Predicting opportunistic fraud from vehicle maintenance records in automobile theft insurance

Abstract

This paper examines whether the policyholders who properly maintain their vehicles will have less chance of committing opportunistic fraud. Using data from Taiwan, we first confirm that the replacement cost endorsement will induce opportunistic fraud. Moreover, we find that the pattern of opportunistic fraud disappears for the group of policyholders with proper maintenance records, whereas the pattern of opportunistic fraud remains for the group of policyholders without proper maintenance records. Our results are robust to the definition of proper maintenance. The potential explanations are also discussed.

1. Introduction

The literature has provided evidence to support the existence of fraud in different insurance markets.[1] Moreover, fraud is very costly to society. For example, according to the Coalition Against Insurance Fraud's (CAIF) estimation, insurance fraud costs the U.S. property and casualty industry $80 billion per year.[2] The Insurance Research Council (2008) reports that the percentage of claims that appear to involve fraud and buildup increases from between 11 and 15 percent of total payments in 2002 to between 13 and 18 percent in 2007.

To combat insurance fraud, the literature has suggested various strategies. A large number of theoretical studies aim at deriving the optimal auditing strategies under costly state verification settings[3] where the policyholders have private information and the insurance companies can verify the claims with an audit cost. The ability of the insurer to commit to an auditing policy is an important factor in this setting.[4] On the other hand, the growing statistical literature on insurance fraud is mainly aimed at designing a better fraud detection system to lower the cost of fraudulent claims or to decrease the error rate.[5] The insurance companies can therefore determine which claims to audit according to the fraud signals/indicators.

---

[1] Cummins and Tennyson (1992, 1996), Tennyson (1997), Carroll and Abrahamse (2001), Dionne and Gagné (2001, 2002), Dionne and Wang (2011), and Pao et al. (2012) have examined fraud in the automobile insurance market. Dionne and St-Michel (1991) and Butler et al. (1996) looked at fraud in workers' compensation insurance, while Dionne (1984), Sparrow (1996) and Hyman (2001, 2002) examined the health care fraud problem.

[2] Please see http://www.insurancefraud.org/80_billion.htm.

[3] Crocker and Morgan (1997) developed a costly state falsification approach whereby the policyholders expend resources for their claims so as not to be detected.

[4] Picard (1996) investigated the optimal audit strategy in a theoretical model. Schiller (2006) asserted that the efficiency of the audit could be improved through conditioning the information from the detection system under costly state verification. Please see Picard (2000) for a detailed survey.

[5] For example, see Artis et al. (2002), Crocker and Tennyson (2002), Derrig and Weisberg (2003), Viaene et al. (2007), and Ai et al. (2012).

Most fraud indicators employed in the literature[6] are variables related to contract types and claim behavior (e.g., delay in filing the accident claim, abnormally high frequencies of accidents, filing on weekends). However, as mentioned in the literature, individual characteristics could also affect the incentive to defraud. For example, Picard (1996) mentioned that not everyone would defraud when there is a chance to cheat since the moral cost of committing fraud differs among persons. Schiller and Lammers (2010) indicated that some dishonest persons would defraud when there is a chance to successfully cheat, but that some honest persons would never defraud. Accordingly, the individual's personality would determine whether such a person would defraud, as well as the probability that the person would defraud.

This paper therefore seeks to identify a new indicator of fraud which could be attributed to individual characteristics.[7] We propose that the policyholders who properly maintain their vehicles will have fewer chances to commit insurance fraud. Bair, Huang, and Wang (2012) suggested that the observable proper maintenance records play a role similar to credit history[8] which signals the unobservable characteristics, such as the degree of risk aversion and the responsibility of the policyholders. They found that the proper maintenance record is significantly negatively correlated with the claim rate on the compulsory automobile liability insurance. Note that fraud is risky and risk-averse individuals would not accept a lottery which is even fairly priced. Thus, on the basis of Bair, Huang, and Wang (2012), we expect that the observable maintenance behavior might be able to serve as an indicator of fraud.

In this paper, we analyze opportunistic fraud which arises owing to the nature of the

[6] For example, see Tennyson and Salsas-Forn (2002), Viaene et al. (2007) and Dionne et al. (2009).
[7] The insured's characteristics such as the age of insured driver and the gender of the insured are considered in Viaene et al. (2007).
[8] Please see Brockett and Golden (2007) regarding the role of the credit record.

insurance contract: the presence of a replacement cost endorsement. By analyzing the automobile theft insurance in Quebec, Dionne and Gagné (2002) found that, compared to policyholders without the replacement cost endorsement, those with the replacement cost endorsement have significantly higher claim rates on automobile theft insurance in the last two months before their policies lapse. The above evidence is significant only in relation to total theft rather than partial theft, which is consistent with the theoretical prediction that the incentive to defraud is larger in the case of total theft when the replacement cost endorsement is purchased. They thus conclude that the evidence is indicative of opportunistic fraud.[9]

We use the data from one large insurance company together with the data from one big car manufacturer in Taiwan.[10] By combining these two data sources, we are able to obtain information regarding the characteristics of the insured and the insured vehicle, the content of the automobile theft insurance contract, the claim records in relation to the contracts during the policy period, and the maintenance behavior of the insured. The information for the last term is obtained from the data set of the repair center of the vehicle manufacturer, and the information for the other terms is obtained from the data set of the insurance company.

We first confirm that the replacement cost endorsement will induce opportunistic fraud in Taiwan. We find that, compared to those without the replacement cost endorsement, the policyholders with the replacement cost endorsement have a higher total theft claim rate near the end of the policy period but the partial theft claim rate is not significantly higher near the end of the policy year. This finding is consistent with the

[9] Many studies have also found that fraud can be caused by the nature of the insurance contract, such as those of Dionne and Wang (2011) and Pao et al. (2012).

[10] The market share of this big insurance company is over 20%, and the market share of this vehicle manufacturer is about 35%.

findings of Dionne and Gagné (2002).

Furthermore, we provide significant evidence depicting the negative relationship between proper maintenance and opportunistic fraud. We evaluate proper maintenance according to whether the maintenance is carried out based on the recommended number of kilometers.[11] We find that in a sub-sample of policyholders without proper maintenance records, the pattern of opportunistic fraud remains. However, in a sub-sample of policyholders with proper maintenance records, the pattern of opportunistic fraud disappears. We then formally test our hypothesis using the full sample. We find that policyholders without proper maintenance records have more significant patterns of opportunistic fraud than policyholders with proper maintenance records. Furthermore, this conclusion is also robust when we vary the chosen threshold to define proper maintenance.

The remainder of this article is organized as follows. Section 2 introduces the automobile theft insurance in Taiwan. Sections 3 and 4 respectively describe the data we collect and the methodologies adopted. Section 5 provides the empirical results. Section 6 concludes the paper.

