The Journal of Risk and Insurance

DOI: 10.1111/j.1539-6975.2011.01433.x

Can Vehicle Maintenance Records Predict Automobile Accidents?

Shyi-Tarn Bair
Rachel J. Huang
Kili C. Wang

Abstract

This article proposes that vehicle maintenance records can provide useful information for predicting the probability that an owner will have an automobile accident. To test the hypothesis, we use a unique data set that is merged from an insurance company and a vehicle manufacturer in Taiwan. We find weak evidence to support our hypothesis. Among all the proxies for proper maintenance, we indicate that proper maintenance defined by the recommended kilometers is significantly negatively correlated with the loss probability in compulsory automobile liability insurance. The average loss probability decreases by 0.23 percent when the insured vehicle is properly maintained according to the recommended number of kilometers in the previous years, whereas the average loss probability for the overall sample is 0.49 percent. We further find that proper maintenance is insignificantly correlated with loss severity.

Introduction

Understanding the factors that affect automobile accidents is an essential and important issue due to the fact that automobile insurance is the most important insurance line in many countries. For example, in the United States, personal automobile insurance is the largest line of insurance in the property and casualty insurance industry. The total net written premiums for personal automobile insurance amounted to $159.7 billion in 2007, accounting for 35.7 percent of property and casualty insurance. The average growth rate of the net written premiums for personal automobile insurance was about 4.2 percent per year for the period from 2000 to 2007.[1]

Shyi-Tarn Bair is CEO, Ho-An Insurance Agency Co., Ltd., Taiwan. Rachel J. Huang is Associate Professor, Graduate Institute of Finance, National Taiwan University of Science and Technology, Taipei, Taiwan. Kili C. Wang is Associate Professor, Department of Insurance, Tamkang University, Taipei, Taiwan. Both Rachel J. Huang and Kili C. Wang are research fellows in Risk and Insurance Research Center, College of Commerce, NCCU. Rachel J. Huang can be contacted via e-mail: rachel@mail.ntust.edu.tw. The authors would like to express their deep gratitude to Larry Tzeng, Yi-Ting Chen, and the two referees who made stimulating comments on the paper.

[1] The data are obtained from the U.S. Census Bureau, Statistical Abstract of the United States: 2010, Table 1185.

A large body of literature has focused on the determinant factors predicting automobile accidents. In addition to the main classificatory variables used in the automobile insurance market, such as gender, age, marital status, zip code, and the claim history, researchers have further found that credit history (Tillman and Hobbs, 1949; Kellison and Brockett, 2003; Miller and Smith, 2003), the market price of an automobile (Andersson, 2005), and drinking and sensation-seeking behavior (such as driving fast and racing cars) (Iversen and Rundmo, 2002; Brockett and Golden, 2007) are correlated with automobile accidents.

Recently, Brockett and Golden (2007) provide a detailed survey of the related literature and classified the risk factors based on biological, psychological, and behavioral characteristics. More importantly, they hypothesized that the reason why credit scoring works in underwriting automobile insurance[2] is that there is an underlying biological and/or psychological and social component of each individual, which regulates their risk-taking behavior, and since this propensity is individual and intrinsic, it spans risk scenarios, from financial decision making to risky-driving decision making.

Inspired by Brockett and Golden (2007), this article intends to study whether maintenance records can serve as another variable that can present the biological and/or psychological and social component of each individual, and could therefore be correlated with automobile accidents. To be specific, in this article, we empirically investigate whether maintenance records can predict the automobile accident rates and automobile accident claim amounts.

Why can maintenance records influence automobile accidents? First, following Brockett and Golden's (2007) hypothesis, people who are more responsible in their vehicle maintenance might also be more responsible in their driving behavior. Second, the vehicle maintenance records could be a signal of the degree of risk aversion of the driver. Jindapon and Neilson (2007) show that a more risk-averse individual is willing to expend more efforts on self-protection for a mean-preserving improvement in the risk. If the car is properly maintained, then the driver may be a more risk-averse individual. Therefore, on the one hand, she may have less of a chance of being involved in an automobile accident. On the other hand, she may not cause a severe automobile accident even if she is involved in one. Third, Angers et al. (2006) demonstrate that failure to pass mechanical inspection is significant in explaining accidents involving trucks. In other words, we would predict that if the car is properly maintained, then, for mechanical reasons, the frequency and severity of automobile accidents may be lower.

Based on the above arguments, we set up two hypotheses. The first one is that proper vehicle maintenance can reduce the probability of an automobile accident. The second hypothesis is that proper vehicle maintenance can decrease the loss severity. To test these two hypotheses, we use a unique data set that is merged from an insurance company and a vehicle manufacturer. The data from the insurance company allow us to observe claim records, insurance contracts, and all variables used by the insurance company in underwriting and pricing. The data from the vehicle manufacturer contain the maintenance dates, kilometers traveled, and items involved in the vehicle maintenance. Only the observations that are recorded by both

[2] For example, please see Kellison and Brockett (2003) for the statistical analysis between credit scores and insured losses.

the manufacturer's repair shop and the insurance company are included in our final sample. Through the help of both companies, we link each individual from these two data sets and construct a pooling data set. In total, we have 65,095 observations. The sample period is from 2003 to 2006.

The insurance we focus on is compulsory automobile liability insurance, which is designed to protect the third party's life and body from the damage caused by the usage of motor vehicles. There are two reasons why we explore the determining factors in predicting automobile accidents within this market. First, we will have the largest sample size possible because each vehicle must by law participate in this insurance. Second, the claim records will serve as a good proxy for accidents in this market. According to Chiappori and Salanie (2000), using claim records as a proxy for accidents may suffer from a bias of unclaimed accidents. However, this problem is less serious because there is no deductible and there must have been a third party involved.

