The Automobile Insurance Pricing Model, Combining Static Premium with Dynamic Premium ——Based on the Generalized Linear Models

Chenghui Han   Dan Yao   Sujin Zheng
School of Insurance,
Central University of Finance and Economics
Beijing, China

Abstract

Since 2015, the reformation of automobile insurance has been restarting in China, which makes the rate of the automobile insurance more and more marketable. In this background, with the development of the Telematics technology, the dynamic factors will be added into the automobile insurance pricing model so that the automobile insurance pricing will become more variable. We combine the static premium with the dynamic premium and built a pricing model based on the Generalized Linear Models (GLM), then compare it with other automobile insurance pricing models: the model considering the static factors such as car models, driving zone and so on; the static premium model considering the factor of driving kilometers; and the dynamic premium model. We discuss the difference between these models in the imitative effect, the influence on cash flows of insurance companies, and so on. Finally, we try to give some advice to the future development of automobile insurance pricing in China.

Key words: market-based pricing, dynamic premium, GLM

1. Introduction of Chinese Automobile Insurance Market

The automobile insurance in China appeared in the 1950s, but since the 1960s this business had been discontinued, with the background of discontinuing all the insurance business in China, until 1980 when it was resumed.

Since China's access to the WTO, the marketization reform of Chinese auto insurance rate has been started. There are some twists and turns in the reform, in total, there are 4 important periods as shown in Table 1.

Table 1 Important Events in Chinese Marketization Reform

Period	**Time**	**Event**
Infancy	Jul. 2001	Floating premium rate was put into practice in Shenzhen.
	Oct. 2001	Marketization reform started in Guangzhou.
	Sep. 2002	CIRC started to accept clauses and rates from companies.
Promotion	Jan. 2003	New system of auto insurance rates started nationwide.
	Mar. 2006	CIRC limited the discount of rates, not low than 70%.
	Jul. 2006	CIRC limited the clauses of contracts, only A, B, C three kinds of products.
	Feb. 2007	Chinese Insurance Industry Association published the unified clauses and rates of common additional risks.
Stagnation	–	–
Restarting	Jun. 2010	Marketization reform restarted in Shenzhen.
	Mar. 2012	Marketization reform restarted nationwide.
	Jan. 2013	CIRC published some documents about the marketization reform to promote it.
	Feb. 2015	
	Mar. 2015	

Before 2001, the rate of auto insurance premium in China had been in a strict regulation, where all the insurance companies used a unified premium rate.

Since 2001, CIRC (China Insurance Regulatory Commission) started the marketization reform of auto insurance premium rate. CIRC hoped to give the insurance companies more freedom and rights to price the auto insurance. However, when the companies could price auto insurance by themselves, there was a large price war in the market of auto insurance. The rates of all the insurers are too low to make them loss. This price war started in 2003, but didn't end until 2007, when CIRC gave up that reform and regulated strictly again.

There was a stagnation period from 2007 to 2010. In 2010, the marketization reform of auto insurance premium rate restarted in Shenzhen. And after that, CIRC published a series of documents to promote the restart of marketization reform.

The marketization of premium rates must make an intense competition of the auto insurance price. In such a competitive market, the insurance companies can't profit from the lower and lower prices, and they should try other methods. For example, if an insurer can improve its technology of pricing, give customers different premiums based on their risks, this insurer may have some advantages in the competitive market.

Many Chinese insurance companies don't have such a pricing technique, because they use and have to use the unified premium rate all the time. We think that with the marketization reform, these insurance companies should and must improve their technology of pricing.

2. Development of Automobile Insurance Premium

The insurance market competition environment is closely related to the development of insurance products pricing, and with the development of insurance pricing technology, the form of premium is changing. In general, there is a growing trend from static premium to dynamic premium.

In this paper, it is defined that static premium is the form of premium which won't change with the usage of insured in an insurance period; and dynamic premium is the form of premium which changes with the usage of insured all the time.

2.1 Traditional Static Premium

Traditional static premium is purely static, both in the period of insurance and in the consideration of pricing factors. The factors of pricing static premium are all static factors, including person's factors as age, gender and car's factors as car kind, new car to purchase value, vehicle type. Because of the limitation of technology, the usage factors as kilometers, speed has not been considered.

In order to make the premium more close to the risk level of the insured, insurance companies use the No-Claim-Discount. However, this form of premium with the posterior information still belongs to traditional static premium.

2.2 Static Premium with Dynamic Factors

Kilometers of driving is a direct reflection of automobile risk exposure. And insurance companies always tried to take the kilometers into consideration of auto insurance pricing.

The most close to traditional method of auto insurance pricing, the kilometers can be treated as a pricing factor of static premium. In the past, because of the lack of actual kilometers, insurers always used estimated kilometers as a replacement of actual kilometers. In this method, like the factors such as age and gender, the kilometers factor was classified into certain intervals. The premium adjustment coefficient of each interval of kilometers could be calculated by a pricing model, like GLM, or decided by actuaries directly.

However, there are some weaknesses of the methods to use kilometers as a factor in static premium of auto insurance pricing. The estimated kilometers were estimated by insured, and the insured always underestimated their own kilometers in order to get a lower premium. So the data of estimated kilometers had a low accuracy. As a result, insurance companies always controlled the weight of kilometers factor in practice. If the coefficients were calculated by pricing model, the weight was difficult to control in the model; and if the coefficients were decided directly, the weight was easy to control, but the coefficients may be not reasonable enough.

2.3 Dynamic Premium

The thought of pricing auto insurance dynamically according to kilometers had been raised in the 1920s. In practice, there are three kinds of dynamic premium based on kilometers.

