

# Empirical Evidence for Advantageous Selection in the Commercial Fire Insurance Market\*

Kili C. Wang<sup>a,b</sup>, Rachel J. Huang<sup>c</sup> and Larry Y. Tzeng<sup>d</sup>

<sup>a</sup>Department of Insurance, TamKang University, Taipei 104, Taiwan.

<sup>b</sup>Cardif Bancassurance Research Center, No. 70 Ta Chih Street, Taipei 104, Taiwan.

<sup>c</sup>Department of Finance, Yuan Ze University, Taipei, Taiwan.

<sup>d</sup>Department of Finance, National Taiwan University, Taipei, Taiwan.

De Meza and Webb (2001) indicated that individuals with a higher degree of risk aversion would demand more insurance and invest in self-protection to reduce risk probability when both the preference type and investment in self-protection are hidden from insurers. They referred to the negative correlation between market insurance and risk type as advantageous selection. However, the relationship between risk type and the degree of risk aversion is debatable in both theoretical and empirical research. This paper therefore proposes that advantageous selection could be supported from another angle by directly examining the relationships that exist among market insurance, self-protection, and risk probability. By focusing on the commercial fire insurance market, information on the purchase of market insurance, investment in self-protection, and fire accident records is hand-collected by means of a unique survey. It is found that firms purchasing market insurance have a greater tendency to channel efforts into self-protection. It is also found that firms expending effort on self-protection are less likely to suffer a fire accident. Furthermore, it is found that firms with commercial fire insurance have less chance of suffering a fire accident than those without such insurance. Each of the above three findings jointly supports the view that advantageous selection could play a critical role in the commercial fire insurance market.

*The Geneva Risk and Insurance Review* (2009) 34, 1–19. doi:10.1057/grir.2008.13

**Keywords:** asymmetric information; advantageous selection; adverse selection; risk aversion

---

## Introduction

Inspired by Rothschild and Stiglitz (1976), researchers have long studied the equilibrium of the insurance market under adverse selection. The traditional theoretical literature shows that, in an insurance market under adverse

---

\*The authors gratefully acknowledge helpful comments received from the participants at the 2006 annual meetings of the American Risk and Insurance Associations and the European Group of Risk and Insurance Economists, and the two anonymous reviewers.

selection, there could exist a separating equilibrium where insurance coverage and risk occurrence are positively correlated with each other. However, recent empirical findings are more debatable. While some of the studies confirm a positive correlation between insurance coverage and risk occurrence in health insurance (Cardon and Hendel, 2001) and the annuities market (Finkelstein and Poterba, 2004), a striking number of them cannot find strong evidence for such a positive correlation, and even find a negative correlation in other insurance markets such as life insurance (Cawley and Philipson, 1999), long-term care insurance (Finkelstein and McGarry, 2006), reverse mortgage insurance (Davidoff and Welke, 2004), dental care insurance (Gronqvist, 2004), and medigap insurance (Fang *et al.*, 2008).

De Meza and Webb (2001) provide an intriguing theory to explain why the previous literature cannot find evidence to support the existence of adverse selection. They show that adverse selection might be replaced by “advantageous selection”<sup>1</sup> if both the individuals’ risk preference and precautionary actions are unobservable to the insurer. They assume that according to the risk preference there are two types of individuals: the timid and the bold. Both types of individuals can invest in self-protection to reduce the loss probability. On the other hand, insurance companies have no information regarding the risk preferences of individuals. However, they try to provide a product menu for individuals to self-select insurance contracts. They show that, under certain conditions, the market could settle on a separating equilibrium where timid individuals purchase market insurance as well as make an effort to reduce the loss probability, whereas bold individuals neither purchase market insurance nor make an effort to reduce the loss probability. Thus, their model predicts that the individuals who purchase insurance suffer a lower accident probability than those who do not. In other words, the purchase of market insurance or the choices of coverage levels are not necessarily a bad signal for the risk level. This finding induces an important empirical issue, which is to verify whether purchasing insurance is a bad or a good signal in insurance markets.