2. Market Background

The major design of automobile theft insurance contracts is similar to that of most countries. When the insured vehicle is totally stolen, the insurance company will

[11] The original manufacturer recommends that the vehicle owners maintain their vehicles every 5 thousand kilometers that they are driven. When we define proper maintenance behavior, we define the owners who delay the maintenance of their vehicles by more than 20% of this recommended standard, in other words, do not maintain them until more than 6 thousand kilometers have passed, as those who do not properly maintain their vehicles. The reason why we use this criterion which allows for a delay of 20% is based on the business experience of the repair houses. They mentioned that their customers who followed the recommendations would be back at the repair house before more than 20% of this recommended standard had been reached. However, later in this paper, we will perform a sensitivity analysis in order to vary this criterion for delaying vehicle maintenance.

reimburse the insured by an amount which is the vehicle's value at the beginning of the contract less the monthly depreciation.[12] To avoid this monthly depreciation in the case where a theft event takes place, the insured can further purchase a replacement cost endorsement. With this endorsement, the indemnity will be free from depreciation. In other words, the indemnity is equal to the value of the insured vehicle at the beginning of the contract.

The details of the theft insurance in Taiwan are a little different from those in certain other countries. For example, in Canada in Quebec, theft insurance contracts are two years in length, whereas the contract period is one year in Taiwan. The endorsement of replacement cost can only be chosen for brand new vehicles in Quebec. In Taiwan, the endorsement is an option for all types of insured. Hence, in comparison to Quebec, the monetary incentive from the replacement cost endorsement in Taiwan's automobile theft insurance contract is much lower. Since the contract period is shorter in Taiwan, the value difference between a vehicle's value with depreciation and without depreciation will be less near the end of the policy period in Taiwan than in Quebec. When Quebec's insurance companies only allow the owners of brand new vehicles to choose the replacement cost endorsement, the brand new vehicle's value is much higher than that of the older vehicles, and the depreciation mechanism will result in more value being deducted for these brand new vehicles than for those vehicles that are more than one year old. However, in Taiwan, the insurance companies do not impose such a restriction on the insured vehicles. Hence, the replacement cost endorsement would provide more of a pecuniary incentive in Quebec than in Taiwan.

[12] The depreciation rate is increasing by months. From the first month to the 12th month, the depreciation rates are as follows: 3%, 5%, 7%, 9%, 11%, 13%, 15%, 17%, 19%, 21%, 23%, and 25%.

It is worth mentioning that Taiwan's automobile theft insurance only indemnifies the insured against the loss resulting from total theft. The insurance companies will not indemnify the insured against the loss of auto parts or accessories under a partial theft event unless the insured has further purchased an auto parts and accessories endorsement. Accordingly, when we need to investigate the problems related to partial theft, we are only able to limit the research sample to a smaller sub-sample in which the insured also extend their contract to the auto parts and accessories endorsement. So, in this sub-sample, the partial theft event could be observed by the claim.

According to the reports of the Taiwan Insurance Institute, about 1.5 million vehicles, which accounts for 26% of the total number of vehicles are covered by automobile theft insurance each year in Taiwan. The average written premium of theft insurance per year is NT$ 3.29 billion, which is about US$0.11 billion from 2007 to 2011. The amount is about 9.10% of the total written premium for automobile insurance per year.[13]

3. Data

Our sample comes from two sources: one insurance company and one vehicle manufacturer. Our sample insurance company accounts for over 20% of the market share in Taiwan's automobile insurance market as a whole. This part of the data contains the information regarding the insurance contract, claim information and the variables used in underwriting including the characteristics of the policyholders and the insured vehicle. In this study, we only direct our attention to privately-used sedans and trucks from all of the insurance provider's automobile theft insurance policyholders. From 2001 to 2007, we obtain information on 1,271,095 policyholders from the sample insurance company.

The automobile manufacturer provides us with the maintenance record information.

[13] The information can be obtained from http://www.tii.org.tw/eindex.asp.

This automobile manufacturer accounts for over 35% of the market share and it is the most popular vehicle brand in Taiwan.[14] We use this part of the data to distinguish whether the vehicle is properly maintained or not. Specifically, proper maintenance is defined if the insured vehicle is regularly maintained at the repair center of the original manufacturer as recommended based on the number of kilometers driven.[15] In order to conscientiously link this proper maintenance record to the individual's behavior which reflects his psychological or biological characteristics, we require this record to be observed continuously for a period of two years, which are the current year and the previous year. Hence, the insured who are included in our investigated sample are those who have at least two years of records in our target insurance company. From the raw data that extend from the year 2000 to 2007, we obtain 1,177,549 observations over that period.

To match this information with the maintenance records from our target automobile manufacturer, we only keep the observations of a particular brand of vehicle since the sample manufacturer only maintains its own products. After combining the data from these two sources, we obtain a total of 495,260 observations over the period from 2001 to 2007.[16]

Table 1 provides the definitions of all variables used in this paper. Table 2 provides a statistical description of our full sample and the sub-sample where partial theft insurance

[14] Although we investigate only one brand of vehicle in the market, this brand of vehicle includes many different models from luxury cars through compact cars to small ones. Judging by the engine capacity, 61.6% of our sample consists of small-sized vehicles, 26.7% medium-sized vehicles, and 11.7% large-sized vehicles. This distribution is roughly representative of the vehicle structure in Taiwan.

[15] It is worth mentioning that some vehicle owners could also maintain their vehicles at a private garage. We do not have that information. It can be observed from our sample that about 80% of the vehicles which are serviced by the original manufacturer are not over 3 years old.

[16] The percentages of the number of observations from the year 2001 to the year 2007 are 11%, 12%, 13%, 14%, 17%, 17% and 16%, respectively.

is also purchased. In the full sample, the total theft claim rate is 0.23%. The average claim amount for these total theft claims is US$ 12,637.[17] The variable *replacement* shows that about 39% of the insured who purchased automobile theft insurance also purchased the replacement cost endorsement. More than 19% of the vehicles are properly maintained during our sample period. The variables related to the age of the policyholder (*age2530, age3060 and ageabv60*) show that the individuals who purchased automobile theft insurance are mostly between 30 and 60 years old. The high percentage of married females (*sexf* and *married*) in our sample mainly comes from insurance premium discounts for them in Taiwan. Since the automobile insurance is not underwritten based on the drivers but rather on the vehicle owners, most families register their vehicles and insure the vehicles under a married middle-aged female member of the family. The car age variables (*carage_1, carage_2, carage_3, carage_4* and *carage_5*) show that over 65% of the insured vehicles are not more than three years old. Over half of the insured are registered in city areas (*city*).

We find that most of the structural information, such as the structure of the age, gender and marital status of the insured and the age, usage and registered area of the insured vehicle for this sub-sample is similar to that for the full sample. A minor difference is that there is a higher percentage of medium to large vehicles that also have the auto parts and accessories endorsement. More insured are also found to have purchased the replacement cost endorsement than the full sample. However, a lower percentage of the insured properly maintained their vehicles. The percentage of theft events is higher. A total of 0.68% of the insured policyholders filed a partial theft claim in the policy year.

[17] This is calculated based on an exchange rate of US$1:NT$30.