2.3.1 Pay-At-The-Pump (PATP)

PATP needed insurance companies had a cooperation with filling stations, and insurance companies would price auto insurance based on the usage of fuel oil and claims data. Insurance companies should calculate the premium per liter fuel oil, then add the premium into the price of fuel oil, so that a car would pay different premium based on its usage of fuel oil.

Compared with treating the kilometers as a factor of static premium directly, PATP translated static premium to dynamic premium. Using fuel oil as a replacement of actual kilometers, solved the problem of lack of actual kilometers data, and guaranteed the accuracy of data at the same time. On the other hand, adding premium into the price of fuel oil would lower the frequency of insured driving, lower carbon emission, so that it had a positive externality.

But there were even some weaknesses. First, the usage of fuel oil was the only pricing factor in this method, and traditional factors, such as age, gender, vehicle type, and so on, weren't taken into consideration. Second, because auto insurance and fuel oil were offered by many different insurance companies and oil companies, and the insured may refuel at different stations, so there were some difficulties to promote PATP in practice.

2.3.2 Pay-As-You-Drive (PAYD)

The appearance of PAYD is closely related to development of telematics technology. With the development of telematics and GPS, insurance companies can use these techniques to get the data of actual kilometers of the insured vehicles in insurance period. PAYD uses the actual kilometers of the insured driving, which guarantees the accuracy of data. The premium of PAYD for insured with different risk levels is different. At the same time, insurance companies don't need to abandon traditional pricing factors because they can calculate the premium per kilometers directly.

The premium of PAYD is also dynamic like PATP. In this method, the insurer uses the data of age, gender, vehicle type, kilometers factor, and so on, to calculate the premium per kilometer of different insured, and in the end, the insurer will calculate the total premium based on the actual kilometers of insured.

In practice, there are many different forms of PAYD. First, the insured can buy certain kilometers, and in a certain insurance period, the kilometers exceeding are not covered. This form may lower the underwriting and can't compensate the risk. Second, the insured can buy a minimum level of kilometers, then pay for the extra premium based on actual kilometers in the end of an insurance period. This form may make the insurer face higher moral hazard, and if insurer don't pay premium in the end of period, the insurer may not get the enough premium, then it may influence the financial stabilization of the insurer. Third, the insured should pay for certain kilometers based on their own estimation at the beginning of an insurance period, and the PAYD will cover all actual kilometers of insured, but the extra actual kilometers will be fined in the end of the insurance period. This form is the combination of the first and second form, so that it will encourage insured to buy enough kilometers covered and protect the financial stabilization of insurer.

2.3.3 Pay-How-You-Drive (PHYD)

PHYD is similar to PAYD, and the insured will pay a dynamic premium which is calculated based on the data from telematics. However, the factors in PHYD pricing are more than PAYD. There are not only the actual kilometers, but also emergency brake, sudden turn, overspeed and other driving behaviors in consideration of pricing. And the insurance companies provide more service by telematics when the insured is driving.

The development of PHYD is closely related to the technology of telematics. The insurer get the data about driving behaviors by the telematics control units, then judge the risk level of insured driving behaviors and calculate the dynamic premium. In practice, the insurer often scores the insured driving based on these data and charge a dynamic premium.

2.4 Combination of Static Premium and Dynamic Premium

Though the development of science and technology solved the technical problem in application of kilometers, and the insurance companies can charge dynamic premium, there are some limitations of these usage-based insurance (UBI) products, such as privacy, regulation, customer acceptance problems and so on.

In practice, the pure dynamic premium has been translating into the form of combination of static and dynamic premium. For example, the insurer decides the discount of rates based on driving behaviors or kilometers, then adjusts the discount regularly.

The UBI in China has been beginning now. Some large-scale property insurance companies, such as People's Insurance Company of China (PICC) and Ping An Property & Casualty Insurance Company of China (Ping An), have started to explore and have a try on the UBI. Now in Chinese market, all the premium of auto insurance is static premium. Considering the international experience, we insist that it is difficult to translate static premium to dynamic premium directly. So we suggest that the Chinese insurance companies can try to use the combination of static and dynamic premium, which will retain a certain weight static premium, and the other part of premium is decided dynamically based on driving behaviors and/or kilometers. This combination can help insurance companies keep finance stabilize and improve customer acceptance, on the other hand, the premium will have the advantages of dynamic premium such as high fairness and positive externality.

3. Models

As a core factor for measuring driving risk, the kilometers is very important for auto insurance pricing. With the development of UBI, it is clear that there is a trend that static premium will change to dynamic premium. This paper will use the generalized linear model (GLM) to discuss the difference between these forms of premium, in order to get some suggestions for Chinese UBI.

3.1 Generalized Linear Model (GLM)

3.1.1 Introduction of GLM

The core theory of GLM is to use the linear combination of a certain function of the risk ranking variables to explain the expectation of a loss variable. We can show it with the following formula:

$$E(Y_i) = \mu_i = g^{-1}(\eta_i) = g^{-1}(X_i\beta)$$

In this formula,

Y_i means the dependent variable, which can be frequency and severity of claims;

$g(\cdot)$ is called linked function;

X_i means the independent variable, which is a vector of risk ranking variables;

β is a coefficient vector of the risk ranking vector.

The keys in application of GLM are the choices of distribution function of dependent variable Y and the linked function $g(\cdot)$.

Based on the previous researches of many scholars, the linked function which is often used to estimate frequency and severity of claims is the logarithmic linked function. So we choose logarithmic function as our linked function directly.

Because of the non-negativity, the distribution of frequency and severity both can't be the normal distribution. We plan to compare some common distribution of frequency and severity, showing in the Table 2, and choose the better one to build the GLM.