This paper provides empirical evidence supporting advantageous selection. It contributes to the literature in two ways. First, we examine advantageous selection from another angle, by directly testing the relationships between the purchase of market insurance, the investment in self-protection, and the risk probability. Second, the market we focus on is that of commercial fire insurance, which is an important line in property and casualty insurance. Moreover, most empirical papers that study asymmetric information use data

---

<sup>1</sup> Hemenway (1990, 1992) referred to this favorable selection for the insurance companies as “propitious selection”.

based on personal insurance. We thus extend the literature to encompass commercial insurance.

In empirical studies, a negative correlation between the coverage provided by the insurance and the occurrence of the risk might not be enough to support the existence of advantageous selection. As shown by Chiappori *et al.* (2007), a negative correlation between the coverage provided by the insurance and the occurrence of the risk can be explained jointly by market power and privately known risk aversions. Thus it is very important to provide other and more direct evidence besides the negative correlation between the coverage of the insurance and the occurrence of the risk in order to support the existence of advantageous selection.

The theory of De Meza and Webb (2001) is able to generate at least two other empirical hypotheses: a positive relationship between the investment in self-protection and the purchase of insurance, and a negative relationship between the investment in self-protection and the occurrence of risk. De Meza and Webb (2001) indicate that individuals purchasing market insurance are more willing to make an effort. This hypothesis certainly deserves a rigorous examination, since as shown by Ehrlich and Becker (1972) market insurance and self-protection could be either substitutes or complements.<sup>2</sup> They also argue that individuals making an effort to reduce their loss probabilities indeed have less chance of suffering a loss. This hypothesis is not trivial. For example, there are two types of individuals: one with a loss probability of 0.01, and the other with a loss probability of 0.02. It is assumed that only high-risk individuals make an effort to reduce their loss probability and reduce it from 0.02 to 0.015. In this case, we may find that individuals making an effort to reduce their loss probabilities still have a greater chance of suffering a loss. Thus, instead of focusing on the test of the negative correlation between the coverage of the insurance and the occurrence of the risk as argued in the literature, we investigate the relationship between the insurance coverage and the risk occurrence as well as the other two hypotheses together.

There are three reasons why we do not measure the degree of risk aversion as other papers do in support of advantageous selection. First, Jullien *et al.* (1999) have shown that an agent with a higher degree of risk aversion will not necessarily invest more in self-protection. Second, the empirical evidence for the correlation between the degree of risk aversion and the risk type is indeed mixed. Finkelstein and McGarry (2006) find that individuals with a higher

---

<sup>2</sup> Ehrlich and Becker (1972) find that market insurance and self-protection might be complements under the assumptions that the insurers have full information regarding the individuals' risk type, but cannot observe their self-protection behavior. They also find that market insurance and self-insurance, which could reduce the risk severity, are shown to be substitutes.

degree of risk aversion are more likely to purchase long-term care insurance and less likely to enter a nursing home. On contrary, Cohen and Einav (2007) find a positive correlation between the individual's degree of risk aversion and loss probability in automobile insurance. Third, not only hidden risk preferences as indicated by De Meza and Webb (2001) but also hidden risk perceptions (Koufopoulos, 2003; Huang *et al.*, 2006), hidden regret (Huang *et al.*, 2008), and hidden patience (Sonnenholzner and Wambach, 2006) are capable of generating advantageous selection. The above papers assume that the *ex ante* risk probabilities for all individuals are the same but the individuals' own characteristic (hidden information) such as risk preference, risk perception, degree of regret, or degree of patience could affect their investment decision regarding self-protection (hidden action), and further change their risk probabilities. They predict that the market could settle on a separating equilibrium where individuals with certain characteristics would expend a great amount of effort reducing their loss probabilities and purchasing more insurance, whereas other types of individuals do not expend any effort on loss prevention and purchase less insurance. Since there are many rationales for advantageous selection and they all give rise to the same predictions for all three hypotheses, focusing on the relationships that exist among loss probability, self-protection, and market insurance rather than examining the individual's degree of risk aversion could provide direct evidence for advantageous selection.