We find that the proper maintenance defined by the recommended kilometers in the previous year is significantly negatively correlated with the occurrence of compulsory automobile liability claims. The average loss probability decreases by 0.23 percent when the insured vehicle is properly maintained according to the recommended number of kilometers in the previous years. The probability decrease accounts for 46.9 percent of the average claim rate. The empirical evidence supports the view that proper maintenance defined by the recommended kilometers in the previous year could be an appropriate proxy for a more responsible behavior, and/or the degree of risk aversion. It might also indicate that the insured vehicle has fewer mechanical problems.

As for the relationship between proper maintenance and loss severity, we find that the maintenance records are not significant factors in terms of predicting the claim amounts for compulsory automobile liability. Although proper maintenance will decrease the loss severity through the use of different proxies for proper maintenance, the relationship is statistically insignificant.

The remainder of this article is organized as follows. In the next section, we provide a brief introduction to the automobile insurance market in Taiwan. The "Data" section describes our data. The "Methodology" section introduces our main empirical methodologies. The "Empirical Results" section reports our empirical findings. The "Sensitivity Analyses" section provides sensitivity analysis by varying the threshold while defining proper maintenance. In the "The Characteristics of the Individual Whose Vehicle Is Properly Maintained" section, we further examine the relationship between the maintenance records and the insured's characteristics. The last section concludes the article.

THE AUTOMOBILE LIABILITY INSURANCE MARKET IN TAIWAN

As in many other countries in the world, there are two systems of automobile liability insurance in Taiwan: compulsory and voluntary. Each motor vehicle must be insured under the system of compulsory liability, which is designed to protect the third party's life as well as protect the third party against bodily injury caused by the usage of the vehicle. Individuals can also purchase voluntary third-party liability insurance, which

covers bodily injury or property damage sustained by the third party, to compensate in the event of insufficient coverage under compulsory insurance.

There are three types of payment in regard to compulsory insurance: life, maiming, and medical expenses. The compulsory insurance coverage for life is NT$1,600,000.[3] The degree of incapacity is divided into 15 levels. If the injured third party becomes disabled as the result of an accident, he or she will be reimbursed by between NT$40,000 and NT$1,600,000 according to the level of incapacity. The medical expenses include the costs of first aid and treatment. The upper-limit indemnity for the medical expenses of the injured third party is NT$200,000.

In 2009, a total of 13,092,421 vehicles were covered by compulsory insurance. The average growth rate of the number of vehicles covered is 2.48 percent per year from 2003 to 2009. Written premiums in 2009 amounted to NT$16,229 million and exhibited an average growth rate of 0.17 percent since 2003. The average loss ratio was 68.37 percent between 2003 and 2009.[4]

DATA

Our empirical data are merged from two sources: one is from a vehicle manufacturer and the other is from an insurance company. The vehicle manufacturer's market share was about 38 percent in Taiwan in 2009. This part of the data provides us with the maintenance information for each vehicle. The written premiums of our sample company accounted for 20 percent of the automobile insurance market in 2009. The insurance data contain the claim records, insurance policies, and information related to the owner and the insured vehicle. The individual characteristics could help us to control the heterogeneity of the observations. The individuals we study are those who purchase compulsory automobile liability insurance in our sample insurance company, and who also have their vehicles maintained or repaired by our sample manufacturer from 2001 to 2006.

Since the purpose of our article is to examine whether maintenance records could predict insurance claims, we trace the maintenance information prior to the insurance policy year. There are two advantages in doing so. The first one is that the previous maintenance behavior will not be affected by the current claim and the current insurance contract. In other words, the moral hazard problem will be less severe and the interactions between maintenance behavior and the claim status might be simplified. Second, the previous maintenance data would be observable to the insurance companies. They could use such data as one of the risk factors to price the insurance in practice.

[3] The average exchange rate during our sample period was approximately NT$33.14 to US$1, and the standard deviation was NT$1.04 per US dollar.

[4] In Taiwan, the average direct written premiums for automobile insurance between 2003 and 2006 amounted to NT$55 billion in each year, which accounted for around 48 percent of the property and casualty insurance market on average. The average annual growth rate of the direct written premiums of the automobile insurance was about 4.53 percent over this period. The above data were obtained from the website of the Taiwan Insurance Institute (http://www.tii.org.tw/).

Furthermore, our sample vehicle manufacturer suggests that a vehicle should be maintained every 10,000 km or every 6 months. Since the average kilometers driven per year per vehicle in our sample is about 17,000 km, it means that regardless of whether the vehicle is maintained according to the suggested kilometers driven or time, we will only have about one or two maintenance records for each vehicle per year. In order to minimize the bias from using one-time records to define proper maintenance, we check 2 years rather than 1 year of maintenance information to classify proper maintenance.

Thus, we delete the observations that do not have continuous 2-year maintenance records prior to the insurance policy year. After the deletion, the total number of observations in our final sample is 65,095 and the sample period extends from 2003 to 2006.[5]

The definitions of our variables are listed in Table 1. The two dependent variables used in the two different empirical models include a dummy variable as to whether the insured files a claim (noted by $claim$) and the claim amount ($amount$).[6] The explanatory variables in this research include the characteristics of the insured individuals and insured cars that are listed in the part for the independent variables in Table 1.

