The Journal of Risk and Insurance

© The Journal of Risk and Insurance, 2013, Vol. 81, No. 4, 735–756
DOI: 10.1111/j.1539-6975.2013.01517.x

Heterogeneity of the Accident Externality from Driving

Rachel J. Huang
Larry Y. Tzeng
Kili C. Wang

Abstract

This article examines the accident externality from driving in terms of loss probability and severity by using a unique individual-level data set with more than 3 million observations from Taiwan. Two types of accident externality are, respectively, measured: the average number of kilometers driven per month per vehicle and the total number of speeding tickets per month. For both variables, we find significant evidence to support the existence of the accident externality. Moreover, we find that the accident externality is heterogeneous in terms of the vehicles' characteristics.

Introduction

The risk of a specific driver is affected by other drivers' driving behavior. This is referred to as the accident externality from driving. Such an externality could be very costly to a society and has received much attention in the literature. For example, by using aggregate panel data for the United States, Edlin and Karaca-Mandic (2006) provide intriguing evidence to support the existence of an accident externality from driving. They find that to correct the substantial accident externalities, a Pigouvian tax could raise over $220 billion per year nationally. By adopting Edlin and Karaca-Mandic's methodology, Saito, Kato, and Shimane (2007) also find evidence of a positive and significant externality in Japan. The estimated nationwide Pigouvian tax is about $16–$51 billion in Japan.

In complementing the above literature that estimates the total size of the accident externality, this article studies two important questions that have so far not been explored to any significant extent in the literature. First, how does the externality affect

Rachel J. Huang is an Associate Professor, Graduate Institute of Finance, National Taiwan University of Science and Technology, Taiwan; Research Fellow, Risk and Insurance Research Center, College of Commerce, National Chengchi University. Larry Y. Tzeng is a Professor, Department of Finance, National Taiwan University, Taiwan; Research Fellow, Risk and Insurance Research Center, College of Commerce, National Chengchi University. Kili C. Wang is an Associate Professor, Department of Insurance, Tamkang University, Taiwan; Research Fellow, Risk and Insurance Research Center, College of Commerce, National Chengchi University. Kili Wang can be contacted via e-mail: kili@mail.tku.edu.tw.

the individual's loss probability and loss severity? Second, who suffers more due to other people's driving? In other words, this article seeks to examine the heterogeneity of the accident externality. The answers to these two questions can guide a social-welfare maximizing government in delicately coping with the externality. It is because information on the accident externality both in terms of frequency and severity is necessary for the government to push the private optimum to the social optimum when drivers are risk averse.[1] In addition, the government could directly compensate the identified victims to improve the social welfare.

Since the heterogeneity of the accident externality from driving cannot be analyzed through aggregate data, we adopt individual-level data. We hand-collect our data by integrating data from a vehicle manufacturer with data from an insurance company in Taiwan. Our insurance data include both the occurrence and the amount of money involved in the accident.[2] We are therefore in a position to investigate the impact of the accident externality on the frequency and severity separately. Our insurance data also contain the individuals' demographic variables that can be used to analyze the heterogeneity of the accident externality.

We use two variables to measure the accident externality.[3] One is the average number of kilometers driven per vehicle since the more kilometers that other drivers cover, the greater the potential for each driver to risk causing an accident. Our data from the vehicle manufacturer contain the kilometers driven for each vehicle. Thus, we could estimate the accident externality conditional on the individual's own driving. Furthermore, even if the average number of kilometers driven per vehicle is high, it might not necessarily mean that the risk is higher if others drive at a reasonable speed.[4] Since speed is one of the major risk factors associated with driving, we further adopt the total speeding tickets per month in Taiwan as another variable for the accident externality.

The major findings are as follows. First, we confirm the existence of the accident externality arising from driving in Taiwan both on the average number of kilometers driven per month per vehicle and on the total number of speeding tickets per month. Moreover, we find that the accident externality exists in terms of both the accident probability and accident severity. With respect to loss probability, an individual will increase her accident probability by 0.2937 percent per month when the average number of kilometers driven per vehicle per month increases by 100 km. Furthermore, the probability will increase by 0.0081 percent when the total number of speeding tickets increases by 1,000 per month. With respect to loss severity, the loss severity will increase by $8.2 per vehicle per month when the average number of kilometers driven per vehicle per month increases by 100 km. We also find that it will increase by $1.8 per vehicle per month when the total number of speeding tickets increases by

[1]Please see the Appendix for the details.
[2]We use insurance claim data as the proxy for accidents.
[3]Edlin and Karaca-Mandic (2006) propose the use of the square of traffic density as the proxy for "the likelihood that two other vehicles are in the same location at the same time." Other papers in the literature (e.g., Belmont, 1953; Lundy, 1965; Turner and Thomas 1986) consider the impact of traffic volume on the accident rate.
[4]We would like to thank an anonymous referee for pointing this out.

1,000 per month. Moreover, we find that the monthly claim cost of $545 occurs when both types of externality are considered.

Second, we find that the accident probability is heterogeneous with respect to vehicle characteristics for both types of accident externality. When other people drive more kilometers, the probability of an accident is larger for the policyholders who drive old cars as opposed to those who drive new cars, small cars as opposed to those who drive large cars, and cars registered in cities as opposed to those who drive cars registered in suburbs. When there are more people violating the speed limit, old cars have a significantly higher accident probability than new cars. On the other hand, we find that policyholders who are less than 20 years old have a higher accident severity than the middle-aged individuals when other people drive more, whereas new cars have a higher accident severity than old cars when there are more people driving above a reasonable speed.

The remainder of this article is organized as follows. The data are described in the second section, the empirical methodology is introduced in the third section, and the empirical results are reported in the fourth section. The fifth section concludes the article.

DATA

Our data are obtained from two sources. One is a large vehicle manufacturer whose market share is about 38 percent in Taiwan for the year 2009. The service and maintenance centers of the manufacturer record the kilometers assumed for each customer who visits the centers for vehicle repairs or maintenance. Thus, the data allow us to calculate the kilometers driven by the individual and the average kilometers driven per vehicle.

The other source is an insurance company whose written premium accounted for about 20 percent of the automobile insurance market in 2009. The insurance data contain claim records, including the claim number and claim amount, so that we can use the claim data as a proxy for accident information. In addition, this part of the data contains the insurance policy and the variables used in underwriting, such as the individual's age, gender, marital status, and the insured vehicle's age, brand, and registered area. With the help of these characteristics, we can control the heterogeneity of observations and further investigate the heterogeneity of the accident externality impacts.

We incorporate the two data sources together and obtain our final sample. The individuals we investigate are those who purchase insurance from our sample insurance company,[5] and also have their vehicles maintained or repaired by our sample manufacturer from the year 2002 to the year 2007.[6]

To compute the number of claims and the claim size, we further consolidate the claim data from different insurance policies for each vehicle where the claims might arise

[5]We delete vehicles other than private usage sedan and small truck, which accounts 47.25 percent of the original sample.