The hypotheses developed in this paper are concerned with three endogenous decisions: the purchase of market insurance, efforts directed toward self-protection activities, and the occurrence of the accident. Since we focus on testing the advantageous selection caused by hidden information, we investigate the significance of the correlation between those three endogenous variables after controlling for observable information. In this paper, we use two methods to examine the correlation between the three endogenous variables. The first involves a two-stage method, similar to the method derived by Dionne *et al.* (2001). The second one is developed by Gourieroux *et al.* (1987).<sup>3</sup>

Both tests suggest that firms purchasing market insurance have a greater tendency to make an effort to engage in self-protection. We also find that firms making such efforts are less likely to suffer a fire accident. Furthermore, we find that firms with commercial fire insurance have less chance of suffering a fire accident than those without such insurance. All the above three findings jointly support the view that advantageous selection could play a critical role in the commercial fire insurance market.

---

<sup>3</sup> Chiappori and Salanie (2000) use this method to examine the existence of the asymmetric information.

The remainder of this paper is organized as follows. The next section describes the data collection process and the design of the questionnaire. The subsequent section provides the summary statistics. The empirical methodology is introduced in the fourth section. The penultimate section reports the empirical results. The last section concludes the paper.

### **Data collection and questionnaire design**

To achieve our goal, we hand collected the data since using the data sets from the insurance companies was not sufficient. As suggested by Chiappori and Salanie (1997), the records kept by insurance companies will of course provide excellent data for conducting empirical research on asymmetric information. However, there are two drawbacks to the data obtained from the insurance companies. First, insurance companies do not have information on the non-insured. Thus, it might not be appropriate to examine the behavioral differences between the insured and non-insured by only using data obtained from insurance companies.<sup>4</sup> The previous literature bypasses this problem by focusing on whether individuals with higher insurance coverage have a higher chance of suffering a loss than those with lower insurance coverage, since insurance companies have information on both the insured with high coverage and the insured with low coverage. Second, and in particular, insurance companies may not have information on individuals' self-protection activities. Without data on individuals' self-protection activities, it may be even more difficult to directly investigate our hypotheses. We therefore overcome these two problems by using survey data.

Since the heterogeneity of the data could give rise to estimation bias as mentioned by Chiappori and Salanie (2000), we restrict our sample in terms of the business locations and business types. We focus on class-A businesses, which are classified according to business type in the standard for the installation of fire safety equipment contained in the Fire Act in Taiwan based on use and occupancy.<sup>5</sup> The government in Taiwan classifies firms as class-A

---

<sup>4</sup> Note that both adverse selection and advantageous selection could cause a separating equilibrium for which one type of individual has insurance and the other type of individual does not have insurance. Thus, examining the behavioral differences between the insured and non-insured rather than between the insured with different levels of insurance coverage is still linked to the theoretical setting.

<sup>5</sup> The Fire Act in Taiwan classifies business into six types. The class-A places include movie projection places, singing rooms, clubs, restaurants, bowling alleys, indoor screen-type golf training fields, game-playing grounds, sightseeing hotels, guest houses, rest houses, stores, markets, department stores, supermarkets, retail markets, exhibition fields, dining rooms, catering shops, coffee bars, tearooms, hospitals, sanitariums, long-term nursing homes, healthcare institutions, elderly service institutions, saunas, and public baths, etc.

due to the belief that this type of firm has a higher probability of loss than those in other places. We further sample the firms from the two largest cities in Taiwan, Taipei and Kaohsiung, because the fire risk factors of rural or urban businesses could be quite different. For example, the response times of the fire department, materials/codes and height of building, number of neighbors, etc., could be diversified between rural and urban areas, which also cause the fire risk to be diversified between rural and urban areas.

In the questionnaires, we collected information regarding (1) market insurance, (2) records of fire accidents,<sup>6</sup> (3) self-protection activities, and (4) other relevant variables on which insurance companies usually collect data to underwrite commercial fire insurance, such as the locations and building characteristics of the firms, and the basic fire safety equipment installed.