Table 2 Assumption of Distribution and Linked Function to Compare

Frequency		**Severity**	
Distribution	**Linked Function**	**Distribution**	**Linked Function**
Poisson	Logarithmic	Gamma	Logarithmic
Negative Binomial	Logarithmic	Inverse Gaussian	Logarithmic

3.1.2 Model Test

- **To choose a distribution**

When we want to know the goodness-of-fit of the GLM, we often compare the used model with Saturated Model. The two models have the same assumptions of distribution and linked function, and their only difference is the number of independent variables: the saturated model has the maximum quantity of independent variables. If there is no or little difference between the goodness-of-fit of these two models, it means that the used model can explain the data well.

Using mathematical symbol language, we can assume the likelihood function of used model by $l(b;y)$, the likelihood function of saturated model by $l(b_{max};y)$. Then, Deviance (D) is defined as following:

$$D = 2 * [l(b_{max};y) - l(b;y)]$$

When the used model fits well, the expectation of deviance should be close to its freedom degree.

We build all the models by the R statistical software. The statistical results from R include the deviance of a null model (no independent variables) and a certain model, and we can denoted them by D_n and D_c. The difference between them is denoted as $\Delta D = D_n - D_c$, which obeys Chi-Square distribution with a freedom degree as $(n - \mathrm{c})$ (n and c are the freedom degrees of D_n and D_c).

So we can choose a better model by comparing the difference between the models and a null model, deviances of models, and the value of AIC.

- **To test a nested model**

The nested model is a model which has the same distribution but more independent variables than another model. When judging the significance of variables in nested model, we always use the deviances of two models.
We will assume that:

$$H_0: \ \beta = \beta_0 = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_q \end{pmatrix}$$

$$H_1: \ \beta = \beta_1 = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \ q < p < n$$

where n is the number of sample; q and p is the number of independent variables of primary model and nested model.
We can use the difference between their deviances: $\Delta D = D_0 - D_1$ to test the hypothesis. If the value of ΔD is bigger than the 95% quantile of the distribution χ^2_{p-q}, then we suggest that the nested model fits better.
We use this method to compare the model of static premium with or without the kilometer factor.

3.2 Data

Because the current Telematics technology in China is developing, and there is not enough public data, so we choose the foreign data, which give details of third party motor insurance claims in Sweden for the year 1977. The data were compiled by a Swedish Committee on the Analysis of Risk Premium in Motor Insurance. This data classified customers based on four factors for pricing, Kilometers (5 ranks), Zone (7 ranks), Bonus (7 ranks), and Type (9 ranks). There were Insured (about 2.38 million), Claim number and Payment statistics in the data, so that we could calculate frequency and severity. This data was the summarized data with 2,182 pieces in total.

3.2.1 To Add Person's Factor

There is no person's factor in this data, such as gender and age. For enriching the factors, we add the gender factor into this data.
In the data, the Insured, Claim number and Payment should be separated based on gender of drivers. We assume that the number of men and women is equal. There have been some research results about the frequency and severity of Chinese men and women (Zhang Y, 2007). The information is shown in the Table 3.

Table 3 Information of Auto Insurance Claim in China (From Zhang Y, 2007)

	Frequency	**Severity**	**Average Loss**
Male	17.30%	2,130	368
Female	19.80%	1,655	328

Based on Table 3, we can see that the frequency of female is higher, the severity of male is higher, and in total the average loss of male is higher. This conclusion is the same as common sense that the female drivers may have more small claims but few of large accidents.
So we can use the information to calculate the separated Claim number and Payment

by the following equations:

$$\begin{cases} m + n = c \\ x + y = p \\ \dfrac{m}{a} \Big/ \dfrac{n}{b} = \dfrac{17.3\%}{19.8\%} \\ \dfrac{x}{a} \Big/ \dfrac{y}{b} = \dfrac{2130}{1655} \end{cases}$$

where a and b are the numbers of male and female insured; m and n are the claim numbers of men and women; x and y are the payments of men and women; the total claim numbers is c; and the total payment is p.

After adding the gender factor into the data, there are 4,364 pieces in total, and next, we will use this adjusted data.

3.2.2 To Simulate Kilometers per Insured

In consideration of the model with dynamic premium, we need the actual kilometers of every risk groups. Because of the lack of actual kilometers data, we will make some assumption by stochastic simulation.

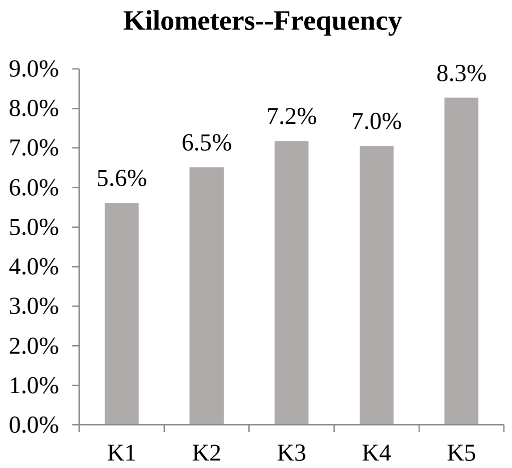
Table 4 Information of Kilometers in Data