[6]When we incorporate the information of mileage usage with the insurance data, we could only study one brand of vehicle that is from our target automobile manufacturer. Hence, we further restrain our research to 38 percent of the insurance data.

due to the same car accident. For example, if a car hit another car and nobody was injured, the policyholder might file claims for first-party car damage insurance and third-party property liability when she is at fault. These two sets of damages are actually caused by one accident. Thus, we merge the claims for each vehicle that are filed on the same date among different insurance policies for the same vehicle as one claim, and sum up the claim amounts to determine the accident size.[7]

We only care about the insurance policies that cover the damage from a car accident. They include compulsory automobile liability insurance, voluntary third-party bodily injury liability insurance, voluntary third-party property liability insurance, and first-party comprehensive coverage insurance. Since this study focuses on the externality caused by driving, in the first-party comprehensive coverage insurance, our sample only includes the vehicle-to-vehicle collision losses. It is worth noting that these kinds of insurance only cover the at-fault accidents. Thus, we can examine whether other people's driving behavior can significantly affect the policyholder's at-fault accident rate and claim amount.

From year 2002 to year 2007, we have observations of 72 calendar months. After the above sample extracted, we obtain 3,796,239 observations. We further delete those observations that do not have at least two maintenance or repair records during our sample period as well as those observations where the vehicles are not insured for the whole calendar month during our research period. There are 631,140 observations deleted. In total, 3,165,099 observations remain in our final monthly data. The samples are constructed as unbalanced panels. On average, we have about 43,960 vehicles per month with a minimum number of 16,290 and a maximum number of 67,662.

Table 1 shows the definitions for all variables used in this study. The basic statistics for each variable are provided in Table 2. Note that the frequency of claims is about 1.59 percent per month. The accident rate is stable over time with an average 1 basis point decrease per month. The average probability of having an accident in each year is about 17.50 percent.[8] It is much higher than for corresponding data from the United States or other developed countries. The high accident probability is mainly contributed by the first-party comprehensive coverage insurance, which accounts for 44.67 percent of the insurance in each year during our sample period. The high probability is due to the fact that Taiwan is a country with a limited territory but a high population density and large numbers of vehicles. The total land area in Taiwan is about 35,980 km^2, which means that it is slightly smaller than the Netherlands and slightly larger than Belgium. Most of the population (about 23 million) are located in the plains, which cover about 27 percent of the territory. In addition, there were a total of about 6.77 million vehicles excluding motorcycles in Taiwan in the year 2009, which had grown at an average annual growth rate of 2.4 percent over the period

[7]Note that an accident can lead to more than one claim, but not all claims will necessarily be reported to the same insurer. Since our data come from one specific insurer, the claim amount in this study might be underestimated. For the accident probability, the underestimation problem is less severe since we use a dummy variable to indicate whether or not the insured vehicle has at least one claim in a month.

[8]The probability of no claim in 12 months is $(1 - 1.59\%)^{12}$. Thus, the average probability of having an accident for a year is equal to $1 - (1 - 1.59\%)^{12} = 17.5\%$.

TABLE 1
Definitions of Variables

Variables	Definition
Dependent variables	
claim	A dummy variable that equals 1 when the insured has filed at least one claim in the current month; otherwise it equals 0
claim amount	The total claim amount (in thousands of U.S. dollars) for the insured who has filed the claim in the current month
Independent variables externalities	
$\overline{km}$	The average hundreds of kilometers driven per month per vehicle that are driven in the whole area of Taiwan
speeding	The total number of speeding tickets (in thousands) per month
The insured's characteristics	
km	The estimated hundreds of kilometers driven for each vehicle in the current month
age2025	A dummy variable that equals 1 if the insured is between the ages of 20 and 25; otherwise it equals 0[a]
age2530	A dummy variable that equals 1 if the insured is between the ages of 25 and 30; otherwise it equals 0[a]
age3060	A dummy variable that equals 1 if the insured is between the ages of 30 and 60; otherwise it equals 0[a]
age60up	A dummy variable that equals 1 if the insured is over the age of 60; otherwise it equals 0[a]
female	A dummy variable that equals 1 if the insured is female; otherwise it equals 0
married	A dummy variable that equals 1 if the insured is married; otherwise it equals 0
carage1	A dummy variable that equals 1 when the car is 1 year old; otherwise it equals 0[b]
carage2	A dummy variable that equals 1 when the car is 2 years old; otherwise it equals 0[b]
carage3	A dummy variable that equals 1 when the car is 3 years old; otherwise it equals 0[b]
carage4	A dummy variable that equals 1 when the car is 4 years old; otherwise it equals 0[b]
capacity2	A dummy variable that equals 1 when the insured car equals or is over 1800 cm^3 and equals or is under 2000 cm^3; otherwise it equals 0[c]
capacity3	A dummy variable that equals 1 when the insured car is over 2000 cm^3; otherwise it equals 0[c]
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
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[d]
south	A dummy variable that equals 1 when the car is registered in the south of Taiwan; otherwise it equals 0[d]
middle	A dummy variable that equals 1 when the car is registered in the middle of Taiwan; otherwise it equals 0[d]

(*Continued*)

TABLE 1
(Continued)

Variables	Definition
Control variables	
season1	A dummy variable that equals 1 when the observation is obtained from January to March[e]
season2	A dummy variable that equals 1 when the observation is obtained from April to June[e]
season3	A dummy variable that equals 1 when the observation is obtained from July to September[e]

[a]The reference group for the dummy variables related to age includes the insured who are under 20 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.
[c]The reference group for the dummy variable related to the vehicle engine capacity is the vehicles that are less than 1800 cm^3.
[d]The reference group for the three dummy variables related to area includes the cars registered in east Taiwan.
[e]The reference group for the three dummy variables related to season includes the data from October to December.

from 1999 to 2009. The average amount of each claim was $1,085.6, with the maximum amount for one claim reaching $354,222.

The key explanatory variables in this article are the average number of kilometers driven per month per vehicle, $\overline{km}$, and the number of speeding tickets per month, $speeding$. These two variables represent different types of accident externality. The former measures how much other people drive on average, and the latter measures how many people drive above the reasonable speed.

To compute $\overline{km}$, we first estimate the kilometers driven per month for each vehicle, km, and then take the average. For each vehicle, we calculate the average kilometers driven per day between two maintenance/repair records, and then times it by the number of days in a month to obtain km. Table 2 shows that the average number of km in our sample is about 1,933.04, whereas the average number of $\overline{km}$ is about 1,918.61. The reason why these two numbers are different is that we calculate $\overline{km}$ before we delete the data with incomplete insurance information, so that the average number of kilometers driven per month per vehicle will be closer to that of the population.

The variable $speeding$ is obtained from the public statistics on the Web site of the National Police Agency, Ministry of Interior.[9] The variable represents the total number of vehicles that violate the speed limits on different types of roadways, including highways and country roads. The average number of speeding tickets issued per month is 212,477; that is, about 3 percent of the vehicles will receive one speeding ticket per month. Note that although this variable could point out how many vehicles

[9]The variable is obtained from the Web site: http://www.npa.gov.tw/NPAGip/wSite/lp?ctNode=12593&CtUnit=2374&BaseDSD=7&mp=1.