The key variable that we focus on is the one that indicates whether the vehicle is properly maintained. We adopt different proxies for proper maintenance. In practice, the vehicle owners can visit the repair shops of the vehicle manufacturer to maintain their cars according to the recommended number of kilometers or dates.[7] As mentioned above, the sample vehicle manufacturer suggests that vehicles should be taken to their service and repair centers every 10,000 km or every 6 months. If the recommended number of kilometers or date is missing, the repair shop can still follow the maintenance instructions to replace or check the suggested items when the vehicle is sent back to the shop. Thus, we use three criteria to identify whether the vehicle has been properly maintained.

The first criterion, vm_km, evaluates proper maintenance according to whether the maintenance is carried out according to the recommended number of kilometers. Note that we allow an extra 20 percent in regard to the recommended number of kilometers.[8] Therefore, if the vehicle is sent in for maintenance within every 12,000 km in the 2 years prior to the policy year, then $vm\_km = 1$, otherwise $vm\_km = 0$. The second criterion, vm_time, evaluates proper maintenance according to whether the maintenance is conducted according to the recommended times. Again, an additional 20 percent of the recommended time is allowed. Thus, if the vehicle is sent for maintenance within 7.2 months of the previous maintenance record in the 2 years prior to the

[5] We also construct a sample that includes these deleted observations and treat them as the observations without proper maintenance. Our findings are still robust.

[6] The claim record in relation to the compulsory automobile liability insurance is the proxy for the occurrence of the risk.

[7] The vehicle maintenance records cover only those repairs and maintenance that are carried out at the brand dealership, so they may understate the results and make the observed significance less than it actually is (i.e., the statistical significance is found to be conservative). This choice of information source for the vehicle maintenance records, while inevitable practically, may mean, for example, that older used cars and younger drivers are underrepresented.

[8] We will further relax the extra 20 percent threshold in the section on robustness analyses.

TABLE 1
Definitions of the Variables

Variables	Definition
Dependent variables	
claim	A dummy variable that equals 1 when the insured has filed an at claim on the compulsory automobile third-party liability insurance, otherwise it equals 0.
amount	Claim amount of the insured for the whole policy year in thousands of NT dollars.
Independent variables	
Vehicle maintenance records	
vm_km	A dummy variable that equals 1 when the vehicle maintenance is done according to the recommended number of kilometers each time it is maintained during both of the 2 previous years, otherwise it equals 0.
vm_item	A dummy variable that equals 1 when the vehicle maintenance is done according to the recommended items each time it is maintained during both of the 2 previous years, otherwise it equals 0.
vm_time	A dummy variable that equals 1 when the vehicle maintenance is done according to the recommended time each time it is maintained during both of the 2 previous years, otherwise it equals 0.
Control variables	
age2530	A dummy variable that equals 1 when the insured is between the ages of 25 and 30, otherwise it equals 0.
age3060	A dummy variable that equals 1 when the insured is between the ages of 30 and 60, otherwise it equals 0.
age60up	A dummy variable that equals 1 when the insured is over the age of 60, otherwise it equals 0.[a]
female	A dummy variable that equals 1 when the owner of the car is female, otherwise it equals 0.
married	A dummy variable that equals 1 when the owner of car is married, otherwise it equals 0.
carage2	A dummy variable that equals 1 when the car is 2 years old, otherwise it equals 0.
carage3	A dummy variable that equals 1 when the car is 3 years old, otherwise it equals 0.
carage4	A dummy variable that equals 1 when the car is 4 years old, otherwise it equals 0.[b]
big	A dummy variable that equals 1 when the insured car equals or is over 2000 c.c., otherwise it equals 0.

(Continued)

TABLE 1
Continued

Variables	Definition
sedan	A dummy variable that equals 1 when the car is a sedan and is for noncommercial or for long-term rental purposes, otherwise it equals 0.[c]
channel_D	A dummy variable that equals 1 when the policy is sold through the channel of a car dealer-owned agency, otherwise it equals 0.
city	A dummy variable that equals 1 when the owner of the car lives in a city, otherwise it equals 0.
north	A dummy variable that equals 1 when the car is registered in the north of Taiwan, otherwise it equals 0.
south	A dummy variable that equals 1 when the car is registered in the south of Taiwan, otherwise it equals 0.
east	A dummy variable that equals 1 when the car is registered in the east of Taiwan, otherwise it equals 0.[d]
year2004	A dummy variable that equals 1 when the data belong to the year 2004, otherwise it equals 0.
year2005	A dummy variable that equals 1 when the data belong to the year 2005, otherwise it equals 0.
year2006	A dummy variable that equals 1 when the data belong to the year 2006, otherwise it equals 0.[e]

[a] The reference group for the dummy variables related to age includes the insured who are under 25 years old.
[b] The reference group for the dummy variables related to the car age is that which includes all the cars used over 4 years. Note that we require our observations to carry 2-year maintenance records. Thus, our sample does not include vehicles that are less than 2 years old.
[c] The reference group for the dummy variable includes the insured cars that are for noncommercial use, other than sedans for long-term rental purposes.
[d] The reference group for those three dummy variables related to area includes the cars registered in central Taiwan.
[e] The reference group for those three dummy variables related to year includes the data for the year 2003 in our regressions.

