and traffic flow characteristics of freeway", *China Safety Science Journal*, Vol. 21 No. 9, pp. 106-112.

About the authors

Dr Zhishuo Liu (1977) is an Associate Professor at Beijing Jiaotong University and a Visiting Scholar at University of California, Irvine (from February 2016-February 2017). His areas of research include control science and engineering, logistics engineering, e-commerce technology and application.

Qianhui Shen (1995)is a Postgraduate student at Beijing Jiaotong University. Her areas of research include control science and engineering. Qianhui Shen is the corresponding author and can be contacted at: 16120772@bjtu.edu.cn

Jingmiao Ma (1990) is a Postgraduate student at Beijing Jiaotong University. Her areas of research include E-commerce technology and application.