TABLE 2
The Basic Statistics of the Variables

Variables	Mean	STD	Min	Max
Dependent variables				
claim	0.0159	0.1251	0.0000	1.0000
claim amount (in $1000)	1.0856	4.3989	0.0059	354.2222
Independent variables externalities				
$\overline{km}$*(in 100 kilometers)*	19.1861	2.0791	13.1922	21.9172
speeding (in thousands)	212.477	30.770	139.575	312.071
The insured's characteristics				
km(in 100 kilometers)	19.3304	23.3996	0.0000	143.5442
age2025	0.0085	0.0919	0.0000	1.0000
age2530	0.0633	0.2435	0.0000	1.0000
age3060	0.8867	0.3170	0.0000	1.0000
age60up	0.0356	0.1852	0.0000	1.0000
female	0.7019	0.4574	0.0000	1.0000
married	0.9203	0.2708	0.0000	1.0000
carage1	0.4894	0.4999	0.0000	1.0000
carage2	0.1989	0.3992	0.0000	1.0000
carage3	0.1331	0.3397	0.0000	1.0000
carage4	0.0866	0.2813	0.0000	1.0000
capacity2	0.2800	0.4490	0.0000	1.0000
capacity3	0.0539	0.2258	0.0000	1.0000
sedan	0.9777	0.1477	0.0000	1.0000
city	0.4947	0.5000	0.0000	1.0000
north	0.4751	0.4994	0.0000	1.0000
south	0.3054	0.4606	0.0000	1.0000
middle	0.1996	0.3997	0.0000	1.0000
Control variables				
season1	0.2318	0.4220	0.0000	1.0000
season2	0.2440	0.4295	0.0000	1.0000
season3	0.2570	0.4370	0.0000	1.0000
Number of observations		3,165,099		

exceed the reasonable speed limits in different area, it cannot identify how fast they travel.

The insured are mostly between 30 and 60 years old. Most of the policyholders are female and married. The percentage of female policyholders is over 70 percent, and the percentage of married policyholders is over 90 percent. The major reason why most of the policyholders are married middle-aged females is that there are insurance premium discounts for them in Taiwan, and the insurance policies cover all the drivers of the insured vehicle. Thus, most families will register their vehicles and insure the vehicles under the name of a married middle-aged female member of the family.

It should be noted that our final sample is limited to a specific car brand since the vehicle manufacturer only looks after its own brand of vehicles. Another drawback is that our observations are insured by a specific insurance company. Thus, our data might have a certain bias.[10]

METHODOLOGY

Our research is based on a set of unbalanced panel data. Probit regressions are adopted to test the effect of the accident externality on the loss probability. We employ random effect model to correct for heteroskedasticity.[11] The random-effects probit model is expressed as:

$$\begin{aligned}\text{Prob}(claim_{it} = 1) = F(&\beta_{\overline{km}}\overline{km}_t + \beta_{sp} speeding_t + X_{it}\beta_x \\ &+ \beta_{q1} season1_t + \beta_{q2} season2_t + \beta_{q3} season3_t),\end{aligned} \tag{1}$$

where F is the cumulative standard normal distribution function. The βs are the corresponding coefficients. If $\beta_{\overline{km}}$ and β_{sp} are significantly positive, it is evidence for the existence of the accident externality: the individual's accident probability will increase when other people drive more or more people drive above a reasonable speed.[12] X_{it} is the vector of explanatory variables for the insured's information, and includes km_{it}, the characteristics of the policyholders and the characteristics of the insured vehicle as defined in Table 1. Since the seasonal effect could be important for monthly data, the seasonal dummy variables, $season1$, $season2$, and $season3$, are also included as control variables.

To examine the effect of the externality on the loss severity, we adopt the random-effects ordinary least squares (OLS) regression with the following form:[13]

$$\begin{aligned}claim\ amount_{it} = \gamma_i + &\gamma_{\overline{km}}\overline{km}_t + \gamma_{sp} speeding_t + X_{it}\gamma_x \\ &+ \beta_{q1} season1_t + \beta_{q2} season2_t + \beta_{q3} season3_t + \varepsilon_{it}.\end{aligned} \tag{2}$$

The dependent variable is the amount of the claim for individual i in month t. γ_i denotes the unobserved heterogeneity term, which is assumed to be individual specific and time invariant. ε_{it} denotes the error term. In the OLS regression, only the observations that have a claim are included. In other words, we estimate the accident externality that is conditional upon the loss events. The size of the accident externality is calculated according to the coefficients $\gamma_{\overline{km}}$ and γ_{sp}. Note that $\gamma_{\overline{km}}$ and γ_{sp} could be positive if an increase in $\overline{km}_t$ increases the chance of an accident involving more cars,

[10]We compare our data with the data from our sample insurance company that contains different car brands and find that our final set of data has significantly higher percentages of those in the 30- to 60-year-old age group, females, new cars, small cars with a capacity of less than 2,000 cm^3, and sedans.

[11]We further examined the effect of the accident externality by adopting the pooled cross-sectional probit model. The results are similar to those from the random effects model and thus are unreported.

[12]Vickrey (1968) predicts that $\beta_{\overline{km}}$ should be positive.

[13]We performed the Hausman test. The chi-square statistics are statistically insignificantly even at 10 percent level. The outcomes of these Hausman test are list in Table 4.

or if an increase in $speeding_t$ indicates that there are more speeding vehicles, and then increases the claim size.

To examine whether the accident externality has a heterogeneous effect on individuals with different characteristics, we use the interaction terms: $X_{it} \times \overline{km}_t$ and $X_{it} \times speeding_t$. In other words, Equations (1) and (2) are modified as

$$\begin{aligned} \text{Prob}(claim_{it} = 1) = F\,(&\theta_{\overline{km}}\overline{km}_t + \theta_{sp} speeding_t + X_{it}\theta_x \\ &+ \theta_{q1} season1_t + \theta_{q2} season2_t + \theta_{q3} season3_t \\ &+ X_{it} \times \overline{km}_t \times \theta_1 + X_{it} \times speeding_t \times \theta_2), \end{aligned} \tag{3}$$

and

$$\begin{aligned} claim\ amount_{it} = \phi_i &+ \phi_{\overline{km}}\overline{km}_t + \phi_{sp} speeding_t + X_{it}\phi_x \\ &+ \phi_{q1} season1_t + \phi_{q2} season2_t + \phi_{q3} season3_t \\ &+ X_{it} \times \overline{km}_t \times \phi_1 + X_{it} \times speeding_t \times \phi_2 + \eta_{it}, \end{aligned} \tag{4}$$

where the θs and ϕs are the coefficients and η_{it} denotes the error term. In Equations (3) and (4), we concentrate on the coefficients θ_1, θ_2, ϕ_1, and ϕ_2. These two equations are used to explore whether or not the accident externality could be decomposed. For example, if the corresponding coefficient of $km_{it} \times speeding_t$ in Equation (3) is significantly positive, it could be explained as meaning that the individual who drives more will have a higher probability of facing an accident than those who drive less when more people drive above the reasonable speed. In other words, θ_1, θ_2, ϕ_1, and ϕ_2 could indicate the characteristics of the victims and represent the heterogeneous effect of the accident externality.