Based on the purchases of market insurance in the year 2004, we separated the insured firms from the non-insured firms. To test for advantageous selection, we analyzed the differences in terms of the self-protection behavior and accident records between firms that buy the basic fire insurance<sup>7</sup> and those that do not buy such coverage.

We obtained information as to whether a firm had faced a fire accident in the year 2004. The fire accident records are more reliable than claim records when it comes to estimating the probability of a fire accident.<sup>8</sup> The claim records kept by the insurance companies underestimate the occurrence of accidents, because the insured may not report accidents involving small amounts of claims due to either the provisions of deductibles or the experience-rating system. Thus, the accident records used in this paper provide more reliable data for estimating the probability of the risk.

Most importantly, we collected information on the self-protection activities of firms, which are made throughout the year 2004. There are two criteria for a variable used to represent a self-protection activity. The literature has defined self-protection as a non-market risk reduction activity to reduce the probability of a loss. Thus, the first criterion for a proxy of self-protection is that the activity could reduce the likelihood that a fire will occur. The second criterion is that this activity to reduce risk is unobservable to insurers. In the models of De Meza and Webb (2001), Koufopoulos (2003), and Huang *et al.* (2006), insurers are assumed to be unable to observe precautions directly, but they

---

<sup>6</sup> It is worth noting that while we can observe the fire accident records, papers using data from insurance companies only have claim records. Chiappori and Salanie (2000) noted that using claim records as a proxy for accident records could result in biased estimates.

<sup>7</sup> The basic commercial fire insurance covers the damage to the property, including chattels and real estate, of the insured.

<sup>8</sup> If we are interested in the total effect of *ex ante* (risk prevention) and *ex post* (claim choice) moral hazard, then claims data may be more useful than loss data.

could estimate the probability of a loss when a hidden action is taken. This criterion rules out certain activities, such as the installation of fire safety equipment.

The first proxy for self-protection is the qualified operation of fire safety equipment. Insurance companies not only can observe the basic safety equipment, but they also take such equipment into account during the process of underwriting the commercial fire insurance and enable the insured to benefit from the point of view of the premium paid. However, the insurance companies cannot be certain how such equipment is maintained from that time on, since they do not follow up on whether such equipment could function well during the policy period. The fire departments are responsible for and implement fire safety inspections every half year, being able to understand how all of the fire safety equipment functions. In our survey data, with the help of certain fire departments, we collected the information regarding the qualified operation of the fire safety equipment.<sup>9</sup> The qualified operation of fire safety equipment is identified by the fire departments through their inspections, and gives us a chance of obtaining a good proxy for self-protection activities.

Another suitable proxy which satisfies the above two criteria is related to the operations of the self-defense fire organizations. A self-defense fire organization is an informal organization within a firm. Every member of this organization has competent knowledge regarding fire control and hazard reduction. To establish knowledge regarding fire control, all of the members need to attend relevant courses every half year. The tasks in the organization include preventing the occurrence of a fire, checking whether all of the fire safety equipments are operable, clearing the obstacles in an emergency entrance, reporting the occurrence of a fire to a fire station and to neighbors, guiding customers or employees to emergency exits, and handling fire safety equipment, etc. All of the above tasks can effectively reduce the possibility of a fire. The operation of such self-defense fire safety organizations is reviewed by the fire departments only every half year. The insurance companies in Taiwan do not collect this information while underwriting.

It should be recognized that we also collected data on all the relevant variables that insurance companies use in pricing commercial fire insurance.<sup>10</sup> There are two effects of these exogenous variables. First, they can be used to

---

<sup>9</sup> The definition of the qualified operation of fire safety equipment is that not only does the fire safety equipment exist, but it should also function well when the fire department conducts a fire safety inspection.