In addition, Table 2 shows that, among the policyholders who have ever filed a claim, 40.7% of the claims are concentrated in the last three months for the total theft insurance, and 20.2% of the claims are concentrated in the last three months for the partial theft insurance. Table 3 further provides the distribution of claim frequencies for the twelve policy months. In the full sample, we can see that, during the first nine months, the claims seem to be randomly distributed during these months. However, the percentage increases significantly during the last three months. In Table 3, we also observe the partial claim distributions across twelve policy months. From Table 3, it can be observed that the timing of claims is randomly distributed among the twelve policy months in the sub-sample. They are not particularly concentrated at the end or based on any specific timing. This is in sharp contrast to the case of total theft claims.

Dionne and Gagné (2002) disentangle the opportunistic fraud from the *ex ante* moral hazard based on the pecuniary incentive. Opportunistic fraud should be induced by a higher monetary incentive, whereas *ex ante* moral hazard is not based on a sufficient monetary incentive.[18] Thus, Table 3 provides a preliminary observation for fraud. Does this preliminary observation, then, imply that there is a bigger incentive to file the total theft claim near the end of the policy year, especially for those contracts with a replacement cost endorsement? We need a formal test of the latter to confirm this.

4. Methodology

According to Dionne and Gagné (2002), the incentive of the opportunistic insured to defraud is related to the expected utility from fraud and the utility of not to defraud. When the difference between the value of the depreciated vehicle and the indemnity from the replacement cost endorsement contract is enlarged as time passes, the expected utility

[18] We will explain this inference in detail in the following section.

from fraud is also enlarged. This makes the incentive to defraud stronger when the contract gets near to the end of the overall policy period. Furthermore, the incentive of the insurer to detect such fraud is lessened as the time passes because the vehicle's value becomes less and less due to the depreciation. Hence, the equilibrium fraud rate would be higher near the end of the policy period. In addition, this makes the probability of filing a theft claim higher near the end of the policy year for those insured who also have the replacement cost endorsement.

The first task of this paper is to identify opportunistic fraud. We distinguish opportunistic fraud according to whether there is a higher probability of filing a theft claim near the end of the policy year for those insured who also have the replacement cost endorsement. We adopt a two-stage method which is similar to the methodology of Dionne et al. (2001) to test this conditional correlation. We try to identify whether this conditional correlation is positive and significant near the end of the policy year.

As mentioned by Dionne and Gagné (2002), *ex ante* moral hazard can also display the same empirical results as opportunistic fraud. The insured with higher insurance coverage, such as those with a larger replacement cost endorsement, could have a bigger incentive to be less cautious near the end of the policy year. This would also cause the probability of theft to rise near the end of the policy. The only difference between *ex ante* moral hazard and opportunistic fraud is that the latter needs a sufficient monetary incentive and the former does not. Hence, we can disentangle opportunistic fraud from *ex ante* moral hazard by performing further tests on the conditional correlation between the replacement cost endorsement and partial theft claims. If the particularly high conditional correlation near the policy's end is only significant when the claim is filed for total theft instead of for partial theft, it is not a sign of *ex ante* moral hazard, and we can confirm

that it is opportunistic fraud.[19] Thus, we not only perform the conditional correlation test between replacement cost endorsement and the total theft claims, but also perform the conditional correlation test between the replacement cost endorsement and the partial theft claims.

There are three endogenous variables in our empirical model, which include: the decision to properly maintain a vehicle, the decision to file a claim, and the decision regarding the choice of contract. The two-stage Probit regressions are as follows:[20]

1st stage:

$$\text{Prob}(PM_i = 1) = \Phi(X_{1i}\beta_1) \tag{1}$$

$$\text{Prob}(claim_i = 1) = \Phi(X_{1i}\beta_2) \tag{2}$$

2nd stage:

$$\begin{aligned}&\text{Prob}(replacement_i = 1)\\ &= \Phi(\theta_1 \widehat{PM}_i + \theta_2 \widehat{claim}_i + \theta_3 PM_i + \theta_4 claim_i + \theta_5 claim_i \times PM_i\\ &\quad + \theta_6 claim_i \times end_i + \theta_7 peak\_1_i + \theta_8 peak\_2_i + \theta_9 X_{2i}).\end{aligned} \tag{3}$$

In the first stage, PM is the dummy variable of whether to properly maintain the vehicle or not. $claim$ is a dummy variable indicating whether or not a total (partial) theft claim has been filed when the full (sub-sample) sample is used. Φ is the CDF of a normal distribution. X_1 is a vector of independent variables including the characteristics of the insured and the insured vehicle which could be related to the possibility of choosing the insurance contract, the possibility of the vehicle being stolen and/or being

[19] If proper maintenance can reduce the car theft rate and is a hidden action to the insurance companies, then the maintenance behavior could be linked to moral hazard. Using both the structural form approach proposed by Dionne et al. (2001) and the reduced form approach proposed by Chiappori and Salanie (2000), we find evidence of an insignificant relationship between proper maintenance and the car theft rate. The statistics are available from the authors upon request.

[20] Random effect models are adopted.

properly maintained, i.e., the gender, marriage status, age of the policyholder, car age, engine capacity, model and the market value of the insured vehicle, and the registration location. They are listed among the independent variables in Table 1.

In the second stage, $replacement = 1$ when the insured has further purchased the replacement cost endorsement on his/her theft insurance contract, otherwise $replacement = 0$. The most important independent variable is $claim \times end$, which represents the claim near the end of the policy year. end is a dummy variable that equals 1 when the claim is filed during the last three months, and 0 otherwise. If the estimated corresponding coefficient $\hat{\theta}_6$ is positive and significantly different from 0 in the full sample but is insignificant in the sub-sample, then it is evidence of opportunistic fraud.

$\widehat{claim}$ and $\widehat{PM}$ are the estimated probabilities of a claim and proper maintenance from the first stage, respectively. Since there are two peaks for months in which new vehicles are purchased during each calendar year in Taiwan, $claim \times end$ might be clustered within certain calendar months. To control for the calendar effect, we include two dummy variables for these two peaks: $peak\_1$ is one when the claim is filed during January and February, and 0 otherwise, and $peak\_2 = 1$ when the claim is filed during July and August, and 0 otherwise. X_2 represents all the variables of our sample insurance company used in underwriting automobile theft insurance. It includes *carage_1, carage_2, carage_3, carage_4, veh_s, veh_m, vehval, mode_1, sedan, city, north, south,* and *central.*

The second task of this paper, which is the major one, is to show that this opportunistic fraud is less significant for those insured whose behavior is characterized

by proper maintenance. To proceed with the analysis, we first separate our sample into two groups: one is for those with proper maintenance records, and the other is for those without proper maintenance records. We test the above conditional correlation between the replacement cost endorsement and theft claims separately within each group. We not only perform the conditional correlation test separately for the above two groups by using maintenance records underlying the total theft , but also use the records underlying the partial theft claim in order to distinguish opportunistic fraud from *ex ante* moral hazard.