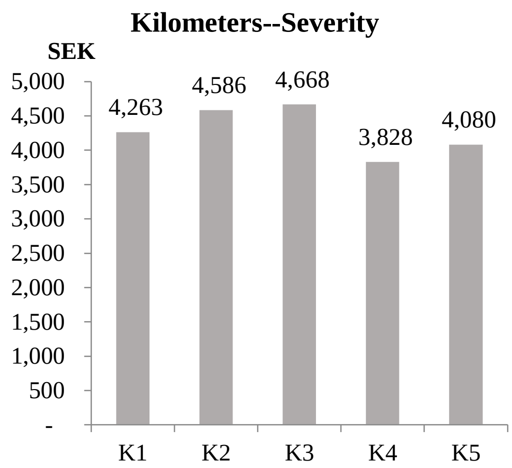
Rank	1	2	3	4	5
Kilometers Interval	< 1,000	1,000~15,000	15,000~20,000	20,000~25,000	> 25,000
Insured	806,801	804,397	477,149	173,150	121,673

Table 4 shows the classification of Kilometers factor. We can see that the range of kilometers is large and it is difficult to simulate a distribution of kilometers based on the insured. So we assume that there is the uniform distribution in each kilometers interval, and then the simulated actual kilometers can be calculated based on the insured in each interval.

3.3 Descriptive Statistics of Single Factor

Except the gender factor which is assumed (G0 – Female, G1 – Male), we analyze the frequency and severity of all the independent variables one by one by descriptive statistics of single factor. The results are shown in Picture 1-4.

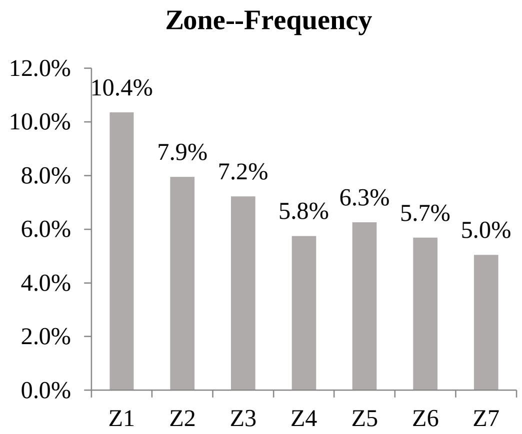


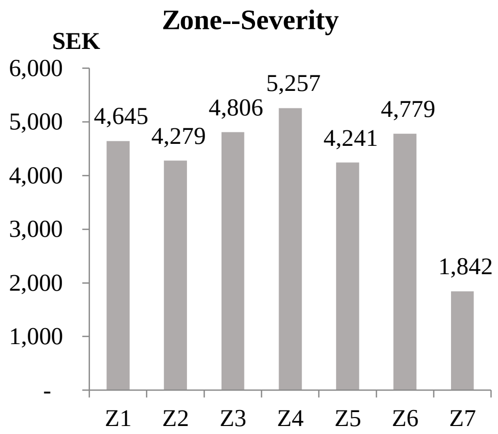


Picture 1 Loss Distribution (Classified by Kilometers)

In Picture 1, rank K1 – K5 mean fewer than 1000, 1000 – 15000, 15000 – 20000, 20000 – 25000 and more than 25000 kilometers.

Picture 1 shows that the more kilometers there are, the higher frequency is, and there is not a clear trend in the severity.

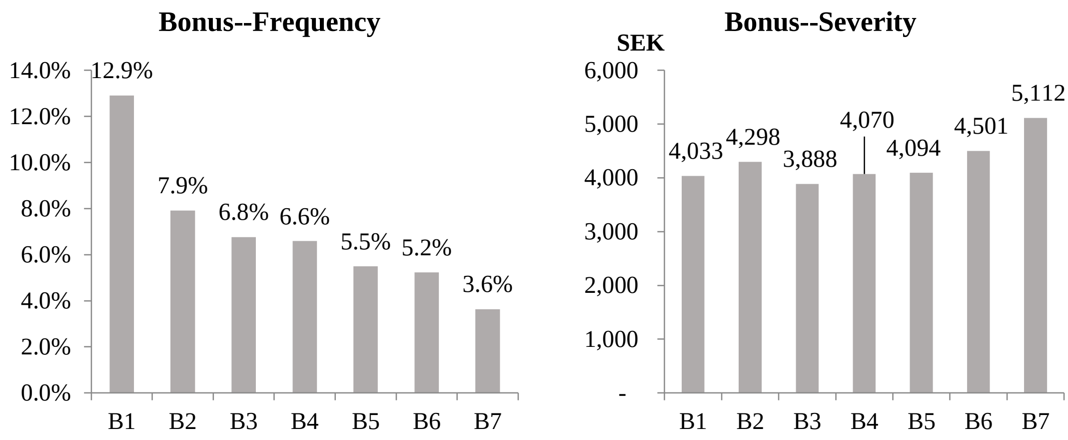




Picture 2 Loss Distribution (Classified by Zone)

In Picture 2, rank Z1 – Z7 mean the 7 zones of Sweden: Z1 – Stockholm, Göteborg, Malmö with surroundings; Z2 – Other large cities with surroundings; Z3 – Smaller cities with surroundings in southern Sweden; Z4 – Rural areas in southern Sweden; Z5 – Smaller cities with surroundings in northern Sweden; Z6 – Rural areas in northern Sweden and Z7 – Gotland.