EMPIRICAL RESULTS

The Accident Externality's Impact on Loss Frequency

Table 3 shows the impact of $\overline{km}$ and $speeding$ on the individual's loss frequency. Model 1 of Table 3 presents the results of Equation (1). Models 2 and 3 provide the results of Equation (1) but only, respectively, include $\overline{km}$ and $speeding$ rather than both of them in the robustness analysis. All models demonstrate that the coefficients of the average number of kilometers driven per month per vehicle and the number of speeding tickets are positively significant at the 5 percent level. These positive coefficients confirm that the accident externality exists in terms of loss frequency.

Specifically, Model 1 shows that the individual's accident probability evaluated at the mean level for all covariates will increase by 0.2937 percent on the basis of the coefficient of $\overline{km}$, 0.0765, when the average number of kilometers driven per month per vehicle increases by 100 km. When the number of speeding tickets per month increases by 1,000 tickets per month, it will increase the individual's accident probability by 0.0081 percent on the basis of the coefficient of $speeding$, 0.0021. When $\overline{km}$ and $speeding$ increase by one standard deviation from their mean value, the accident probabilities will increase by 0.6106 percent and 0.2492 percent,

TABLE 3
The Effect of Average Driving on Loss Probability

	Model 1		Model 2		Model 3	
Variables	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Intercept	−3.1766	<0.0001	−3.1535	<0.0001	−3.0596	<0.0001
Externalities						
$\overline{km}$	0.0765	0.0060	0.0888	0.0010		
speeding	0.0021	0.0140			0.0033	0.0032
The insured's characteristics						
km	0.0031	<0.0001	0.0031	<0.0001	0.0031	<0.0001
age2025	−0.0232	0.8330	−0.0232	0.8340	−0.0232	0.8330
age2530	0.0573	0.5970	0.0574	0.5970	0.0574	0.5970
age3060	0.0819	0.4500	0.0819	0.4500	0.0819	0.4500
age60up	0.0542	0.6190	0.0542	0.6190	0.0540	0.6200
female	0.1325	<0.0001	0.1324	<0.0001	0.1324	<0.0001
married	−0.0126	0.0780	−0.0125	0.0810	−0.0124	0.0820
carage1	0.6055	<0.0001	0.6058	<0.0001	0.6068	<0.0001
carage2	0.2334	<0.0001	0.2336	<0.0001	0.2345	<0.0001
carage3	0.1139	<0.0001	0.1138	<0.0001	0.1149	<0.0001
carage4	0.0512	<0.0001	0.0512	<0.0001	0.0520	<0.0001
capacity2	−0.0345	<0.0001	−0.0346	<0.0001	−0.0347	<0.0001
capacity3	−0.0039	0.6530	−0.0038	0.6650	−0.0039	0.6610
sedan	0.2166	<0.0001	0.2167	<0.0001	0.2175	<0.0001
city	0.0057	0.0260	0.0057	0.0280	0.0057	0.1260
north	−0.0372	0.0080	−0.0372	<0.0001	−0.0370	0.0080
south	0.0771	0.0000	0.0771	<0.0001	0.0773	<0.0001
middle	0.0376	0.0080	0.0376	<0.0001	0.0378	0.0080
Control variables						
season1	−0.0928	<0.0001	−0.0901	<0.0001	−0.0867	<0.0001
season2	−0.0419	0.0020	−0.0389	0.0040	−0.0392	0.0060
season3	−0.0375	0.0050	−0.0395	0.0040	−0.0353	0.0120
Log likelihood	−247046		−247047		−247050	
Number of observations	3,165,099		3,165,099		3,165,099	

respectively.[14] When these increments are compared to the average accident probability of 1.59 percent, the impact of the externality can not be neglected.

In addition, all models reveal that the number of kilometers driven by the individual is also significantly positively correlated with the loss probability. When the number of kilometers driven by the policyholder increases by one standard deviation above its mean, the individual's accident probability will increase by 0.2747 percent on the basis of the coefficients of km in Model 1. Models 1 and 2 show that the coefficient of km is smaller than the coefficient of $\overline{km}$. It means that the loss probability increase due to km is lower than that due to the same increment in $\overline{km}$. This is because $\overline{km}$

[14]The standard deviation of $\overline{km}$ is 208 kilometers and the standard deviation of $speeding$ is 30,770 tickets per month.

represents the "average kilometers" among all vehicles. A 1-km increase in $\overline{km}$ could be viewed as everyone in the society driving one more kilometer. The traffic density will have a significant increase compared to the case where the policyholder alone decides to drive one more kilometer.

The results for the other control variables in all models are quite consistent and are generally consistent with the findings in the literature. Specifically, we find that a single female has a higher probability of a car accident. New cars have a higher accident probability than old cars. These findings are generally consistent with the findings of Wang, Chung, and Tzeng (2008). If the vehicle is registered in a city, then the probability of an accident is also higher. This result is consistent with Belmont (1953) and Lundy (1965) who find that accident rates increase with traffic volume. We also find that smaller sedans have higher accident rates. For the seasonal effect, we find that the accident rate is higher during winter.

The Accident Externality's Impact on Loss Severity

Table 4 reports the effects of the OLS regression on the loss severity that is conditional on a loss being obtained. Similar to Table 3, Model 1 of Table 4 includes both $\overline{km}$ and *speeding*, whereas Models 2 and 3 only, respectively, include $\overline{km}$ and *speeding*. All models show that the coefficients of $\overline{km}$ and *speeding* are positive and significant at the 5 percent level. This finding supports the view that both types of accident externalities exist in terms of loss severity. In Table 4, Model 1 shows that a one-hundred kilometer increase in $\overline{km}$ at its mean level will increase the policyholder's accident cost by $8.2, whereas a 1,000-ticket increase in *speeding* at its mean level will increase the claim amount by $1.8. Since the average of $\overline{km}$ is 1,918.61 kilometers per vehicle per month and the average of *speeding* is 212,477 per month, for each vehicle a cost of $162.5[15] per month will be incurred due to other people's driving kilometers and a cost of $382.5[16] per month will be incurred due to other people's speeding behavior. In total, a monthly externality cost of $545 occurs when both types of externality are considered.

All models in Table 4 demonstrate that the effects of other explanatory variables on loss severity are similar regardless of which type of accident externality is considered. We find that the policyholder's own number of kilometers driven has an insignificantly negative effect on loss severity. This finding is contrary to the findings in relation to loss probability. Intuitively, the more a person drives the higher is the risk exposure and so the higher is the amount of the damage. On the other hand, it is also true that driving experience plays an important role and so acquired driving expertise could be a reason to explain why kilometers driven would not influence loss frequency.[17]

We further find that females or vehicles registered in cities have lower accident costs. We suspect that these phenomena might be correlated with driving speed. On average, the driving speed of female drivers might be lower than that of male drivers. People

[15]The number is obtained by $8.2 per hundred km per month $\times$19.1861 hundred km.
[16]The number is obtained by $1.8 per thousand speeding tickets per month $\times$212.477 thousand speeding tickets.
[17]We would like to thank an anonymous referee for pointing out the possible reason.