TABLE 2
Basic Statistics of Full Sample

	All Observations		Claimed Observations	
Variables	Mean	Std. Dev.	Mean	Std. Dev.
claim	0.0049	0.0698		
amount			100.602	227.852
vm_km	0.0645	0.2275	0.0408	0.1980
vm_item	0.0857	0.2590	0.0627	0.2428
vm_time	0.3116	0.3340	0.3103	0.4634
age2530	0.0603	0.2425	0.0502	0.2186
age3060	0.8955	0.3144	0.8966	0.3050
age60up	0.0380	0.1977	0.0439	0.2052
female	0.6926	0.4615	0.7273	0.4461
married	0.9292	0.2664	0.9436	0.2311
carage2	0.3119	0.4432	0.3009	0.4594
carage3	0.1904	0.3715	0.1693	0.3756
carage4	0.0990	0.3019	0.0909	0.2879
big	0.2467	0.4391	0.2132	0.4102
sedan	0.9824	0.1309	0.9812	0.1361
channel_D	0.4899	0.4886	0.4545	0.4987
city	0.4975	0.4996	0.4389	0.4970
north	0.4941	0.4997	0.3793	0.4860
south	0.2794	0.4451	0.3574	0.4800
east	0.0220	0.1421	0.0345	0.1828
year2004	0.2010	0.3793	0.2351	0.4247
year2005	0.2679	0.4272	0.2445	0.4305
year2006	0.3890	0.4634	0.3762	0.4852
Number of observations	65,095		319	

policy year, then $vm\_time = 1$, otherwise $vm\_time = 0$. The third criterion, vm_item, evaluates proper maintenance according to whether the maintenance is conducted according to the recommended items. If the maintenance record in the 2 years prior to the policy year shows that the vehicle is checked according to the maintenance instructions or more than required by those instructions, then $vm\_item = 1$, otherwise $vm\_item = 0$.

The control variables include the underwriting variables of the compulsory automobile liability insurance, such as the individual's age and gender. We also include other variables to control for the heterogeneity of the insured, such as the individual's marital status, the vehicle's age, size and model, the registration area, the sales channel of the insurance policy, and the year dummies.

The basic statistics for the final sample are listed in the first two columns of Table 2. The probability of having filed a compulsory liability insurance claim is around 0.49 percent. Our sample claim probability is lower than the average claim probability for the compulsory insurance market as a whole, which was 1.17 percent from 2003 to 2006. In our sample, most of the insured who have previously filed claims only filed

one claim per accident during our sample period. Only two individuals filed two claims, and none of them filed more than two claims during the sample period.

Based on the index for *vm_km*, 6.45 percent of insured cars are properly maintained; based on the index for *vm_item*, 8.57 percent of insured cars are properly maintained; and based on the index for *vm_time*, 31.16 percent of insured cars are properly maintained.[9] Note that *vm_time* is higher than *vm_km* and *vm_item*. This may arise because *vm_time* indicates that some of the car owners maintain their cars because they are advised to do so. Our sample vehicle manufacturer will record the date of maintenance for its customers, and call the customers to remind them of the scheduled maintenance when the recommended date is approaching. Although our sample vehicle manufacturer also records the kilometers for its customers, it has difficulty knowing whether the recommended number of kilometers is approaching. In other words, *vm_time* includes both the active and passive maintenance records, whereas *vm_km* represents the active maintenance behavior.

Most of the insured are married and between 30 and 60 years old, the percentage of insured in this age range is over 89 percent, and the percentage of married insured is 92.92 percent. Female insured account for over half of the total, and the percentage is 69.26 percent. One of the reasons why most vehicles are owned by married middle-aged females is that the insurance premium is discounted for them. In Taiwan, the insurance premium is dependent on the characteristics of the owner of the vehicle rather than the drivers. Since married middle-aged females can enjoy a premium discount, most families register their vehicle under the name of a married middle-aged female member of the family.

Since each observation is required to carry at least 2 years of maintenance records prior to the insurance policy year, our sample excludes vehicles less than 2 years old. The percentage of cars between 2 and 4 years old is around 60 percent. The bigger cars (over 2000 cc.) account for about 24.67 percent of the total. Most (about 98 percent) of them are noncommercial sedans. Around 48 percent of the insurance contracts in our empirical sample were sold by the car dealers' own agents. About 50 percent of insured cars are used in cities, and 49 percent are registered in the northern area of Taiwan.

In our empirical work, we further investigate whether the cars' maintenance records are able to explain the severity of third-party liability risk. Therefore, we constrain our data into a subsample of those insured who have filed a claim. The basic statistics for the subsample of the claimed insured are listed in the last two columns of Table 2. The average claim amount is NT$100,602.[10] Our average claim amount is higher than that for the whole market, which is NT$78,730 during our sample period. The structures of the characteristics of the individual and the vehicle, e.g., sex, age, the

[9] One might wonder whether cars involved in accidents will be maintained on a timely basis simply because of their accident involvement. Our data show that among the claims sample, only 1.183 percent, 1.164 percent, and 1.085 percent of the observations face this kind of problem when the proper maintenance is defined according to kilometers, items, and time, respectively.

[10] The maximum value of the claim amount is NT$1,712,240. The minimum value of the claim amount is NT$2,000.

marital status of the insured, and the age and usage of the insured cars, are similar to those in the full sample. The maintenance indices according to kilometer and item, *vm_km* and *vm_item*, are about 2 percent lower in this subsample than in the full sample, whereas *vm_time* is similar to the full sample. This may imply that proper maintenance evaluated by kilometer and item could predict the risk of third-party liability, and the ones who properly maintain their vehicles may have a lower risk probability.

It is worth noting that we only include one particular brand instead of different brands of cars in our final sample. This is because motor vehicle manufacturers only take care of their own brand of vehicles. Since our sample only contains one brand of car, it has different features than the observations based on purchases of compulsory liability from our insurance company, which contains several brands of cars. To be specific, our sample is insignificantly different from the data obtained from the sample insurance company with respect to the vehicle registration location. However, our data set has a significantly higher percentage of females and new cars (2 or 3 years old). The individuals in our data set also have a higher probability of purchasing insurance policies through the channel of a car dealer-owned agency.