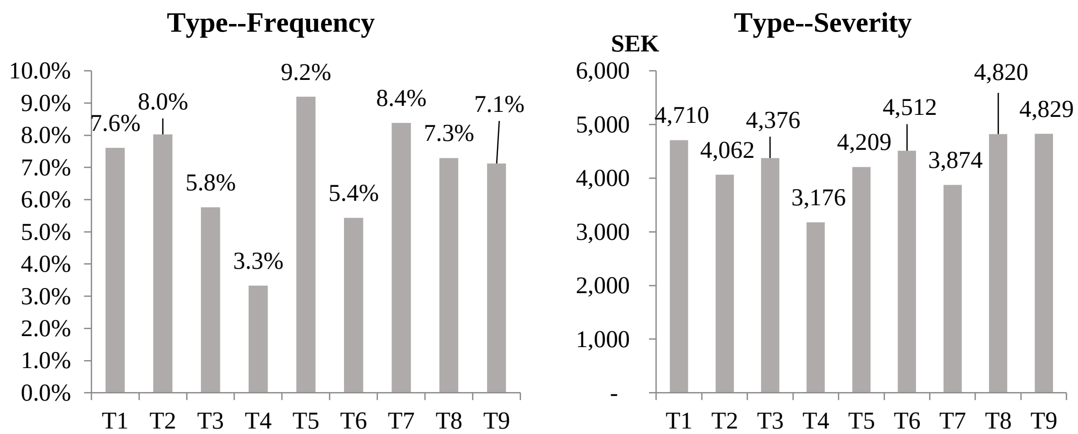
Picture 2 shows that the frequency of cities is higher than the one of rural areas, and frequency in southern Sweden is higher than northern Sweden. There are also differences between the severities, but the trend is not the same as frequency, which may be related to the labor cost and price in different zones.



Picture 3 Loss Distribution (Classified by Bonus)

In Picture 3, rank B1 – B7 mean the no claims bonus, equal to the number of years since last claim, and plus 1.

Picture 3 shows that the bigger the number of years since last claim are, the higher the frequency is, and the severities of the middle ranks are lower. We suggest the reasons may be that when these insured have an accident, the accident may be a large one, because they may want to hold the smaller claims in order to get a no-claims discount.



Picture 4 Loss Distribution (Classified by Type)

In Picture 4, rank T1 – T9 mean the types of vehicles, where T1 – T8 represent 8 different common vehicle types and T9 represent all other types.

Picture 4 shows that there are different frequencies and severities between different vehicle types. It is important to pricing the auto insurance with consideration of vehicle types.

3.4 Models and Analysis

Based on the adjusted data, we will build the models and analyze the results.

3.4.1 To Choose Distributions of Frequency and Severity

First, we compare the distributions of frequency and severity, and choose the better one. For the frequency, we build GLMs with static premium (including kilometers) based on Poisson and Negative Binomial distribution separately, and the other assumptions of them are completely the same. The results of statistic testing is shown in Table 5.

Table 5 Statistic test of Frequency (Different Distributions)

	Negative Binomial		**Poisson**	
	DF	Value	DF	Value
Null Deviance	4267	689733	4267	34877.6
Residual Deviance	4242	658110	4242	3244.1
ΔD	25	31623	25	31623.5
AIC	765714		–	

In Table 5, Null Deviance represents the deviance of a null model which includes intercept and deviant only; Residual Deviance represents the deviance of the model with static premium (not including kilometers); ΔD represents the difference between the two deviances; DF means the degree of freedom.

The values of ΔD of Negative Binomial model and Poisson model are very close, and their degrees of freedom are the same. It is obvious that the value of ΔD, 31623, is bigger than the 95% quantile of χ^2_{25}. So both Negative Binomial model and Poisson model are significant.

In consideration of the residual deviances of two models, the value of deviance of the Poisson model is more close to its degree of freedom. The value of deviance of the Negative Binomial is so much bigger than its degree of freedom, which is a result from the problems of this model. As a consequence, the Poisson distribution is better for our GLM of frequency.

For the severity, it is similar to frequency, but the distributions are Gamma and Inverse Gaussian. The results of statistic testing is shown in Table 6.

Table 6 Statistic test of Severity (Different Distributions)

	Gamma		**Inverse Gaussian**	
	DF	Value	DF	Value
Null Deviance	3363	9712.5	3363	2.2792
Residual Deviance	3338	4607.6	3338	1.232
ΔD	25	5104.9	25	1.0472
AIC	1875474		1901578	

Similar to Table 5, from Table 6 we can see that the value of ΔD of the Gamma model, 5104.9, is bigger than the 95% quantile of χ^2_{25}, but the value of ΔD of the Inverse Gaussian model, 1.0472, is smaller than the 95% quantile of χ^2_{25}. So the Inverse Gaussian model can be regarded as a null model, and the Gamma model is significant. However, the value of residual deviance of the Gamma model is more close to its degree of freedom, and the AIC of Gamma is smaller, so that the Gamma distribution is better for our GLM of severity.

3.4.2 Comparison of Models with Static Premium (Including Kilometers or Not)

Using the Poisson and Gamma as the distribution of frequency and severity, we build the GLMs with static premium (not including kilometers and including kilometers) by R statistical software. The results of model is shown in Table 7 and Table 8.

Table 7 Models of Frequency and Severity (Not Including Kilometers)