In addition, we perform a formal test by combining the two groups together once again. We add an interaction term between $claim$, end and PM, i.e., $claim \times end \times PM$, in the second stage of the two-stage model to formally test the effect of proper maintenance on opportunistic fraud. Equation (3) is then modified as follows:

$$\begin{aligned}&\text{Prob}(replacement_i = 1)\\ &= \Phi(\gamma_1 \widehat{PM}_i + \gamma_2 \widehat{claim}_i + \gamma_3 PM_i + \gamma_4 claim_i + \gamma_5 claim_i \times PM_i\\ &\quad + \gamma_6 claim_i \times end_i + \gamma_7 claim_i \times end_i \times PM_i + \gamma_8 peak\_1_i\\ &\quad + \gamma_9 peak\_2_i + \gamma_{10} X_{2i}).\end{aligned} \tag{4}$$

The key estimated coefficient is $\hat{\gamma}_7$. If $\hat{\gamma}_7$ is positive and significantly different from 0 in the full sample but is insignificant in the partial theft sample, then it is evidence that an insured vehicle owner who has a proper maintenance record also has less of a tendency to defraud.

There is still one concern that remains. When we define proper maintenance records, we allow for a 20% delay in the recommended maintenance standard. As to whether the conclusions of our analysis are robust while this threshold varies, we also perform sensitivity analysis for this threshold.

5. Empirical results

Table 4 provides evidence of opportunity fraud. The empirical results of the second stage in the two-stage model for total theft insurance and partial theft insurance are listed in Table 4. The first column in both total theft and partial theft insurance is the result of Equation (3) excluding the interaction term $claim \times end$. The second column is the result for Equation (3).

For total theft insurance, the coefficients of $claim$ and $claim \times PM$ in the first column of Table 4 are positive but insignificant. These results imply that after controlling the non-linear relationship between the replacement cost endorsement and claim and between the replacement cost endorsement and proper maintenance, the relationship between risk type and the replacement cost endorsement is insignificant regardless of the maintenance records. In other words, we do not find evidence for asymmetric information for total theft insurance at first glance.

However, is there really not any force of asymmetric information in the total theft insurance market? In the second column of Table 4, the estimated coefficient of the interaction term $claim \times end$ is significantly positive. This result indicates that an insured who files a total theft claim when the timing is near the end of the policy year has a significant higher probability of purchasing the replacement cost endorsement. The coefficient of $claim$ becomes negative after $claim \times end$ is controlled, although it is not significantly different from 0. These findings imply that the total theft insurance market might suffer from either fraud or moral hazard instead of adverse selection. If the phenomenon is because of *ex ante* moral hazard, then we should also observe a significant coefficient of $claim \times end$ in the partial theft insurance sample.

The last column of Table 4 shows that the coefficient of $claim \times end$ is not

significantly different from 0. The insured who files a partial theft claim at the end of the policy year does not have a significantly high probability of purchasing the replacement cost endorsement. These findings in Table 4 jointly support the view that the monetary incentive can affect the policy-year-end claim effect, which is evidence that opportunistic fraud exists.

Table 5 provides a preliminary analysis for whether there is less incentive to engage in opportunistic fraud for those who properly maintain their vehicles. The first two columns in Table 5 show the results of Equation (3) for total and partial theft for the group with proper maintenance records, whereas the last two columns show the results for the group without proper maintenance records. For the group with proper maintenance behavior, the coefficients of $claim$ and $claim \times end$ are significantly different from 0 regardless of whether total theft or partial theft is examined. These findings show that, in the group with proper maintenance records, we do not find evidence for adverse selection, *ex ante* moral hazard or opportunistic fraud.

By contrast, in the group without proper maintenance behavior, the estimated coefficient of $claim \times end$ is significantly positive for total theft, but it is insignificant for partial theft. This means that within the group without proper maintenance records, we still observe opportunistic fraud.

Thus, we can conclude that the phenomenon of opportunistic fraud, instead of *ex ante* moral hazard, exists in the subgroup made up of the insured without proper maintenance records. The insured who do not properly maintain their vehicles have the incentive to commit opportunistic fraud.

The results of the formal test of the incentive to engage in opportunistic fraud for those insured who properly maintain their vehicles versus the insured who do not

properly maintain them as represented is shown in Table 6. The estimated coefficient of $claim \times end \times PM$ is negative and significantly different from 0 for total theft insurance, whereas it is insignificant for partial theft insurance. Thus, we can conclude that the insured who would properly maintain their vehicles are less likely to commit opportunistic fraud which is induced by the replacement cost endorsement.

In Table 7, we perform sensitivity analysis for different criteria of delay in relation to the recommended maintenance. We find that the results are quite robust when we vary the threshold to a more stringent standard, such as moving from 20% to 19%, or when we vary the threshold to a less stringent standard, such as moving from 20% to 22%. By using the criterion of 20% we can arrive at a satisfying and robust conclusion when we identify the time pattern of conditional correlation for opportunistic fraud. Furthermore, the inference of our hypothesis that the insured who properly maintains his or her vehicle is less likely to commit fraud also receives the significant and strongest support when we set this criterion as 20%.

6. Conclusion

The literature has documented that, when there is fraud that has existed in the insurance market and an audit is a necessary strategy, employing signals revealed by the insured could help to improve the audit's efficiency and lead to a saving in cost. There are many different suspicious fraud signals revealed by the claimant that have been discussed in the literature. However, they seldom focus on the insured's other behavior beyond the insurance.

To respond to the view of Picard (1996) and Schiller and Lammers (2010) that the characteristics of individuals could affect the incentive to engage in fraud, which coincides with the findings in Brockett and Golden (2007), and Bair, Huang, and Wang

(2012) that the credit records or maintenance records could reflect some of the individual's unobservable characteristics, such as the degree of risk aversion and the responsibility of the policyholders, we employ the individual's proper vehicle maintenance behavior to predict the tendency to commit fraud.

We have found that individuals who properly maintain their vehicles tend to be less willing to commit an opportunistic fraud that is induced by the replacement cost endorsement. This conclusion is robust when we vary the criterion of the threshold to define proper maintenance. Our paper contributes to the literature in that we identify a relevant fraud signal that the individual's proper maintenance record could be a good signal of less of a tendency to engage in fraud.

One possible reason why maintenance records could predict the opportunistic fraud is that the records reflect the characteristic of the policyholders, such as risk tolerance and responsibility. Brockett and Golden (2007) have shown that credit records could also reflect the characteristic of the policyholders and thus have certain prediction powers on automobile insurance losses. Thus, a future work of collecting credit records to distinguish the effect of credit scores from that of maintenance records on predicting opportunistic fraud will be fruitful.

In addition, the negative relationship between fraud and proper maintenance could be also enhanced when the individuals confront financial difficulties. When the insured are financially constrained, they might not be able to maintain their vehicles. Meanwhile, they commit fraud to escape from financial difficulties. Dionne and Wang (2012) have proved that fraud fluctuates in an opposite pattern to the business cycle, which implies that opportunistic individuals tend to have a bigger incentive to commit fraud when they are confronted with financial difficulties. Thus, it is also interesting to collect information

regarding the policyholders' financial situation, and to further to examine the relationship between proper maintenance and opportunistic fraud which is conditional on the policyholders' financial situation.