Methodology

There are two major hypotheses in this article that may be described as follows:

Hypothesis 1: *The accident probability will decrease when the insured vehicle is properly maintained.*

Hypothesis 2: *The accident severity will decrease when the insured vehicle is properly maintained.*

To test the first hypothesis, we employ pooled probit models[11] as follows:

$$\begin{aligned} \Pr(claim_{it} = 1 | vm\_km_{it}, X_{it}, year) \\ = F(vm\_km_{it}\beta + X_{it}\gamma_x + year_t\gamma_y), \end{aligned} \tag{1}$$

$$\begin{aligned} \Pr(claim_{it} = 1 | vm\_item_{it}, X_{it}, year) \\ = F(vm\_item_{it}\beta + X_{it}\gamma_x + year_t\gamma_y), \end{aligned} \tag{2}$$

$$\begin{aligned} \Pr(claim_{it} = 1 | vm\_time_{it}, X_{it}, year) \\ = F(vm\_time_{it}\beta + X_{it}\gamma_x + year_t\gamma_y), \end{aligned} \tag{3}$$

[11] Note that most of our variables do not vary over time. The major time-varying variable is the vehicle maintenance variable. To better control the individual heterogeneity and emphasize the individual characteristics, we adopt pooled probit regressions and pooled OLS regressions to test the relationship between proper maintenance and loss probability and between proper maintenance and loss severity, respectively.

where $claim_{it} = 1$ when the insured has filed the claim based on compulsory automobile liability insurance during the policy year, otherwise $claim_{it} = 0$. F is the cumulative distribution function of the probit regression and is normally distributed. The vectors X_{it} denote the explanatory variables that are listed in Table 1. $year_t$ is the vector of year dummy[12] which controls the time effect. β, γ_x, and γ_y are the coefficients for the proper maintenance variable, X_{it}, and $year_t$.

The prediction ability of the maintenance record in regard to the loss probability is captured by the significance of the estimated coefficient $\hat{\beta}$. The predicted sign of $\hat{\beta}$ is negative since our hypothesis is that the risk of third-party liability is negatively correlated with the maintenance records. We further calculate the change in the estimated claim probability due to proper maintenance.

Note that maintenance behavior might be endogenously determined. However, the endogeneity will not affect the estimation in our regressions since whether or not a vehicle is properly maintained is based on the maintenance records in the previous years rather than in the same year of the insurance policy. In other words, the vehicle maintenance variables in our regressions are observable. Thus, even though the maintenance behavior is endogenous, the endogeneity does not need to be taken care of in our regressions.

To test our second hypothesis, we perform the regression on the claim amount of the compulsory automobile liability insurance under the subsample of claimed policies. In controlling the time effect, we set up the ordinary least squares models as follows:

$$amount_{it} = \alpha + \theta vm\_km_{it} + X_{it}\delta_x + year_t\delta_y + \varepsilon_{it}, \tag{4}$$

$$amount_{it} = \alpha + \theta vm\_item_{it} + X_{it}\delta_x + year_t\delta_y + \varepsilon_{it}, \tag{5}$$

$$amount_{it} = \alpha + \theta vm\_time_{it} + X_{it}\delta_x + year_t\delta_y + \varepsilon_{it}, \tag{6}$$

where $amount_{it}$ denotes the claim amount in thousands NT dollars, and ε_{it} is the random residual. The key coefficient that we focus on is the significance of the estimated coefficient $\hat{\theta}$. We also predict that the sign of $\hat{\theta}$ should be negative if the hypothesis of negative correlation between risk severity and maintenance is sustained.

Empirical Results

As shown in Table 3, we find that proper maintenance defined by the recommended kilometers is significantly negatively correlated with risk probability. Proper maintenance defined by the recommended items is also negatively correlated with risk probability, but it is only significant at the 10 percent level. Our results suggest that *vm_km* could be a proper proxy for a more responsible driving behavior, a higher degree of risk aversion, or for an insured vehicle with fewer mechanical problems.

[12] In this article, we use the dummy variables *year2004*, *year2005*, and *year2006* to control the time effect.

TABLE 3
Pooled Probit Regression of Compulsory Liability Claim

	Model 1		Model 2		Model 3	
Variables	Coefficient	*p*-value	Coefficient	*p*-value	Coefficient	*p*-value
Intercept	−2.3584	< 0.0001	−2.3576	< 0.0001	−2.3793	< 0.0001
vm_km	−0.2080	0.0026				
vm_item			−0.1480	0.0614		
vm_time					0.0151	0.7102
age2530	−0.2231	0.3230	−0.2241	0.3209	−0.2260	0.3164
age3060	−0.1657	0.4360	−0.1672	0.4318	−0.1681	0.4289
age60up	−0.0995	0.6682	−0.1024	0.6590	−0.1049	0.6507
female	0.0393	0.3725	0.0390	0.3752	0.0393	0.3715
married	0.0890	0.2972	0.0863	0.3115	0.0870	0.3073
carage2	−0.0524	0.2582	−0.0489	0.2907	−0.0469	0.3137
carage3	−0.0900	0.1103	−0.0855	0.1287	−0.0839	0.1383
carage4	−0.0663	0.3628	−0.0624	0.3913	−0.0606	0.4064
big	−0.0503	0.2886	−0.0503	0.2880	−0.0502	0.2889
sedan	−0.0086	0.9522	−0.0097	0.9463	−0.0109	0.9398
channel_D	−0.0910	0.0231	−0.0896	0.0259	−0.0744	0.0599
city	−0.0519	0.1986	−0.0515	0.2014	−0.0500	0.2150
north	−0.1286	0.0125	−0.1284	0.0126	−0.1291	0.0122
south	0.0423	0.4277	0.0417	0.4342	0.0420	0.4304
east	0.0867	0.4652	0.0886	0.4558	0.0809	0.4964
year2004	0.0642	0.3363	0.0652	0.3285	0.0639	0.3380
year2005	−0.0172	0.7956	−0.0156	0.8145	−0.0161	0.8083
year2006	−0.0032	0.9594	−0.0009	0.9884	0.0000	0.9995
Likelihood ratio	38.4091		36.7822		33.0406	