TABLE 4
The Effect of Average Driving on Loss Severity

	Model 1		Model 2		Model 3	
Variables	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Intercept	3.3176	0.0070	3.0977	0.0110	3.4259	0.0050
Externalities						
$\overline{km}$	0.0082	<0.0001	0.0094	<0.0001		
speeding	0.0018	0.0120			0.0019	0.0100
The insured's characteristics						
km	−0.0024	0.4970	−0.0021	0.5530	−0.0024	0.5012
age2025	−0.0314	0.9790	−0.0315	0.9790	−0.0305	0.9800
age2530	−0.9226	0.4330	−0.9229	0.4330	−0.9219	0.4340
age3060	−1.0097	0.3910	−1.0087	0.3910	−1.0091	0.3910
age60up	−1.0202	0.3880	−1.0177	0.3890	−1.0208	0.3880
female	−0.1990	<0.0001	−0.1984	<0.0001	−0.1993	<0.0001
married	0.0370	0.6340	0.0329	0.6720	0.0376	0.6280
carage1	−0.4297	0.0010	−0.4441	<0.0001	−0.4212	<0.0001
carage2	−0.2031	0.1290	−0.2122	0.1120	−0.1961	0.1410
carage3	−0.1960	0.1800	−0.1938	0.1850	−0.1886	0.1950
carage4	−0.2194	0.1830	−0.2203	0.1810	−0.2154	0.1900
capacity2	0.0660	0.1420	0.0732	0.1030	0.0644	0.1510
capacity3	0.0100	0.9180	0.0037	0.9700	0.0101	0.9180
sedan	−0.2546	0.1720	−0.2592	0.1650	−0.2475	0.1840
city	−0.0991	0.0140	−0.1008	0.0120	−0.0993	0.0140
north	−0.1785	0.2440	−0.1758	0.2520	−0.1774	0.2470
south	−0.2394	0.1170	−0.2360	0.1230	−0.2377	0.1200
middle	−0.0746	0.6320	−0.0737	0.6360	−0.0733	0.6380
Control variables						
season1	−0.0109	0.8580	−0.0362	0.5490	−0.0048	0.9360
season2	0.0827	0.1580	0.0569	0.3230	0.0850	0.1430
season3	0.0312	0.5900	0.0515	0.3670	0.0337	0.5560
Hausman test	16.28	0.8430	14.96	0.8639	15.01	0.8618
Overall R^2	0.0240		0.0220		0.0230	
Number of observations	50,298		50,298		50,298	

could also drive slowly in high-population density areas. Thus, the lower accident severity might be caused by the lower driving speed. In addition, we find that the vehicles that are 1 year old or less have significantly lower claim amounts at the 1 percent level than the vehicles that are more than 4 years old.

The Heterogeneity Effect of the Accident Externality

The results for Equation (3) are shown in Table 5. Due to the fact that the number of observations is huge and the dependent variables included in Equation (3) are many, our computer's capacity does not allow us to directly estimate Equation (3). To overcome this problem, we construct a subsample by randomly selecting half of the observations in each month. In the subsample, we have 1,582,479 observations.

TABLE 5
The Heterogeneity Effect of Average Driving on Loss Probability

	Main Effects		Interaction Effects			
			$X \times \overline{km}$		$X \times speeding$	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Intercept	−3.5789	0.0320				
Externalities						
$\overline{km}$	0.0081	0.7520				
speed	0.0052	0.3930				
The insured's characteristics						
km	0.0088	<0.0001	0.0032	0.0010	$2.99E-06$	0.3180
age2025	−0.2377	0.8870	0.0365	0.6820	−0.0021	0.7290
age2530	−0.2120	0.8980	0.0432	0.6220	−0.0024	0.6930
age3060	−0.3236	0.8440	0.0508	0.5620	−0.0024	0.6880
age60up	−0.4239	0.7980	0.0443	0.6140	−0.0014	0.8090
female	0.0849	0.1990	0.0014	0.6920	0.0001	0.5840
married	0.0710	0.5040	−0.0042	0.4400	−0.0001	0.8800
carage1	0.8304	<0.0001	−0.0024	0.0900	−0.0007	0.0360
carage2	0.5296	<0.0001	−0.0002	0.9810	−0.0009	0.0300
carage3	0.2553	0.1250	0.0028	0.7080	−0.0001	0.2060
carage4	0.4091	0.0260	−0.0109	0.1860	−0.0001	0.3890
capacity2	0.0412	0.5090	−0.0069	0.0310	−0.0003	0.2270
capacity3	0.3508	0.0080	−0.0122	0.0040	−0.0005	0.2000
sedan	0.3938	0.0400	0.0005	0.9640	−0.0001	0.2040
city	0.0794	0.1640	0.0056	0.0590	0.0001	0.6390
north	0.1701	0.3950	−0.0070	0.4980	−0.0001	0.6660
south	0.1739	0.3850	0.0043	0.6770	−0.0002	0.2110
middle	0.0886	0.6640	0.0021	0.8410	−0.0004	0.5560
Control variables						
season1	−0.0874	<0.0001				
season2	−0.0446	0.0030				
season3	−0.0322	0.0280				
Log likelihood			−123311			
Number of observations			1,582,479			

The basic statistics of the subsample, which are shown in the Appendix, Table A1, are similar to the basic statistics for the full sample.

The first column of Table 5 demonstrates the estimated values of $\theta_{\overline{km}}$, θ_{sp}, θ_x, θ_{q1}, θ_{q2}, and θ_{q3}. The third and the fifth columns, respectively, show the estimated values of θ_1 and θ_2. The corresponding p-values are shown in the even columns. We find that the coefficients of $\overline{km}$ and $speeding$ become insignificant when the interaction terms are considered. These findings indicate that the effect of the accident externality can be decomposed according to the insured characteristics. In other words, the accident externality is heterogeneous.

TABLE 6
The Marginal Probability of the Externality for Individuals With Different Characteristics

	Probability (1)	Counter Group (2)	Difference (1)–(2)	p-value
$\overline{km} \times carage1$[a]	0.0027364	0.0034001	−0.0006637	0.0010
$\overline{km} \times capacity2$[b]	0.0032559	0.0034701	−0.0002142	0.0080
$\overline{km} \times capacity3$[c]	0.0028880	0.0033699	−0.0004819	0.0060
$\overline{km} \times city$[d]	0.0034604	0.0033598	0.0001006	0.0760
$speeding \times carage1$[a]	0.0001002	0.0001316	−0.0000314	0.0450
$speeding \times carage2$[e]	0.0001085	0.0001319	−0.0000234	0.0630

[a]The counter group includes the cars that are more than 1 year old.
[b]The counter group comprises the vehicles that have an engine capacity of less than 1,800 cm^3 or more than 2,000 cm^3.
[c]The counter group consists of the vehicles that have an engine capacity of less than 2,000 cm^3.
[d]The counter group consists of the vehicles that are registered in suburbs.
[e]The counter group includes the cars that are not yet 2 years old.