	Frequency		**Severity**	
	Value	Pr(>|z|)	Value	Pr(>|t|)
(Intercept)	-1.577191	<2e-16***	8.213164	<2e-16***
G1	-0.135781	<2e-16***	0.386594	<2e-16***
Z2	-0.23569	<2e-16***	0.022099	0.058477.
Z3	-0.385496	<2e-16***	0.045835	0.000119***
Z4	-0.574709	<2e-16***	0.128802	<2e-16***
Z5	-0.336486	<2e-16***	0.045206	0.011433*
Z6	-0.520094	<2e-16***	0.143862	<2e-16***
Z7	-0.740246	<2e-16***	-0.004206	0.933218
B2	-0.445306	<2e-16***	0.045961	0.001976**
B3	-0.643529	<2e-16***	0.071548	1.60e-05***
B4	-0.762646	<2e-16***	0.061438	0.000586***
B5	-0.849754	<2e-16***	0.037808	0.026775*
B6	-0.913463	<2e-16***	0.076764	6.18e-08***
B7	-1.263193	<2e-16***	0.121847	<2e-16***
T2	0.119403	1.83e-08***	-0.035973	0.168136
T3	-0.190219	3.21e-14***	0.079354	0.010060*
T4	-0.786149	<2e-16***	-0.183844	4.95e-10***
T5	0.141873	2.35e-12***	-0.092693	0.000197***
T6	-0.371112	<2e-16***	-0.044535	0.036845*
T7	-0.068417	0.00338**	-0.124914	1.39e-05***
T8	0.044879	0.15456	0.20129	2.18e-07***
T9	-0.098709	<2e-16***	-0.056659	3.43e-06***

Table 8 Models of Frequency and Severity (Including Kilometers)

	Frequency		**Severity**	
	Value	Pr(>|z|)	Value	Pr(>|z|)
(Intercept)	-1.747551	<2e-16***	8.19652	<2e-16***
G1	-0.135457	<2e-16***	0.386645	<2e-16***
K2	0.212348	<2e-16***	0.025275	0.005965**
K3	0.319985	<2e-16***	0.021655	0.040961*
K4	0.404433	<2e-16***	0.041572	0.004897**
K5	0.575857	<2e-16***	0.037163	0.018023*
Z2	-0.238082	<2e-16***	0.022033	0.058071.
Z3	-0.386013	<2e-16***	0.046008	0.000104***
Z4	-0.581681	<2e-16***	0.12797	<2e-16***
Z5	-0.325688	<2e-16***	0.04619	0.009451**
Z6	-0.525908	<2e-16***	0.143934	<2e-16***
Z7	-0.73154	<2e-16***	-0.003889	0.937949
B2	-0.478867	<2e-16***	0.042545	0.004093**
B3	-0.693054	<2e-16***	0.066916	5.34e-05***
B4	-0.827282	<2e-16***	0.056083	0.001693**
B5	-0.925304	<2e-16***	0.031789	0.063193.
B6	-0.993413	<2e-16***	0.070816	7.02e-07***
B7	-1.327286	<2e-16***	0.11702	<2e-16***
T2	0.076684	0.000306***	-0.039179	0.132063
T3	-0.246851	<2e-16***	0.074808	0.014970*
T4	-0.6537	<2e-16***	-0.173037	4.93e-09***
T5	0.154972	1.88E-14***	-0.092096	0.000203***
T6	-0.335445	<2e-16***	-0.041404	0.051389.
T7	-0.055164	0.018117*	-0.123629	1.55e-05***
T8	-0.043581	0.167904	0.196679	3.84e-07***
T9	-0.067987	8.55E-12***	-0.054768	6.79e-06***

The models including kilometers are the nested models of the ones not including kilometers. So we can use the ANOVA test by R statistical software to compare nested models and original models. The results of test is shown in Table 9 and Table 10.

Table 9 ANOVA of Frequency

Original Model (1): Frequency ~ Gender + Zone + Bonus + Type					
Nested Model (2): Frequency ~ Gender + Kilometers + Zone + Bonus + Type					
	DF of Deviance	Deviance	DF	ΔD	Pr(>Chi)
1	4246	6154.3	4	2899.2	< 2.2e-16***
2	4242	3255.1			
Significance：0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 10 ANOVA of Severity

Original Model (1): Severity ~ Gender + Zone + Bonus + Type					
Nested Model (2): Severity ~ Gender + Kilometers + Zone + Bonus + Type					
	DF of Deviance	Deviance	DF	ΔD	Pr(>Chi)
1	3342	4627.5	4	19.855	0.01004*
2	3338	4607.6			
Significance：0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

In Table 9 and 10, DF means the degree of freedom; ΔD represents the difference between the deviances of nested model and original model. Because the values of ΔD are bigger than the 95% quantile of the corresponding Chi-Square distribution, we can reach a conclusion that nested models, the models with static premium (including kilometers), are better fitted.

So we insist that comparing with traditional static premium, the static premium with kilometers factor is better to explain the risk of insured.

3.4.3 Comparison of Models with Static and Dynamic Premium

Based on the models with static premium (including kilometers), we can calculate the exponents of the values simulated in models and get a rate table of static premium as Table 11.

When an insured purchases a policy of auto insurance, the insurer will collect the basic information of the insured and confirm the corresponding risk ranks of every factor. Then the insurer can calculate the premium based on the basic rate by multiplying the adjustment coefficients of the corresponding risk ranks. This premium is static, and fixed in a period of insurance (always one year) for the insured.

For calculating dynamic premium, we can first calculate the static premium per kilometer based on the static premium rate table. The method is to take the mid-value of each kilometers interval as the mean kilometers, as shown in Table 12.

For example, if an insured drives 10,000 kilometers per year, then the kilometers rank of this insured is K2. Based on Table 12, the mean kilometers of the insured is 8,000.

Then the premium per kilometer equals to the static premium divided by the mean kilometers. Dynamic premium can be calculated based on the premium per kilometer and the simulated actual kilometers.