To examine whether the effects of *vm_km* are economically significant, we further calculate the average reduction in loss probability from maintenance from Model 1 in Table 3. We find that the average loss probability decreases by 0.23 percent when the insured vehicle is properly maintained according to the recommended number of kilometers in the 2 years prior to the policy year.[13] In addition, we respectively examine the decrease in probability when the insured vehicle is properly maintained according to the recommended number of kilometers 1 year and 2 years before the policy year. We find that in both time horizons, *vm_km* is also economically and statistically significant. The average loss probability decreases by 0.17 percent and 0.19 percent when the insured vehicle is properly maintained according to the recommended number of kilometers 1 year and 2 years before the policy year, respectively. These differences are significant at the 1 percent level by using the normal test.

Model 2 in Table 3 indicates that *vm_item* is also statistically significant. When the insured vehicle is properly maintained according to the recommend items each time when the vehicle visits the service center in the 2 years prior to the policy year, the accident probability on average is 0.33 percent. It is 0.51 percent otherwise. The difference is 0.18 percent, which is significant at the 1 percent level by using the Normal test.

[13] This difference is significant at the 1 percent level by using the Normal test.

TABLE 4
Pooled OLS Regression on Compulsory Liability Claim Amount in the Claimed Subsample

	Model 1		Model 2		Model 3	
Variables	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Intercept	94.2655	< 0.0001	94.4288	< 0.0001	84.2907	< 0.0001
vm_km	−73.3396	0.2761				
vm_item			−20.4150	0.7122		
vm_time					−22.3110	0.4390
age2530	−33.9583	0.8145	−29.8322	0.8371	−29.0456	0.8410
age3060	−12.9923	0.9231	−12.5960	0.9256	−10.3951	0.9385
age60up	61.2645	0.6775	59.7890	0.6854	65.0956	0.6592
female	0.6685	0.9821	1.3747	0.9633	4.1100	0.8914
married	75.4694	0.0983	79.7121	0.0754	76.0571	0.0954
carage2	−64.3458	0.0385	−65.5552	0.0352	−63.6301	0.0415
carage3	−38.5142	0.3011	−40.1791	0.2812	−41.6586	0.2633
carage4	−13.4430	0.7812	−13.6028	0.7790	−11.3216	0.8156
big	−32.1660	0.3207	−31.7041	0.3285	−32.1876	0.3209
sedan	−4.4093	0.9638	−1.9530	0.9840	3.1694	0.9740
channel_D	21.7962	0.4186	18.0584	0.5019	16.2190	0.5408
city	3.4012	0.8999	2.7571	0.9191	0.2581	0.9924
north	−47.1293	0.0763	−48.1638	0.0676	−51.2313	0.0441
south	24.0889	0.4995	21.7362	0.5423	19.0582	0.5933
east	7.1595	0.9268	11.0847	0.8875	15.1237	0.8456
year2004	13.1524	0.7671	10.8499	0.8072	13.9652	0.7545
year2005	−28.5044	0.5282	−31.1626	0.4931	−29.5581	0.5140
year2006	−47.5816	0.2610	−50.3628	0.2363	−47.6039	0.2637
Adj. R^2	0.092		0.057		0.027	

It is worth noting that *vm_time* is insignificantly positive in Model 3. As mentioned in the previous section, this variable includes both active and passive maintenance behavior. The driving force behind a vehicle being maintained according to the suggested time could be either the individual's characteristics or the reminder. This might be the reason why we find an insignificant relationship between *vm_time* and the loss probability.

Table 3 also demonstrates that the individuals who purchase the insurance policy through the channel of a car dealer-owned agency will have a lower claim probability regardless of the proxies of proper maintenance. This result suggests that car dealer-owned agencies might have a certain ability to screen the insured during underwriting. In addition, we find that the loss probability is lower in northern Taiwan than in other parts of Taiwan in all models. This might be because the population density in northern Taiwan is higher than in other places and traffic moves more slowly in the high population density areas.

For our second hypothesis, we cannot find statistical support. As shown in Table 4, we find that if the vehicle is properly maintained according to the recommended

number of kilometers, items, or time, then the loss severity will be lower. However, the estimated coefficients of the proper maintenance variable in all models are insignificant. These findings imply that the information regarding whether the car owners properly maintain their car or not cannot provide valuable information for predicting the severity of risk for those who have already filed a claim.

In addition to the proper maintenance variable, there are several variables that have significant explanatory power in terms of loss severity but not loss probability. We find that married vehicle owners have a significantly higher loss severity than single owners. The difference in the claim amounts between married owners and single owners is about NT$77,000 per claim. This finding is not, however, consistent with standard insurance heuristics. One possible explanation[14] could be that, due to the particularities of the Taiwan insurance market, the women are the registered owners when there are married people, even when the men are mostly doing the driving (due to cultural effects), while the single drivers are the ones who actually drive, with males and females being spread roughly evenly. Thus, if men have a higher accident severity, they will then contribute to most of the exposure of married driving, but only to half the exposure of the single drivers. We also find that older cars (over 4 years old) have a higher loss severity than younger cars. The average claim amount for vehicles that are 2 years old is NT$64,500 per year less than that of vehicles that are more than 4 years old. In addition, we find that vehicles registered in north Taiwan have a lower loss severity than those registered in central Taiwan.