References

Ai, J., Brockett, P. L., Golden, L. L., and Guillén, M., 2012, A robust unsupervised method for fraud rate estimation, ***Journal of Risk and Insurance***, forthcoming.

Artis, M., Ayuso, M., and Guillen, M., 2002, Detection of automobile insurance fraud with discrete choice models and misclassified claims, ***Journal of Risk and Insurance***, 69: 325-340.

Bair, S. N, Huang, R. J., and Wang, K. C., 2012, Can vehicle maintenance records predict automobile accidents? ***Journal of Risk and Insurance***, 79: 567-584.

Brockett, P. L. and Golden, L. L., 2007, Biological and psychobehavioral correlates of credit scores and automobile insurance losses: Toward an explication of why credit scoring works, ***The Journal of Risk and Insurance***, 74:1; 23-63.

Butler, R. J., Durbin, D., and Helvacian, M. N., 1996, Increasing claims for soft tissue injuries in workers' compensation: Cost shifting and moral hazard, ***Journal of Risk and Uncertainty***, 13: 73-87.

Carroll, S. and Abrahamse, A., 2001, The frequency of excess auto personal injury claims, ***American Law and Economics Review***, 3: 228-249.

Chiappori, P.A. and Salanie, B., 2000, Testing for asymmetric information in insurance markets, ***Journal of Political Economy***, 108: 56-78.

Cummins, J. D. and Tennyson, S., 1992, Controlling automobile insurance costs, ***Journal of Economic Perspectives***, 6: 95-115.

Cummins, J. D. and Tennyson, S., 1996, Moral hazard in insurance claiming: Evidence from automobile insurance, ***Journal of Risk and Uncertainty***, 12: 29-50.

Crocker, K. J., Morgan, J., 1998, Is honesty the best policy? Curtailing insurance fraud through optimal incentive contracts, ***Journal of Political Economy***, 106: 355-375.

Crocker, K. J. and Tennyson, S., 2002, Insurance fraud and optimal claims settlement strategies. ***Journal of Law and Economics***, 45: 469-507.

Derrig, R. A. and Weisberg, H. I., 2003, Auto bodily injury claim settlement in Massachusetts, Document, Automobile Insurers Bureau of Massachusetts, Boston.

Dionne, G.., 1984, The effect of insurance on the possibility of fraud, ***Geneva Papers on Risk and Insurance***, 9: 304-321.

Dionne, G., Giuliano, F., and Picard, P., 2009, Optimal auditing with scoring: Theory and application to insurance fraud, ***Management Science***, 22: 58-70.

Dionne, G., Gourieroux, C., and Vanasse, C. 2001, Testing for evidence of adverse selection in the automobile insurance market: A comment. ***Journal of Political Economy***, 109(2): 444-455.

Dionne, G., Gagné, R., 2001, Deductible contracts against fraudulent claims: evidence from automobile insurance, ***The Review of Economics and Statistics***, 83: 290-301.

Dionne, G. and Gagné, R., 2002, Replacement cost endorsement and opportunistic fraud in automobile insurance, ***Journal of Risk and Uncertainty***, 24: 213-230.

Dionne, G. and St-Michel, P., 1991, Workers' compensation and moral hazard, ***Review of Economics and Statistics***, 73: 236-244.

Dionne, G. and Wang, K. C., 2012, Does opportunistic fraud in automobile theft insurance fluctuate with the business cycle? ***Journal of Risk and Uncertainty***, forthcoming.

Hyman, D. A., 2001, Health care fraud and abuse: Market change, social norms, and the trust "reposed in the workmen," ***Journal of Legal Studies***, 30: 531-567.

Hyman, D. A., 2002, HIPAA and health care fraud: An empirical perspective, ***CATO Journal***, 22: 151-178.

Insurance Research Council, 2008, Fraud and buildup in auto injury insurance claims, Report by IRC (Malvern, PA).

Pao, T. I., Tzeng, L. Y., and Wang, K. C., 2012, Typhoons and opportunistic fraud—Claim patterns of automobile theft insurance in Taiwan, ***Journal of Risk and Insurance***, forthcoming.

Picard, P., 1996, Auditing claims in the insurance market with fraud: The credibility issue, ***Journal of Public Economics***, 63: 27-56.

Picard, P., 2000, Economic analysis of insurance fraud, in: G. Dionne, ed., ***Hand Book of Insurance***, 315-363.

Schiller, J., 2006, The impact of insurance fraud detection systems, ***Journal of Risk and Insurance***, 73: 421-438.

Schiller, J. and Lammers, F., 2010, Contract design and insurance fraud: An experimental investigation, ***World Risk and Insurance Economics Congress 2010***, Singapore.

Sparrow, M., 1996, *License to Steal* (Boulder, Colorado: Westview Press).

Tennyson, S., 1997, Economic institutions and individual ethics: A study of consumer attitudes toward insurance fraud, ***Journal of Economic Behavior and Organization***, 32: 247-265.

Tennyson, S. and Salsas-Forn, P. ,2002, Claims auditing in automobile insurance: fraud detection and deterrence objectives, ***Journal of Risk and Insurance***, 69: 289-308.

Viaene, S., Ayuso, M., Guillen, M., Van Gheel, D., and Dedene, G., 2007, Strategies for detecting fraudulent claims in the automobile insurance industry, ***European Journal of Operational Research***, 176: 565-583.

Table 1 Variable Definitions

Variable	Definition
Dependent variables:	
PM	A dummy variable that equals 1 when the individual has properly maintained his/her vehicle in the current year and previous year, and 0 otherwise.
claim	A dummy variable that equals 1 when the insured has filed a total theft claim, and 0 otherwise. [1]
replacement	A dummy variable that equals 1 when the insured with automobile theft insurance is also covered by the replacement cost endorsement, and 0 otherwise.
Independent variables:	
end	A dummy variable that equals 1 when the claim is filed during the last three months, and 0 otherwise.
peak_1	A dummy variable that equals 1 when the accident date is during the first automobile theft rate peak period in the whole calendar year, and 0 otherwise.
peak_2	A dummy variable that equals 1 when the accident date is during the second automobile theft rate peak period in the whole calendar year, and 0 otherwise. [2]
carage_1	A dummy variable that equals 1 when the car is one year old, and 0 otherwise.
carage_2	A dummy variable that equals 1 when the car is two years old, and 0 otherwise.
carage_3	A dummy variable that equals 1 when the car is three years old, and 0 otherwise.
carage_4	A dummy variable that equals 1 when the car is four years old, and 0 otherwise. [3]
veh_s	A dummy variable that equals 1 when the insured car has an engine capacity that equals or is under 1800 c.c., and 0 otherwise.
veh_m	A dummy variable that equals 1 when the insured car has an engine capacity that is above 1800 c.c. and equals or is under 2000 c.c., and 0 otherwise. [4]
vehval	The logarithm of per ten thousand dollars of vehicle market value at the beginning of the contract.
mode_1	A dummy variable that equals 1 when the car is mode_1 of our target brand vehicle, and 0 otherwise. [5]
sedan	A dummy variable that equals 1 when the car is a sedan and is for non-commercial or for long-term rental purposes, and 0 otherwise.
age2530	A dummy variable that equals 1 if the insured is between the ages of 25 and 30, and 0 otherwise.
age3060	A dummy variable that equals 1 if the insured is between the ages of 30 and 60, and 0 otherwise.
ageabv60	A dummy variable that equals 1 if the insured is above 60 years old, and 0 otherwise. [6]