We find that the coefficient of the interaction term $\overline{km} \times km$ is significantly positive at the 1 percent level. The individual who drives 100 km more will increase the accident probability by an additional 0.0123 percent when the average number of kilometers driven per month per vehicle increases by 100 km. In addition, the insured vehicles which are 1 year old, with a capacity above 1,800 cm^3, and registered in the suburbs face a lower accident externality in terms of loss probability compared to the vehicles that are 4 years old, with a capacity less than 1,800 cm^3, and registered in the cities, respectively, when the average number of kilometers driven per month per vehicle increases. Table 5 further shows that the accident probability for cars 1 and 2 years old significantly decreases more than that for vehicles that are more than 4 years old when the number of speeding tickets increases.

Table 6 presents the marginal probability of the externality estimated by the coefficients obtained from Table 5. When $\overline{km}$ increases by 100 km, the first row in Table 6 shows that vehicles that are more than 1 year old face an additional 0.066 percent accident probability compared to 1-year-old (or less) vehicles. The accident probability of vehicles with a capacity of less than 2,000 cm^3 increases by 0.048 percent more than that of vehicles with a capacity of more than 2,000 cm^3 when $\overline{km}$ increases by 100 km. If $\overline{km}$ increases by 100 km, vehicles registered in cities suffer an additional 0.01 percent accident probability compared to those registered in suburban areas. Furthermore, Table 6 demonstrates that vehicles that are more than 1 year old face an additional 0.003 percent accident probability compared to 1-year-old (or less) vehicles when the average number of speeding tickets increases by 1,000.

Both Tables 5 and 6 lead us to conclude that the accident externality in terms of probability is heterogeneous with respect to the vehicle's characteristics. We find that the drivers of older cars, smaller cars, and the vehicles registered in cities that have higher numbers of kilometers driven are more likely to be the victims due to the

TABLE 7
The Heterogeneity Effect of Average Driving on Loss Severity

	Main Effects		Interaction Effects			
			$X \times \overline{km}$		$X \times speeding$	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Intercept	3.4509	0.8910				
Externalities						
$\overline{km}$	0.5011	0.6860				
speed	0.0432	0.3760				
The insured's characteristics						
km	0.0049	0.7120	−0.0005	0.4630	1.3E-05	0.5490
age2025	−4.8859	0.8470	0.7597	0.1250	−0.0663	0.1710
age2530	−2.5701	0.9190	−0.7541	0.1410	−0.0553	0.2480
age3060	−3.6614	0.8840	−0.9853	0.0980	−0.0551	0.2490
age60up	−3.3375	0.8950	0.7406	0.1490	−0.0555	0.2480
female	0.3574	0.4970	−0.0213	0.4330	−0.0006	0.6950
married	0.5499	0.5120	−0.0216	0.6120	−0.0004	0.8830
carage1	−0.4333	0.7110	0.0062	0.2440	0.1694	0.0900
carage2	−0.6159	0.6290	0.0084	0.1370	0.1193	0.1040
carage3	−1.4649	0.3110	0.0092	0.1450	0.0725	0.2220
carage4	0.6478	0.6870	0.0065	0.3880	0.0363	0.5820
capacity2	−0.2338	0.6270	0.0250	0.3050	−0.0009	0.5790
capacity3	−0.8139	0.4450	0.0072	0.8980	0.0031	0.3280
sedan	0.7189	0.6520	−0.0591	0.4930	0.0006	0.9280
city	−0.6759	0.1270	0.0313	0.1710	−0.0002	0.9090
north	0.5971	0.6990	−0.0874	0.2760	0.0042	0.4090
south	0.9622	0.5330	−0.1098	0.1710	0.0042	0.4030
middle	1.3990	0.3730	−0.1097	0.1800	0.0030	0.5660
Control variables						
season1	−0.0101	0.8800				
season2	0.0765	0.2370				
season3	0.0303	0.6350				
Overall R^2			0.0294			
Number of observations			50,298			

kilometers driven by others, whereas only the drivers of older cars are the victims of others' speeding behavior.

For the robustness checks, we also use the full sample to estimate Equation (3) but only one type of externality is considered each time when we run the regressions. When the variables $X \times \overline{km}$ are included but not $X \times speeding$, the results generated from the full sample exhibit the same conclusions as those from the subsample; that is, the accidents are more likely to involve older cars, smaller cars and vehicles registered in cities and cars with more kilometers driven. When $X \times speeding$ are included but not $X \times \overline{km}$, the full sample predicts that the vehicles that are less than 2 years old and have an engine capacity of less than 2,000 cm^3 have a lower accident externality in terms of the loss probability compared to the vehicles which are more than 4 years old and an engine capacity of less than 1,800 cm^3, respectively.

As regards the effect of the accident externality in terms of severity, we find that the externality could be decomposed according to either the individual's or the vehicle's age as shown in Table 7. Table 7 is estimated using all the observations that have at least one claim in a month. The sample is the same as in Table 4. The first column in Table 7 shows the estimated value of $\phi_{\overline{km}}$, ϕ_{sp}, ϕ_x, ϕ_{q1}, ϕ_{q2}, and ϕ_{q3} in Equation (4), the third column is the estimated value of ϕ_1, and the fifth column is the estimated value of ϕ_2. Table 7 indicates that when the average number of kilometers driven per vehicle increases by 100 km, the loss severity for individuals who are less than 20 years old increases by \$985 per month more that for than policyholders who are between 30 and 60 years old. When the number of speeding tickets in the society increases by 1,000, cars that are 1 year old or less will face fines of \$169 per month more than the cars that are more than 4 years old.[18]

CONCLUSIONS

The literature has indicated that the accident externality from driving is significant in different countries. However, most articles estimate the cost of the externality by using aggregate data, and the details regarding the externality are not clear. Our article examines the accident externality from driving by using data at the individual level in Taiwan.

We examined two types of externality: one is the average kilometers driven per month and the other is the numbers of speeding tickets issued in the society. We found that the impacts of the accident externality in each month on the loss probability are about 0.6106 percent and 0.2492 percent when the average number of kilometers driven per vehicle per month and the total number of speeding tickets increases by one standard deviation, respectively. Being conditional upon an accident, a monthly externality cost of \$545 on the claim amount occurs when both types of externality are considered.

In addition, the externality is heterogeneous in terms of the insured's characteristics. We found that older cars, smaller cars, and the vehicles registered in cities with higher numbers of kilometers driven suffer higher accident probabilities due to the higher numbers of kilometers driven by others, whereas only older cars are the victims in terms of accident probability due to others' speeding behavior. When the average number of kilometers driven per vehicle increases by 100 km, the loss severity for individuals who are less than 20 years old increases more than that for policyholders who are between 30 and 60 years old. When the number of speeding tickets in the society increases, the claim amount for cars that are 1 year old or less will increase more than that of the cars that are over 4 years old.