Sensitivity Analyses

In this section, we vary the chosen threshold (20 percent extra time/miles) while defining proper maintenance to examine the sensitivity of the relationship between proper maintenance and loss probability.[15] We allow an extra 15 percent to 25 percent in regard to the recommended number of kilometers and recommended time.[16]

Table 5 reports the coefficients and the *p*-values of the proper maintenance variables of Models 1 and 3 in the pooled probit regressions of the full sample. It shows that no matter which threshold is chosen, proper maintenance according to the recommended number of kilometers could significantly reduce the probability of having a compulsory liability claim.

From the coefficients, we find that the largest probability deduction occurs when a 20 percent threshold is chosen. The negative correlation between vm_km and the loss probability increases in the threshold chosen when the threshold is below 20 percent, whereas it decreases in the threshold chosen when the threshold is above 20 percent. As in Brockett and Golden's (2007) hypothesis, we argue that maintenance records

[14] We thank an anonymous referee for providing a possible explanation.

[15] The sensitivity of the relationship between proper maintenance and loss severity is also examined. The relationship is consistently insignificant when using different thresholds. Since the results are insignificant and consistent with the previous findings, we do not report them. The results are, however, available from the authors upon request.

[16] Since we require the exact items suggested, in this section, we do not consider the variable vm_item.

TABLE 5
Sensitivity Analyses of Compulsory Liability Claims

	Model 1		Model 3	
Threshold	Coefficient of *vm_km*	*p*-value	Coefficient of *vm_time*	*p*-value
15%	−0.0775	0.0387	0.0268	0.5233
16%	−0.0845	0.0244	0.0206	0.6238
17%	−0.0899	0.0138	0.0201	0.6277
18%	−0.0978	0.0096	0.0177	0.6681
19%	−0.0995	0.0035	0.0223	0.5854
20%	−0.2080	0.0026	0.0151	0.7102
21%	−0.0784	0.0404	0.0118	0.7720
22%	−0.0760	0.0481	0.0145	0.7188
23%	−0.0651	0.0473	0.0190	0.6352
24%	−0.0605	0.0579	0.0278	0.4847
25%	−0.0589	0.0628	0.0188	0.6354

could be a signal of the individual's responsibility and the degree of risk aversion. An individual who is more responsible and more risk averse will maintain his/her vehicle properly. Thus, if we allow a lower level of extra kilometers when defining proper maintenance, we might group some individuals who have a high degree of responsibility and/or a high degree of risk aversion but only miss the recommended number of kilometers once as the individuals who have a low degree of responsibility and/or a low degree of risk aversion. Similarly, if we allow a higher level of extra kilometers when defining proper maintenance, some individuals who have a low degree of responsibility and/or a low degree of risk aversion might be pooled with the individuals who have a high degree of responsibility and/or high degree of risk aversion. In both cases, the differences between individuals with vehicles with and without proper maintenance will be lessened, and the negative correlation between proper maintenance and loss probability will decrease.

Table 5 further shows that the coefficients of *vm_time* are positive for all different thresholds ranging from 15 percent to 25 percent. Consistent with the previous findings, the relationship is insignificant for all different levels of thresholds. This finding suggests that *vm_time* might not be a good proxy for the driver's degree of risk aversion or responsibility, since it is also affected by the manufacturer's behavior.

THE CHARACTERISTICS OF THE INDIVIDUAL WHOSE VEHICLE IS PROPERLY MAINTAINED

Our major findings imply that the vehicle maintenance records might be a signal of the insured's characteristics. Therefore, in this section, we further adopt pooled probit regressions to examine the relationship between the proper maintenance and the insured's characteristics. Specifically, the factors that impact the decision to maintain the insured vehicle properly according to the recommended number of kilometers (*vm_km*) and the recommended items (*vm_item*) are examined as

$$\Pr(vm\_km_{it} = 1 | X_{it}, year) = F(X_{it}\phi_x + year_t\phi_y)\text{, and} \tag{7}$$

TABLE 6
Pooled Probit Estimation on the Probability of Proper Maintenance

	vm_km		*vm_item*	
Variables	Coefficient	*p*-value	Coefficient	*p*-value
Intercept	−1.2992	< 0.0001	−1.1086	< 0.0001
age2530	0.0270	0.8072	−0.0416	0.6667
age3060	0.1085	0.3088	−0.0070	0.9398
age60up	0.3286	0.0035	0.1791	0.0700
female	0.0119	0.5018	0.0241	0.0395
married	0.0441	0.0731	−0.0086	0.7650
carage2	0.2068	< 0.0001	0.0658	0.0003
carage3	0.2040	< 0.0001	0.0111	0.5926
carage4	−0.1728	< 0.0001	−0.0041	0.8756
big	−0.0263	0.1770	−0.0215	0.2269
sedan	0.0799	0.1975	0.0173	0.7523
channel_D	−0.8880	< 0.0001	−0.9007	< 0.0001
city	−0.0915	< 0.0001	−0.0729	< 0.0001
north	0.0099	0.6555	0.0142	0.4800
south	0.0611	0.0127	0.0191	0.3966
east	0.1609	0.0015	0.2293	< 0.0001
year2004	0.0513	0.0694	0.1090	< 0.0001
year2005	0.0085	0.7571	0.1177	< 0.0001
year2006	−0.0999	0.0001	0.0294	0.2391
Likelihood ratio	1,921.774		1,852.1803	

$$\Pr(vm\_item_{it} = 1 | X_{it}, year) = F(X_{it}\phi_x + year_t\phi_y). \tag{8}$$