Table 11 Static Premium Rate (Including Kilometers)

Rate Table (Including Kilometers)	**Frequency** (1)	**Severity** (2)	**Pure Premium** (1)×(2)
Basic Rate	0.17	3628.30	632.05
Adjustment Coefficient			
Gender0	1.00	1.00	1.00
Gender1	0.87	1.47	1.29
Kilometres1	1.00	1.00	1.00
Kilometres2	1.24	1.03	1.27
Kilometres3	1.38	1.02	1.41
Kilometres4	1.50	1.04	1.56
Kilometres5	1.78	1.04	1.85
Zone1	1.00	1.00	1.00
Zone2	0.79	1.02	0.81
Zone3	0.68	1.05	0.71
Zone4	0.56	1.14	0.64
Zone5	0.72	1.05	0.76
Zone6	0.59	1.15	0.68
Zone7	0.48	1.00	0.48
Bonus1	1.00	1.00	1.00
Bonus2	0.62	1.04	0.65
Bonus3	0.50	1.07	0.53
Bonus4	0.44	1.06	0.46
Bonus5	0.40	1.03	0.41
Bonus6	0.37	1.07	0.40
Bonus7	0.27	1.12	0.30
Type1	1.00	1.00	1.00
Type2	1.08	0.96	1.04
Type3	0.78	1.08	0.84
Type4	0.52	0.84	0.44
Type5	1.17	0.91	1.06
Type6	0.72	0.96	0.69
Type7	0.95	0.88	0.84
Type8	0.96	1.22	1.17
Type9	0.93	0.95	0.88

Table 12 Mean Kilometers

Kilometers Ranks	**Kilometers Interval**	**Mean Kilometers**
K1	< 1,000	500
K2	1,000 – 15,000	8,000
K3	15,000 – 20,000	17,500
K4	20,000 – 25,000	22,500
K5	> 25,000	30,000

To compare static and dynamic premium, the results is shown in Table 13.

Table 13 Comparison of Static and Dynamic Premium

	Static Premium	Dynamic Premium
Ratio of the Residuals Sum of Squared (Static Premium / Dynamic Premium)	1.08	
Stabilization of Premium	Only depend on the number of insured	Depend on the number of insured and kilometers
Positive Externality	No	Yes

In Table 13, we can see that the residual sum of squared of static premium is a little bigger than the one of dynamic premium. However, we calculate the dynamic premium by simulated kilometers, so this comparison may be not so significant. It is only said that dynamic premium may fit the data better than static premium.

In consideration of the stabilization of premium, the cash flows of static premium are more stable, because it is only dependent on the number of insured. Dynamic premium is also dependent on the kilometers which the insured drives actually, so its volatility is higher.

Considering attractiveness for consumers, dynamic premium has a positive externality, and it is attractive to the consumers who is more sensitive about the premium rate. So an insurance company using dynamic premium may improve its market share.

3.4.4 Combination of Static and Dynamic Premium

Dynamic premium is really the premium paid as you drive, however, there are some limitations from insurance companies and consumers to make it difficult to change static premium to dynamic directly. As a consequence, the combination of static and dynamic premium can be considered. A certain percentage of premium can be the static premium, which is fixed during the period of insurance, and the other percentage is the dynamic premium depending on the kilometers.

We can calculate the percentage of static premium, which make the residual sum of squared minimum. Based on this adjusted data and the simulated kilometers, the percentage of static premium should be 60%. It should be emphasized that the kilometers, which should be actual, are simulated in our data, so the result may have some deviation. In practice, the insurance companies can decide this percentage based on the actual kilometers, and take the financial stabilization and other purposes into consideration at the same time.

For sure, the combination of static and dynamic premium has some advantages, such as to keep the insurer's finance stable, to attractive consumers by dynamic premium, to improve fairness, and so on. This form of premium may be a good transitional one to help the insurers change their static premium to dynamic premium.

4. Conclusion

With the promotion of marketization reform in China, the prices of auto insurance products will be very important in the competition of insurance companies. But the vicious price competition may make the market confused, make the insurance companies lose their credits, and take threaten to the business stabilization of insurance

companies. It is sure that the Chinese insurance companies should improve their technology of pricing and classify the risks of different consumers.
This paper discuss the static and dynamic premium of auto insurance, and build models to compare the different forms of premium. By empirical analysis, we can get some conclusions as following.

1. **Adding the factor Kilometers can improve the fitness of the model with static premium.**

Kilometers is a very important factor to show the actual risk exposures of insured, and is paid extensive attention by insurance companies. In 2006, when the CIRC published the clauses A/B/C of auto insurance, CIRC said that the kilometers should be treated as a factor to adjust premium rate. But there were no applications in practice.
In our empirical analysis, with the comparison of the models with static premium (Including kilometers or not), it is a conclusion that the model including kilometers fit better, and all the coefficients are significant.

2. **Dynamic premium may improve the precision of model and has a positive externality.**

In theory, it is fairer to charge different premiums for insured with different driving kilometers, and the precision of model can be improved. In our empirical analysis, we compare the models with static and dynamic premium based on the residual sum of squares, and conclude that dynamic premium is a little better. However, the kilometers is simulated stochastically, and the data is grouped data, so there may be more detailed data needed to judge the precision of models.
In addition, dynamic premium is attractive to the consumers who is sensitive to premium rate, and it has a positive externality. But because the dynamic premium is related to the actual kilometers of insured, which can’t be known at the beginning of the insurance period, so the stabilization of dynamic premium is lower than static premium.

3. **Combination of static and dynamic premium may be a necessary transition.**

According to international experience, it is difficult to promote the pure dynamic premium directly. There are some advantages to keep a certain percentage of static premium, such as to make it easier to promote dynamic premium, to accumulate more information and data, to help the insurers improve their pricing models of dynamic premium in practice, and so on.
So we suggest that the Chinese insurance companies should not change the static premium to the pure dynamic premium directly, they should try the combination of static and dynamic premium as a transition.

References

[1] Guochen Pan. The Innovation of Usage-Based Insurance: Foreign Theory and Practice [J]. China Insurance, 2011, 05: 62-64.