Table 1 (Continued)

Variable	Definition
sexf	A dummy variable that equals 1 if the insured is female, and 0 otherwise.
married	A dummy variable that equals 1 if the insured is married, and 0 otherwise.
city	A dummy variable that equals 1 when the owner of the car lives in a city, and 0 otherwise.
north	A dummy variable that equals 1 when the car is registered in the north of Taiwan, and 0 otherwise.
south	A dummy variable that equals 1 when the car is registered in the south of Taiwan, and 0 otherwise.
central	A dummy variable that equals 1 when the car is registered in the central part of Taiwan, and 0 otherwise. [7]

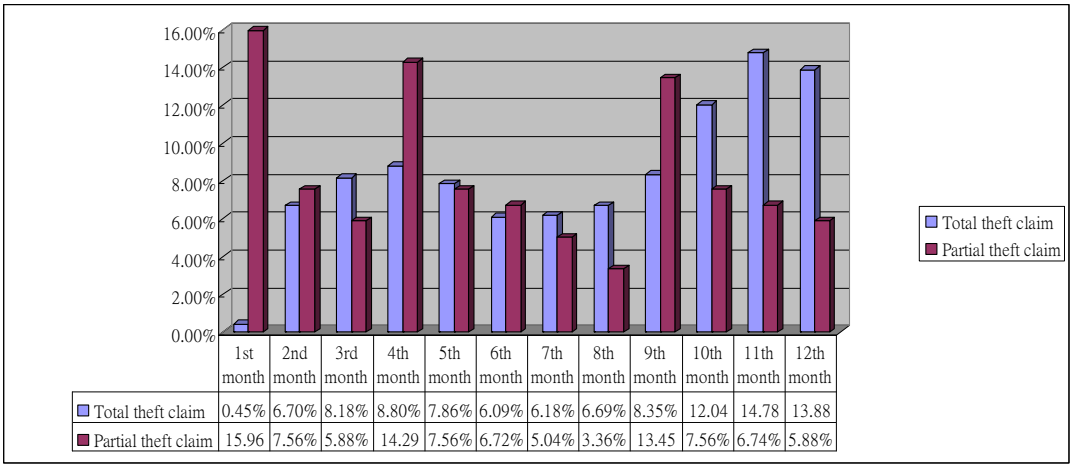
Note: 1. The average claim amount for these total theft claims is NT$ 382,950 (US$ 12,637).
2. There are two automobile theft rate peaks in the whole calendar year. One is from January to February. The other is from July to August.
3. The counterpart of carage_1, carage_2, carage_3 and carage_4 is the vehicle which is over 4 years old.
4. The counterpart of veh_s and veh_m is the vehicle whose engine capacity is over 2000 c.c.
5. Our research targets one particular vehicle brand. Within this brand, the insurer further separates the vehicles into two categories according to the vehicle model when the insurer underwrites or sets the premium price. We name *mode_1* as one of the model categories.
6. The counterpart of age2530, age3060, and ageabv60 is the insured who is younger than 25 years old.
7. The counterpart of north, south, and central is the vehicle which is registered in the east of Taiwan.

Table 2 Descriptive Statistics

Variables	Full sample		Partial theft sub-sample	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
claim	0.0023	0.0484	0.0068	0.0824
replacement	0.3896	0.4877	0.4258	0.4945
PM	0.1839	0.3874	0.1556	0.3625
*end**	0.4071	0.5022	0.2018	0.3797
*peak_1**	0.2486	0.2414	0.2429	0.2500
*peak_2**	0.2946	0.2627	0.2024	0.2418
carage_1	0.2191	0.4136	0.1211	0.3262
carage_2	0.1764	0.3812	0.1668	0.3728
carage_3	0.1411	0.3481	0.1696	0.3753
carage_4	0.1152	0.3192	0.1483	0.3554
veh_s	0.6164	0.4863	0.4760	0.4994
veh_m	0.2674	0.4426	0.3053	0.4605
vehval	4.4671	1.4110	4.5918	1.3241
mode_1	0.5587	0.4965	0.6100	0.4878
sedan	0.9846	0.1232	0.9823	0.1320
age2530	0.0500	0.2179	0.0425	0.2018
age3060	0.8950	0.3065	0.8955	0.3060
ageabv60	0.0495	0.2169	0.0565	0.2310
sexf	0.6741	0.4687	0.6475	0.4778
married	0.9253	0.2629	0.9284	0.2579
city	0.5242	0.4994	0.5441	0.4981
north	0.4734	0.4993	0.4871	0.4998
south	0.2818	0.4499	0.2226	0.4160
central	0.2271	0.4190	0.2747	0.4464
Observations	495260		17069	

Note: * The basic statistics for these variables are calculated only for the observations with claims.

Table 3 The claim frequencies during the twelve policy months



	1st month	2nd month	3rd month	4th month	5th month	6th month	7th month	8th month	9th month	10th month	11th month	12th month
Total theft claim	0.45%	6.70%	8.18%	8.80%	7.86%	6.09%	6.18%	6.69%	8.35%	12.04	14.78	13.88
Partial theft claim	15.96	7.56%	5.88%	14.29	7.56%	6.72%	5.04%	3.36%	13.45	7.56%	6.74%	5.88%

Table 4 The Probit regression on the replacement cost endorsement

This table presents the second stage results of the two-stage model (Equation (3)) to test the conditional correlation between the theft claim and the choice of the replacement cost endorsement. The dependent variable is *replacement*, which equals 1 if the replacement cost endorsement is chosen and 0 if it is not. $\widehat{PM}$ and $\widehat{claim}$ are estimated from the first stage. The estimation is presented in the Appendix, Table A1. See Table 1 for the definitions of the independent variables.