<sup>10</sup> With the help of a few high-ranking executives in insurance companies, we double-checked that we had collected data on all the variables that insurance companies use in underwriting commercial fire insurance.

control the insurer's known information. Advantageous selection is caused by a firm's hidden risk preference or a firm's hidden risk perception, which the insurance companies do not know. Controlling the information that insurance companies do know is essential for a proper test of advantageous selection. Second, data are obtained from a heterogeneous sample. As noticed by Chiappori and Salanie (1997), the estimator used to test for asymmetric information would be biased if a test did not appropriately control for the heterogeneity of the sample.

### **Basic statistics**

With the help of the fire departments of the cities of Taipei and Kaohsiung, we sent out 6,249 questionnaires, and received 3,361 of them back. After eliminating 523 questionnaires with missing variables, 2,838 effective questionnaires were left.

To test our hypothesis of advantageous selection, we further excluded observations from the sample if firms had purchased commercial fire insurance due to the requirements of a loan, since in that case the theories of adverse selection and advantageous selection cannot be applied. After excluding those firms, we were left with a sample of 2,592 questionnaires for our empirical tests.

Tables 1 and 2 define variables and give descriptive statistics. The sample is well distributed in terms of business types as shown in Table 2. Within class-A places, Taiwan's Fire Act further classifies firms into seven sub-classes. Table 2 shows that the final sample does not concentrate on any specific business. For the seven types of business, the percentages are mostly near 10 percent, and the share of the highest one is 30 percent, which includes a dining room, catering shop, coffee bar, and tearoom. Almost 87 percent of the sample is located in Taipei city. Within Taipei city, 15 percent of the sample is in the Ta-an district (with the postal area code 106), which is a shopping and commercial district in Taipei.

From Table 2, we can observe some basic characteristics of our final sample. About 40 percent of the firms purchased basic commercial fire insurance, while about 5 percent of them had experienced a fire accident. As for self-protection, 74.6 percent of firms had all four kinds of fire safety equipment functioning properly in the most recent half year, and 75.3 percent of the firms in our sample had operated a self-defense fire organization. Almost 80 percent of the firms were located in buildings with less than 10 floors, and over 80 percent of them had basic fire extinguishing equipment, fire alarm equipment, and an area of refuge as well as escape equipment. The percentage of firms with the basic equipment needed for a rescue in the event of a fire was a little bit lower, but still reached 74.01 percent.



**Table 1** Definitions of variables

<i>Dependent variables</i>	
<i>ins</i>	A dummy variable that equals 1 when the firm had bought basic fire insurance last year, otherwise 0.
<i>accid</i>	A dummy variable that equals 1 when a fire occurred last year, otherwise 0.
<i>sp_qualified</i>	A dummy variable that equals 1 when all four kinds of fire safety equipment <sup>a</sup> are qualified for operation in the last half year, otherwise 0.
<i>sp_org</i>	A dummy variable that equals 1 when the self-defense fire organization has operated in the last half year, otherwise 0.
<i>Independent variables</i>	
<i>Group A variables</i>	
<i>busi_i</i>	A dummy variable that equals 1 when an individual firm is one of the <i>i</i> businesses, where <i>i</i> = 1–6, otherwise 0. The reference group is the sauna or public baths.
<i>area_aaa</i>	A dummy variable that equals 1 when aaa is one of the zip codes, otherwise 0. The reference code is that for Zuoying.
<i>build</i>	A dummy variable that equals 1 when the height of the building is less than 10 floors, otherwise 0.
<i>el</i>	A dummy variable that equals 1 when there is basic fire extinguishing equipment, otherwise 0.
<i>al</i>	A dummy variable that equals 1 when there is basic fire alarm equipment, otherwise 0.
<i>ll</i>	A dummy variable that equals 1 when there is a basic area of refuge and escape equipment, otherwise 0.
<i>rl</i>	A dummy variable that equals 1 when there is basic equipment needed for fire rescue, otherwise 0.
<i>Group B variables</i>	
<i>eh</i>	A dummy variable that equals 1 when there is other fire extinguishing equipment that goes beyond the requirements of the law, otherwise 0.
<i>ah</i>	A dummy variable that equals 1 when there is other fire alarm equipment that goes beyond the requirements of the law, otherwise 0.
<i>lh</i>	A dummy variable that equals 1 when there are other areas of refuge and escape equipment that go beyond the requirements of the law, otherwise 0.
<i>rh</i>	A dummy variable that equals 1 when there is other necessary equipment for fire rescue that goes beyond the requirements of the law, otherwise 0.
<i>manage</i>	A dummy variable that equals 1 when the person who is responsible for the decision regarding preventing fire hazards is at least at the managerial level.