[2] Huiqing Zhao, Hanzhang Wang. An Empirical Study of Rate Making of Automobile Insurance in China——Analysis Based on Generalized Linear Models [J]. Journal of Tianjin University of Commerce, 2011, 05:8-12.

[3] Yali Chen. Pay-as-you-drive pricing model: Research and Reference [D]. Dongbei University of Finance, 2013.

[4] Yin Zhang. The Study on the rating of automobile insurance with human factors [D]. Hunan University, 2007.

[5] Lianzeng Zhang, Dinghai Lv. Applications of Generalized Linear Models in Non-Life Insurance Ratemaking Analysis [J]. Application of Statistics and Management, 2013, 05:903-909.

[6] Xinjun Wang, Yajuan Wang. The Empirical Research on Classified Ratemaking of Automobile Insurance Based on Generalized Linear Models [J]. Insurance Studies, 2013, 09:43-56+85.

[7] Shengwang Meng. An Application of Generalized Linear Model to Auto motor Insurance Pricing [J]. Application of Statistics and Management, 2007, 01:24-29.

[8] Yan Gao. Foreign Automobile Insurance with Each Own Characteristics [J]. Chinese Credit Card, 2012, 03:77-79.

[9] Baige Duan, Dongfa Yu, Lianzeng Zhang. Foreign Automobile Insurance Mileage Pricing Theory and Practice [J]. Insurance Studies, 2012, 02:72-79.

[10] Lianzeng Zhang, Baige Duan. Study of the Impacts of Mileage on Vehicle Insurance Net Premiums——From the Perspective of the Impacts of Mileage on Accident Losses [J]. Insurance Studies, 2012, 06:29-38.

[11] Ruiyao Niu. Classification Ratemaking of Automobile based on Generalized Linear Models [D]. Jilin University, 2011.

[12] Xiaofeng Chen. The Enlightenment from European and American PAYD Auto Insurance to the Auto Insurance Pricing Reform in China [N]. China Insurance News, 2010-09-03002.

[13] Wei Zhang. Automobile Insurance: Mileage Pricing Will Become the Mainstream [N]. Financial Times, 2011-10-26010.

[14] Jian Yang. Ratemaking Study of Automobile Insurance in the Background of Marketing [D]. Dongbei University of Finance, 2012.

[15] Yajuan Wang. Research on Classification Ratemaking of Automobile Insurance Based on Generalized Linear Models [D]. Shandong University, 2013.

[16] Litman T. Pay-as-you-drive pricing and insurance regulatory objectives [J]. Journal of Insurance Regulation, 2005, 23(3): 35.

[17] Litman T A. Pay-As-You-Drive Pricing For Insurance Affordability [J]. May, 2004, 17: 1-17.

[18] Litman T. Pay-as-you-drive vehicle insurance in British Columbia [M]. Pacific Institute for Climate Solutions, University of Victoria, 2011.

[19] Boucher J P, Pérez-Marín A M, Santolino M. Pay-as-you-drive insurance: the effect of the kilometers on the risk of accident [C]//Anales del Instituto de Actuarios

Españoles. Instituto de Actuarios Españoles, 2013 (19): 135-154.
[20] Knoop V L, Li H, van Arem B. Variable insurance premium for safer driving: A survey result [C]//Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on. IEEE, 2011: 439-444.
[21] Guensler R, Amekudzi A, Williams J, et al. Current state regulatory support for Pay-as-You-Drive automobile insurance options [J]. Journal of Insurance Regulation, 2003, 21(3): 31-52.
[22] Litman T. Distance-based vehicle insurance as a TDM strategy [J]. Transportation Quarterly, 1997, 51: 119-137.
[23] Bolderdijk J W, Knockaert J, Steg E M, et al. Effects of Pay-As-You-Drive vehicle insurance on young drivers' speed choice: Results of a Dutch field experiment [J]. Accident Analysis & Prevention, 2011, 43(3): 1181-1186.
[24] Paefgen J, Staake T, Thiesse F. Evaluation and aggregation of pay-as-you-drive insurance rate factors: A classification analysis approach [J]. Decision Support Systems, 2013, 56: 192-201.
[25] Azzopardi M, Cortis D. Implementing Automotive Telematics for Insurance Covers of Fleets [J]. Journal of technology management & innovation, 2013, 8(4): 59-67.
[26] Litman T A. Implementing Pay-As-You-Drive Vehicle Insurance [J]. Policy Options, The Institute for Public Policy Research, London, 2002.
[27] Insurance fraud: The crime you pay for [M]. Coalition Against Insurance Fraud, 2005.
[28] Parry I W H. Is Pay-as-You-Drive insurance a better way to reduce gasoline than gasoline taxes? [J]. American Economic Review, 2005: 288-293.
[29] Abou-Zeid M, Ben-Akiva M, Tierney K, et al. Minnesota pay-as-you-drive pricing experiment [J]. Transportation Research Record: Journal of the Transportation Research Board, 2008, 2079(1): 8-14.
[30] Bordoff J. Pay-as-you-drive car insurance [J]. Democracy Journal, 2008, 8.
[31] Agerholm N, Waagepetersen R, Tradisauskas N, et al. Preliminary results from the Danish intelligent speed adaptation project pay as you speed [J]. Intelligent Transport Systems, IET, 2008, 2(2): 143-153.
[32] Troncoso C, Danezis G, Kosta E, et al. PriPAYD: Privacy-friendly pay-as-you-drive insurance [J]. Dependable and Secure Computing, IEEE Transactions on, 2011, 8(5): 742-755.