Our findings may generate further policy implications. The literature has proposed that a Pigouvian tax on gasoline, automobile insurance based on miles driven, and

[18]In Table A2, we report the estimated results for Equation (4) by using the claim data in the subsample. In addition to the coefficients of $age3060 \times \overline{km}$ and $carage1 \times speeding$, the coefficient of $carage2 \times speeding$ is significantly positive.

pay-as-you-drive insurance could reduce the accident externality from driving.[19] Indeed, the above policies could reduce both km and $\overline{km}$ and further decrease the externality. In addition to these policies, our findings suggest that alternative mechanisms exist to correct the accident externality from driving.

To handle the heterogeneous effect regarding the insured's characteristics, a government could launch an insurance premium tax due to the accident externality. Rather than apply a constant tax rate as with most premium taxes, the proposed tax rate could be designed to be negatively correlated with the characteristics that disclose whether the insured is suffering more as a result of the accident externality. To avoid a significant increase in the number of uninsured after the insurance premium tax, we suggest that this type of tax be levied on compulsory insurance rather than on voluntary insurance.

Another possible way of correcting the accident externality from driving is that the government can design a bonus-malus system for compulsory insurance that is based on who is affected by the accident externality. The current bonus-malus system on compulsory automobile insurance in Taiwan considers three factors that are used to adjust the premium: the policyholder's gender, age, and past record. Our study suggests that the bonus-malus system could depend on the vehicle's characteristics to ease the accident externality. By using Tunisian data, Dionne and Ghali (2005) show that the bonus-malus system can reduce the probability of reported accidents for good risks but has no effect on bad risks because the bad risks can switch insurance companies. Thus, if the goal of the government is to compensate the individuals who face a greater accident externality in terms of probability, then older cars, smaller cars and the vehicles registered in cities should receive some premium discounts (i.e., bonuses), whereas new cars, cars with an engine capacity above 2,000 cm^3 and the vehicles registered in suburbs should pay an extra premium. Since the proposed bonus-malus system depends on observable vehicle characteristics, switching companies can not help the bad risks to avoid the extra premium. Therefore, the victims of other people's driving can be compensated by this type of system.

To compensate the policyholders who suffer higher claim amounts due to the accident externality, the government could consider launching a relief system that depends on $\overline{km}$, *speeding*, and the insured's characteristics. For example, for the vehicles involved in a car accident, the vehicle that is 1 year old or less and policyholders who are less than 20 years old could receive such relief. The amount of relief for 1-year-old (or less) cars should be positively correlated with the numbers of speeding tickets in the society, whereas that for policyholders who are less than 20 years old should be positively correlated with the traffic density. In this system, the victims of the accident externality in terms of severity are compensated. The budget for this kind of relief could come from the fines for traffic violations.

[19]For example, see Parry (2005), Edlin and Karaca-Mandic (2006), Saito, Kato, and Shimane (2007), and Parry, Walls, and Harrington (2007).

APPENDIX

As mentioned in the Introduction, understanding the accident externality from driving in terms of its probability and severity can help the government correct the externality more accurately. Let us use a simple model, which is modified on the basis of Dupor and Liu (2003) and Huang and Tzeng (2008), to illustrate how governments could use the information from our article to correct the externality.

For simplicity, let us consider only one type of accident externality. Assume that the individuals are homogeneous and endowed with wealth w. The representative individual might face an accident loss L, which is a function of the number of kilometers driven, km, and the average number of kilometers driven by the society, $\overline{km}$, and is denoted by $L(km,\overline{km})$. The loss probability, π, is also a function of km and $\overline{km}$ and is denoted by $\pi(km,\overline{km})$. Assume that driving could bring the individual some benefit, for example, saving time for the individual or having fun with driving. Let $B(km)$ denote the corresponding monetary reward from driving with $B\prime(km) > 0$. Thus, the individual will take $\overline{km}$ as given and choose the optimal number of kilometers driven to maximize her expected utility:

$$Eu = \pi(km,\overline{km})u(w - L(km,\overline{km}) + B(km)) + (1 - \pi(km,\overline{km}))\, u(w + B(km)), \quad \text{(A1)}$$

where u denotes the individual's utility function with $u' > 0$ and $u'' \leq 0$. The corresponding first-order condition is

$$\begin{aligned} M = {} & \frac{\partial \pi(km,\overline{km})}{\partial km} u(w - L(km,\overline{km}) + B(km)) \\ & + \pi(km,\overline{km})u'(w - L(km,\overline{km}) + B(km)) \left(-\frac{\partial L(km,\overline{km})}{\partial km} + B'(km) \right) \\ & - \frac{\partial \pi(km,\overline{km})}{\partial km} u(w + B(km)) \\ & + \left(1 - \pi(km,\overline{km})\right) u'(w + B(km))B'(km) \\ = {} & 0. \end{aligned} \quad \text{(A2)}$$

On the other hand, the government knows that everyone's decision regarding km will also affect the average kilometers driven, $\overline{km}$. Thus, a benevolent social planner will take the externality into consideration. Thus, the first-order condition of the government will be

$$\begin{aligned} N = M + {} & \frac{\partial \pi(km,\overline{km})}{\partial \overline{km}} u(w - L(km,\overline{km}) + B(km)) \\ & + \pi(km,\overline{km})u'(w - L(km,\overline{km}) + B(km)) \left(-\frac{\partial L(km,\overline{km})}{\partial \overline{km}} \right) \\ & - \frac{\partial \pi(km,\overline{km})}{\partial \overline{km}} u(w + B(km)) \\ = {} & 0. \end{aligned} \quad \text{(A3)}$$

The private optimum will reach the social optimum if and only if $N = M = 0$, that is,

$$\begin{aligned}
&\frac{\partial \pi(km,\overline{km})}{\partial \overline{km}} u(w - L(km,\overline{km}) + B(km)) \\
&\qquad + \pi(km,\overline{km}) u'(w - L(km,\overline{km}) + B(km)) \left(-\frac{\partial L(km,\overline{km})}{\partial \overline{km}}\right) \\
&\qquad - \frac{\partial \pi(km,\overline{km})}{\partial \overline{km}} u(w + B(km)) = 0.
\end{aligned} \tag{A4}$$

As argued by Dupor and Liu (2003) and Huang and Tzeng (2008), to correct the externality, the government should design a system such that Equation (A4) will hold. In other words, understanding $\frac{\partial \pi(km,\overline{km})}{\partial \overline{km}}$ and $\frac{\partial L(km,\overline{km})}{\partial \overline{km}}$ can help the government to design a system to reach the social optimum.