Variables	Total theft				Partial theft			
	(1)		(2)		(1)		(2)	
	Est. coef	P-value	Est. coef	P-value	Est. coef	P-value	Est. coef	P-value
constant	-0.8348	<.0001	-0.8346	<.0001	-0.6347	0.0012	-0.6343	0.0012
$\widehat{PM}$	0.8772	<.0001	0.8771	<.0001	0.6083	0.0094	0.6085	0.0093
$\widehat{claim}$	96.2905	<.0001	96.2849	<.0001	8.1146	<.0001	8.1174	<.0001
PM	0.0047	0.3966	0.0047	0.4012	0.0986	0.0023	0.0986	0.0023
claim	0.0586	0.4301	-0.0674	0.4337	0.3673	0.0518	0.3779	0.0547
$claim \times PM$	0.2871	0.1034	0.2757	0.1204	-0.5270	0.2037	-0.5235	0.2081
$claim \times end$			0.3908	0.0033			-0.0745	0.8460
peak_1	-0.2033	0.3394	-0.1643	0.4446	-0.1182	0.7497	-0.1157	0.7551
peak_2	0.0716	0.6606	0.0899	0.5847	-0.6495	0.1417	-0.6331	0.1578
carage_1	0.4282	<.0001	0.4281	<.0001	0.4590	<.0001	0.4591	<.0001
carage_2	0.0972	<.0001	0.0973	<.0001	0.2694	<.0001	0.2693	<.0001
carage_3	0.0742	<.0001	0.0742	<.0001	0.1890	<.0001	0.1890	<.0001
carage_4	0.0983	<.0001	0.0984	<.0001	0.1395	0.0004	0.1395	0.0004
veh_s	-0.5160	<.0001	-0.5159	<.0001	-0.6064	<.0001	-0.6064	<.0001
veh_m	-0.5099	<.0001	-0.5098	<.0001	-0.5448	<.0001	-0.5448	<.0001
vehval	0.0208	<.0001	0.0207	<.0001	0.0174	0.0422	0.0174	0.0421
sedan	0.0955	0.0007	0.0955	0.0007	0.1478	0.3204	0.1473	0.3222
mode_1	0.0191	0.0066	0.0190	0.0066	-0.2295	<.0001	-0.2295	<.0001
city	-0.1449	<.0001	-0.1450	<.0001	0.0809	0.0009	0.0810	0.0009
north	0.5083	<.0001	0.5082	<.0001	0.6235	<.0001	0.6234	<.0001
south	-0.0506	0.0019	-0.0507	0.0019	-0.0806	0.4398	-0.0805	0.4399
central	0.5302	<.0001	0.5301	<.0001	-0.0651	0.5312	-0.0652	0.5311
Log Likelihood	-2877.09		-2877.93		-793.871		-793.851	

Table 5 The Probit regression on the choice of the replacement cost endorsement for the groups with and without proper maintenance records

This table presents the second stage results of the two-stage model (Equation (3)) to test the conditional correlation between the theft claim and the choice of the replacement cost endorsement for the groups with and without proper maintenance records. The dependent variable is *replacement*, which equals 1 if the replacement cost endorsement is chosen and 0 if it is not. $\widehat{PM}$ and $\widehat{claim}$ are estimated from the first stage. The estimation is presented in the Appendix, Table A2. See Table 1 for the definitions of the independent variables.

Variables	Sub-sample with proper maintenance				Sub-sample without proper maintenance			
	Total theft		Partial theft		Total theft		Partial theft	
	Est. coef	P-value	Est. coef	P-value	Est. coef	P-value	Est. coef	P-value
constant	-0.3632	0.0008	-0.5602	0.0053	-0.9375	<.0001	-0.6851	0.0011
$\widehat{claim}$	16.4076	0.0395	3.5992	0.1254	14.2000	0.0521	7.9297	<.0001
claim	0.3519	0.1461	-0.0633	0.8976	-0.0901	0.3147	0.3678	0.0704
$claim \times end$	0.1810	0.6187	0.3840	0.7098	0.4458	0.0020	-0.2361	0.5724
peak_1	-3.8345	0.8343	-4.9008	0.9691	-0.0689	0.7559	0.0178	0.9636
peak_2	-0.1077	0.7740	-0.6645	0.5177	0.1089	0.5546	-0.6777	0.1808
carage_1	0.4586	0.0059	4.0436	0.9745	0.3163	<.0001	0.3973	<.0001
carage_2	0.0881	<.0001	0.0440	0.6044	0.3595	<.0001	0.4126	<.0001
carage_3	0.0858	<.0001	0.0088	0.9097	0.2860	<.0001	0.3323	<.0001
carage_4	0.0621	0.0001	-0.0086	0.9189	0.2188	<.0001	0.2110	<.0001
veh_s	-0.2692	<.0001	-0.3512	0.0010	-0.4544	<.0001	-0.5791	<.0001
veh_m	-0.1396	<.0001	-0.1618	0.1359	-0.4670	<.0001	-0.5409	<.0001
vehval	-0.0081	0.0547	0.0030	0.9030	0.0254	<.0001	0.0204	0.0261
sedan	-0.0319	0.7296	0.3557	0.5140	0.1241	<.0001	0.1540	0.3247
mode_1	-0.2216	<.0001	-0.2177	0.0020	-0.0420	<.0001	-0.3216	<.0001
city	-0.1859	<.0001	0.0873	0.1268	-0.1103	<.0001	0.0875	0.0007
north	0.3473	<.0001	0.3471	0.1776	0.5703	<.0001	0.7047	<.0001
south	-0.1075	0.0039	-0.1863	0.4726	-0.0301	0.1003	-0.0502	0.6648
central	0.6511	<.0001	-0.1484	0.5640	0.5042	<.0001	-0.0371	0.7486
Loglikelihood	-5404.09		-1555.27		-2305.81		-1809.68	

Table 6 The Probit regression on the choice of the replacement cost endorsement by considering the effect of proper maintenance on the policy-year-end claims

This table presents the second stage results of the two-stage model (Equation (4)) to test the conditional correlation among proper maintenance, the theft claim and the choice of the replacement cost endorsement. The dependent variable is *replacement*, which equals 1 if the replacement cost endorsement is chosen and 0 if it is not. $\widehat{PM}$ and $\widehat{claim}$ are estimated from the first stage. The estimation is presented in the Appendix, Table A1. See Table 1 for the definitions of the independent variables.

Variables	**Total theft**		**Partial theft**	
	Est. coef	**P-value**	**Est. coef**	**P-value**
constant	-0.8346	<.0001	-0.6310	0.0013
$\widehat{PM}$	0.8771	<.0001	0.6103	0.0091
$\widehat{claim}$	96.2764	<.0001	8.1461	<.0001
PM	0.0047	0.4018	0.0986	0.0023
claim	-0.0840	0.3381	0.4112	0.0400
$claim \times PM$	0.4008	0.0647	-0.6997	0.1267
$claim \times end$	0.4474	0.0019	-0.2183	0.5995
$claim \times end \times PM$	-0.3750	0.0133	0.7838	0.4279
peak_1	-0.1591	0.4595	-0.1099	0.7672
peak_2	0.0823	0.6180	-0.7190	0.1136
carage_1	0.4281	<.0001	0.4589	<.0001
carage_2	0.0973	<.0001	0.2691	<.0001
carage_3	0.0743	<.0001	0.1883	<.0001
carage_4	0.0984	<.0001	0.1390	0.0004
veh_s	-0.5159	<.0001	-0.6069	<.0001
veh_m	-0.5098	<.0001	-0.5454	<.0001
vehval	0.0207	<.0001	0.0174	0.0424
sedan	0.0955	0.0007	0.1452	0.3294
mode_1	0.0190	0.0066	-0.2296	<.0001
city	-0.1450	<.0001	0.0811	0.0009
north	0.5082	<.0001	0.6231	<.0001
south	-0.0507	0.0019	-0.0810	0.4375
central	0.5301	<.0001	-0.0655	0.5292
Log Likelihood	-2877.93		-793.846	