The left-hand side of Table 6 demonstrates the impact of the insured's characteristics on the maintenance records according to the recommended number of kilometers. The results show that senior owners (over 60 years old) are more likely to maintain their vehicles according to the suggested kilometers than young owners (under 20 years old). Married owners also have a higher probability of maintaining their vehicles according to the kilometers than single owners. New cars (less than 3 years old) have a higher probability of being properly maintained than vehicles more than 4 years old. It is usually the case that new cars have higher values than old cars. This finding is consistent with the argument that people have a higher probability of maintaining their property when the value of the property is higher. We also find that if the individuals live in central Taiwan, they have a lower probability of properly maintaining their vehicles.

The right-hand side of Table 6 demonstrates the relationship between individual characteristics and the probability of proper maintenance according to the recommended items. The characteristics that determine whether a vehicle is properly maintained according to the recommended kilometers are similar to those affecting whether a vehicle is properly maintained according to the recommended items. We further find that female owners are more likely to maintain their vehicles according to the suggested

items than male owners, whereas the gender differences are not significant when examining the probability of maintenance according to the suggested kilometers.

CONCLUSION

Much of the literature has explored the determining factors in terms of predicting automobile accidents. Recently, the focus of risk classification factors for automobile accidents has been on the biological, psychological, and behavioral characteristics of human beings. Inspired by the literature, this article seeks to explore new valuable risk prediction factors based on the individual's biopsychological behavior. We examine whether or not maintenance records can predict automobile accidents.

We find that maintenance records could contain valuable information in predicting automobile accidents. We evaluate proper maintenance based on three different criteria: kilometers, time, and items maintained. We find that proper vehicle maintenance measured in kilometers as well as recommended items are negatively correlated with the loss probability by using the previous years' maintenance records. The loss probability of a vehicle properly maintained according to the recommended number of kilometers is reduced by 0.23 percent and that according to the recommended items is reduced by 0.18 percent. Our findings also show that proper vehicle maintenance measured in kilometers is a better proxy. Although we find that vehicle maintenance could reduce loss probability, we cannot reject the hypothesis that maintenance records are not correlated with loss severity.

Our article contributes to the literature by linking the vehicle maintenance records with automobile liability claims. We provide further support for Brockett and Golden's (2007) hypothesis in that people who are more responsible in their vehicle maintenance are also more responsible in their driving behavior, even after adjusting for standard underwriting variables. Brockett and Golden supply the biological and psychological underpinnings from the finances and responsibility with credit domain, whereas we provide additional evidence based on the vehicle maintenance domain.

There are some extensions that could be performed in future studies. First, our data are limited to a particular vehicle brand. Therefore, collecting more data for different brands of vehicle might be useful to better understand the relationship between vehicle maintenance, loss probability, and loss severity. Second, we only focus on liability automobile insurance. The relationship between vehicle maintenance and automobile property insurance is also worth examining. Third, both credit scoring and maintenance records tap the underlying dimension of the biological and/or psychological and social component of individuals. If collecting both types of information were possible, it would have a certain contribution to the credit scoring and insurance debate by examining the relationship between credit scores and vehicle maintenance.

REFERENCES

Andersson, H., 2005, The Value of Safety as Revealed in the Swedish Car Market: An Application of the Hedonic Pricing Approach, *Journal of Risk and Uncertainty*, 30(3): 211-239.

Angers, J., D. Desjardins, G. Dionne, and F. Guertin, 2006, Vehicle and Fleet Random Effects in a Model of Insurance Rating for Fleets of Vehicles, *ASTIN Bulletin*, 36(1): 25-77.

Brockett, P. L., and L. L. Golden, 2007, Biological and Psychobehavioral Correlates of Credit Scores and Automobile Insurance Losses: Toward an Explication of Why Credit Scoring Works, *Journal of Risk and Insurance*, 74(1): 23-63.

Chiappori, P. A., and B. Salanie, 2000, Testing for Asymmetric Information in Insurance Markets, *Journal of Political Economy*, 108(1): 56-78.

Greene, W., 2001, Fixed and Random Effects in Nonlinear Models, Manuscript, Department of Economics, Stern School of Business, New York University.

Iversen, H., and T. Rundmo, 2002, Personality, Risky Driving and Accident Involvement Among Norwegian Drivers, *Personality and Individual Differences*, 33: 1251-1263.

Jindapon, P., and W. S. Neilson, 2007, Higher-Order Generalizations of Arrow–Pratt and Ross Risk Aversion: A Comparative Statics Approach, *Journal of Economic Theory*, 136: 719-728.

Kellison, B., and P. Brockett, 2003, Check the Score Credit Scoring and Insurance Losses: Is There a Connection? *Texas Business Review*, 1-6. World Wide Web: http://www.ic2.utexas.edu/bbr/back-issues/2003-texas-business-review/index.php (accessed October 4, 2010).

Miller, M., and R. A. Smith, 2003, The Relationship of Credit-based Insurance Scores to Private Passenger Automobile Insurance Loss Propensity, An Actuarial Study, Epic Actuaries. World Wide Web: http://www.epicactuaries.com (accessed October 4, 2010).

Tillman, W. A., and G. E. Hobbs, 1949, The Accident-Prone Automobile Driver: A Study of the Psychiatric and Social Background, *American Journal of Psychiatry*, 106: 321-331.