It is worth noting that if the representative individual is assumed to be risk neutral, then the expected cost of the accident externality is sufficient to determine the difference between the social and private optima. Specifically, let $u\prime = 1$. Equation (A4) may then be written as

$$\begin{aligned}
&\frac{\partial \pi(km,\overline{km})}{\partial \overline{km}} (w - L(km,\overline{km}) + B(km)) + \pi(km,\overline{km}) \left(-\frac{\partial L(km,\overline{km})}{\partial \overline{km}}\right) \\
&\qquad - \frac{\partial \pi(km,\overline{km})}{\partial \overline{km}} (w + B(km)) \\
&= \frac{\partial \pi(km,\overline{km})}{\partial \overline{km}} (-L(km,\overline{km})) + \pi(km,\overline{km}) \left(-\frac{\partial L(km,\overline{km})}{\partial \overline{km}}\right) \\
&= -\frac{\partial \left[\pi(km,\overline{km}) L(km,\overline{km})\right]}{\partial \overline{km}}.
\end{aligned}$$

In other words, estimating the expected costs associated with the accident externality, $\frac{\partial[\pi(km,\, \overline{km})L(km,\, \overline{km})]}{\partial \overline{km}}$, would help the government to correct the externality when all the individuals are risk neutral.

TABLE A1
The Basic Statistics of the Variables in the Subsample

Variables	Mean	STD	Min	Max
Dependent variables				
claim	0.0159	0.1252	0.0000	1.0000
claim amount (in $1000)	0.9966	3.9775	0.0054	166.1162
Independent variables externalities				
$\overline{km}$*(in 100 kilometers)*	19.1865	2.0788	13.1922	21.9172
speeding (in thousands)	212.477	30.770	139.575	312.071
The insured's characteristics				
km(in 100 kilometers)	19.3114	23.3749	0.0000	143.5442
age2025	0.0084	0.0911	0.0000	1.0000
age2530	0.0631	0.2432	0.0000	1.0000
age3060	0.8869	0.3167	0.0000	1.0000
age60up	0.0412	0.1988	0.0000	1.0000
female	0.7020	0.4574	0.0000	1.0000
married	0.9210	0.2697	0.0000	1.0000
carage1	0.4894	0.4999	0.0000	1.0000
carage2	0.1992	0.3994	0.0000	1.0000
carage3	0.1331	0.3397	0.0000	1.0000
carage4	0.0864	0.2810	0.0000	1.0000
capacity2	0.2801	0.4490	0.0000	1.0000
capacity3	0.0536	0.2253	0.0000	1.0000
sedan	0.9778	0.1474	0.0000	1.0000
city	0.4944	0.5000	0.0000	1.0000
north	0.4750	0.4994	0.0000	1.0000
south	0.3057	0.4607	0.0000	1.0000
middle	0.1993	0.3995	0.0000	1.0000
Control variables				
season1	0.2318	0.4220	0.0000	1.0000
season2	0.2440	0.4295	0.0000	1.0000
season3	0.2570	0.4370	0.0000	1.0000
Number of observations			1,582,479	

TABLE A2
The Heterogeneity Effect of Average Driving on Loss Severity in the Subsample

	Main Effects		Interaction Effects			
			$X \times \overline{km}$		$X \times speeding$	
	Coefficient	*p*-value	Coefficient	*p*-value	Coefficient	*p*-value
Intercept	6.9569	0.8500				
Externalities						
$\overline{km}$	0.0563	0.9810				
speed	0.0205	0.8570				
The insured's characteristics						
km	0.0039	0.8380	−0.0005	0.6070	1.8*E*-05	0.5560
age2025	−4.2276	0.9090	0.0867	0.7910	−0.0299	0.7910
age2530	−1.3567	0.9700	−0.1682	0.2710	−0.0016	0.9890
age3060	−2.9748	0.9350	−0.6408	0.0944	−0.0018	0.9870
age60up	−3.5284	0.9230	0.1429	0.2530	0.0029	0.9800
female	−0.6433	0.3800	0.0151	0.6910	0.0009	0.6940
married	0.4863	0.6830	−0.0039	0.9480	−0.0018	0.6520
carage1	0.0457	0.9770	−0.0234	0.7580	0.2271	0.0420
carage2	0.0917	0.9580	−0.0047	0.9550	0.1962	0.0790
carage3	−1.5473	0.4400	0.0176	0.8510	0.0961	0.1830
carage4	2.3703	0.2860	−0.1139	0.2770	0.0004	0.9700
capacity2	−0.2148	0.7490	0.0278	0.4150	−0.0012	0.6040
capacity3	−1.7937	0.2290	0.0301	0.6980	0.0059	0.1880
sedan	−1.8225	0.4120	0.0764	0.5160	0.0020	0.8180
city	−1.2716	0.0400	0.0064	0.3810	−0.0057	0.7170
north	−0.3125	0.8820	−0.0010	0.8990	0.0014	0.8100
south	−0.0219	0.9920	−0.0013	0.8610	0.0015	0.9060
middle	1.3212	0.5400	−0.0012	0.8430	0.0051	0.7370
Control variables						
season1	−0.0252	0.7860				
season2	0.0324	0.7160				
season3	0.0554	0.5290				
Overall R^2			0.0550			
Number of observations			25,226			

REFERENCES

Belmont, D. M., 1953, Effect of Average Speed and Volume on Motor-Vehicle Accidents on Two-Lane Tangents, *Proceedings of Highway Research Board*, 32: 385-395.

Dionne, G., and O. Ghali, 2005, The (1992) Bonus-Malus System in Tunisia: An Empirical Evaluation, *Journal of Risk and Insurance*, 72(4): 609-633.

Dupor, B., and W. Liu, 2003, Jealousy and Equilibrium Overconsumption, *American Economic Review*, 93(1): 423-428.

Edlin, A. S., and P. Karaca-Mandic, 2006, The Accident Externality From Driving, *Journal of Political Economy*, 114(5): 931-955.

Huang, R. J., and L. Y. Tzeng, 2008, Consumption Externality and Equilibrium Underinsurance, *Journal of Risk and Insurance*, 75(4): 1039-1054.

Lundy, R. A., 1965, Effect of Traffic Volumes and Number of Lanes on Freeway Accident Rates, *Highway Research Record*, 99: 138-156.

Parry, I. W. H., 2005, Is Pay-as-You-Drive Insurance a Better Way to Reduce Gasoline Than Gasoline Taxes? *American Economic Review*, 95(2): 288-293.

Parry, I. W. H., M. Walls, and W. Harrington. 2007. Automobile Externalities and Policies, *Journal of Economic Literature*, 45(2): 373-399.

Saito, K., T. Kato, and T. Shimane, 2007, Traffic Congestion and Accident Externality: A Japan-U.S. Comparison, Presented at the of Third Annual Asia-Pacific Economic Association (APEA) Conference, July 25–26, Hong Kong, China.

Turner, D. J., and R. Thomas, 1986, Motorway Accidents: An Examination of Accident Totals, Rates and Severity and Their Relationship With Traffic Flow, *Traffic Engineering and Control* 27(July/August): 377-383.

Vickrey, W., 1968, Automobile Accidents, Tort Law, Externalities, and Insurance: An Economist's Critique, *Law and Contemporary Problems* 33(Summer): 464-487.

Wang, J. L, C. Chung, and L. Y. Tzeng, 2008, An Empirical Analysis of the Effects of Increasing Deductibles on Moral Hazard, *Journal of Risk and Insurance* 75(3): 551-566.