Table 7 Sensitivity analysis by varying the criterion of proper maintenance

Criterion of delay	Second stage (regression (4))	
	$claim \times end$	$claim \times end \times maintenance$
17%	0.4460	-0.0024
	(0.0048)	(0.8860)
18%	0.4458	-0.1808
	(0.0100)	(0.2699)
19%	0.4469	-0.3744
	(0.0028)	(0.0169)
20%	0.4474	-0.3750
	(0.0019)	(0.0133)
21%	0.4489	-0.3711
	(0.0010)	(0.0193)
22%	0.4389	-0.3510
	(0.0033)	(0.0330)
23%	0.4372	-0.2977
	(0.0030)	(0.1072)
24%	0.4381	-0.1032
	(0.0033)	(0.3104)
25%	0.4365	-0.0039
	(0.0043)	(0.8351)

Note: When we vary the chosen threshold by 16% to define proper maintenance, there is no total theft claim in the group of policyholders with proper maintenance records. Hence, it is meaningless to conduct the test when the threshold is more stringent than 16%.

Appendix A

Table A1 The Probit regressions on the maintenance and claim

This table presents the first stage results of the two-stage model (Equations (1) and (2)). The dependent variables are PM, which is the dummy variable of whether to properly maintain the vehicle or not, and $claim$, which is defined according to whether a total (partial) theft claim is filed when the total theft (partial theft) sample is used. See Table 1 for the definitions of the independent variables.

Variables	Total theft				Partial theft			
	Proper maintenance		Claim		Proper maintenance		Claim	
	Est. coef	P-value	Est. coef	P-value	Est. coef	P-value	Est. coef	P-value
constant	-2.1158	<0.0001	-4.9785	<.0001	-1.8187	<0.0001	-6.9789	<.0001
carage_1	-2.2198	<0.0001	0.3027	<.0001	-2.2341	<0.0001	0.8351	<.0001
carage_2	0.8236	<0.0001	0.1451	0.0055	0.5072	<0.0001	0.7175	<.0001
carage_3	0.7655	<0.0001	0.1281	0.0199	0.6088	<0.0001	0.5341	0.0002
carage_4	0.5261	<0.0001	-0.0018	0.9780	0.3892	<0.0001	0.1269	0.5174
veh_s	0.4486	<0.0001	-0.0584	0.3098	0.5690	<0.0001	0.2299	0.1325
veh_m	0.4188	<0.0001	0.1455	0.0126	0.5309	<0.0001	0.5759	<.0001
vehval	-0.0023	0.2730	-0.0570	<.0001	0.0045	0.7250	-0.0367	0.2385
sedan	0.1860	<0.0001	2.6396	0.9498	0.1844	0.4080	-0.2072	0.6225
mode_1	-0.4267	<0.0001	-0.1042	0.0072	-0.5721	<0.0001	0.5079	<.0001
age2530	0.2489	<0.0001	-0.2664	0.0327	-0.1964	0.3930	0.1456	0.7612
age3060	0.3657	<0.0001	-0.4072	0.0005	-0.1113	0.6140	-0.1972	0.6704
ageabv60	0.3142	<0.0001	-0.4210	0.0027	-0.0360	0.8760	-0.6472	0.2675
sexf	0.0749	<0.0001	-0.0018	0.9558	0.0605	0.0630	-0.0275	0.7656
married	0.0348	0.0010	-0.0368	0.5068	0.0726	0.2110	-0.1496	0.3150
city	0.1185	<0.0001	0.0011	0.9725	0.1474	<0.0001	-0.0501	0.5772
north	0.0748	<0.0001	-0.0962	0.3997	0.0114	0.9310	3.7980	0.9857
south	-0.0199	0.3330	-0.0261	0.8191	-0.0585	0.6570	3.6779	0.9861
central	-0.0137	0.5120	0.0589	0.6087	0.0037	0.9770	3.7544	0.9858
Log Likelihood	-1512.006		-3384.89		-471.992		-441.424	

Table A2 The Probit regressions on the maintenance and claim for the groups with and without proper maintenance records

This table presents the first stage results of the two-stage model (Equations (1) and (2)) for the groups with and without proper maintenance records. The dependent variables are PM, which is the dummy variable of whether to properly maintain the vehicle or not, and $claim$, which is defined according to whether a total (partial) theft claim is filed when the total theft (partial theft) sample is used. See Table 1 for the definitions of the independent variables.

Variables	**Sub-sample with proper maintenance**				**Sub-sample without proper maintenance**			
	Total theft		**Partial theft**		**Total theft**		**Partial theft**	
	Est. coef	**P-value**	**Est. coef**	**P-value**	**Est. coef**	**P-value**	**Est. coef**	**P-value**
constant	-5.1007	0.0063	-10.8456	0.0756	-5.0025	0.0082	-10.6761	0.0728
carage_1	-3.7461	0.9984	-2.4387	0.9983	0.3185	<.0001	0.8817	<.0001
carage_2	-0.1659	0.2391	0.5577	0.1844	0.1650	0.0064	0.7866	<.0001
carage_3	-0.1717	0.2169	0.2954	0.4984	0.1666	0.0077	0.5899	0.0003
carage_4	-0.5599	0.0112	-0.1735	0.7650	0.0733	0.2809	0.1612	0.4657
veh_s	0.2490	0.3696	-0.5482	0.1555	-0.0818	0.1702	0.3680	0.0346
veh_m	0.3919	0.1633	-0.3392	0.3577	0.1372	0.0229	0.7311	<.0001
vehval	-0.0998	0.0020	-0.0151	0.8939	-0.0535	<.0001	-0.0344	0.3009
sedan	3.2453	0.9976	2.5729	0.9933	2.6344	0.9516	-0.2367	0.5827
mode_1	-0.1323	0.2331	0.4240	0.1173	-0.1186	0.0051	0.5546	<.0001
age2530	-0.3406	0.3718	-0.5776	0.4949	-0.2591	0.0502	3.8791	0.9891
age3060	-0.5287	0.1451	-1.1584	0.1373	-0.3933	0.0015	3.5530	0.9900
ageabv60	-3.8018	0.9883	-4.1027	0.9526	-0.3592	0.0143	3.1768	0.9910
sexf	-0.0120	0.8957	0.1040	0.7155	-0.0029	0.9357	-0.0465	0.6449
married	0.0201	0.8972	0.1664	0.7538	-0.0446	0.4538	-0.1973	0.2174
city	-0.0550	0.5134	-0.1907	0.4561	0.0116	0.7272	-0.0201	0.8395
north	-0.0692	0.8247	3.4321	0.9812	-0.1022	0.4049	3.5714	0.9789
south	0.0874	0.7791	3.0941	0.9831	-0.0483	0.6943	3.5064	0.9792
central	0.0061	0.9848	0.0064	1.0000	0.0612	0.6198	3.6466	0.9784
Loglikelihood	-4670.96		-418.46		-2912		-429.66	