<sup>a</sup>The four kinds of fire safety equipment include extinguishing equipment, fire alarm equipment, areas of refuge and escape equipment, and fire rescue equipment.

Notes: 1. The exogenous variables in the  $X_{1i}$  vector include the group A and group B independent variables in the above table.

2. The exogenous variables in the  $X_{2i}$  vector include the group A independent variables in the above table.

3. The exogenous variables in the  $X_{3i}$  vector include the group A and group B independent variables in the above table.

**Table 2** Summary statistics for the final sample

<i>Variables</i>	<i>Mean</i>	<i>Standard deviation</i>
<i>Dependent variables</i>		
<i>ins</i>	0.4247	0.4896
<i>accid</i>	0.0579	0.2336
<i>sp_qualified</i>	0.7460	0.4354
<i>sp_org</i>	0.7530	0.4313
<i>Independent variables</i>		
<i>busi_1</i>	0.2448	0.4301
<i>busi_2</i>	0.0949	0.2931
<i>busi_3</i>	0.0932	0.2908
<i>busi_4</i>	0.1187	0.3235
<i>busi_5</i>	0.3055	0.4607
<i>busi_6</i>	0.1346	0.3413
<i>area_100</i>	0.0535	0.2251
<i>area_103</i>	0.0293	0.1685
<i>area_104</i>	0.1100	0.3128
<i>area_105</i>	0.0381	0.1913
<i>area_106</i>	0.1483	0.3555
<i>area_108</i>	0.1041	0.3298
<i>area_110</i>	0.0652	0.2469
<i>area_111</i>	0.1074	0.3097
<i>area_112</i>	0.0632	0.2762
<i>area_114</i>	0.0681	0.2520
<i>area_115</i>	0.0380	0.0933
<i>area_116</i>	0.0318	0.1754
<i>area_800</i>	0.0117	0.1076
<i>area_801</i>	0.0100	0.0997
<i>area_804</i>	0.0109	0.1037
<i>area_806</i>	0.0251	0.1564
<i>area_807</i>	0.0460	0.2095
<i>area_811</i>	0.0096	0.0976
<i>area_812</i>	0.0084	0.0911
<i>area_813</i>	0.0109	0.1037
<i>build</i>	0.7919	0.4060
<i>el</i>	0.9887	0.1056
<i>eh</i>	0.0915	0.2884
<i>al</i>	0.8905	0.3123
<i>ah</i>	0.0464	0.2104
<i>ll</i>	0.8625	0.3444
<i>lh</i>	0.0514	0.2209
<i>rl</i>	0.7401	0.4387
<i>rh</i>	0.0364	0.1872
Number of observations	2,592	

*Note:* All variables are dummy variables.

## Empirical methodology

To test the three hypotheses generated from the theories of advantageous selection, which include the existence of a positive relationship between market insurance and self-protection, a negative relationship between self-protection and the probability of a fire accident, and a negative relationship between market insurance and the risk type, two different methods are employed in this paper.

The first empirical model involves a two-stage method, similar to that of Dionne *et al.* (2001). There are three interactive endogenous variables: the purchase of the insurance, the occurrence of the fire accident, and the self-protection effort. In this model, we simultaneously estimate the probability of each endogenous variable that is conditional upon the other two endogenous variables as well as other exogenous variables. Since the non-linearity among the endogenous variables is both important and crucial to the results as suggested by Dionne *et al.* (2001), we employ the two-stage regression presented below.

More precisely, in the first stage, we run three probit regressions as follows:

$$\text{Prob}(ins_i = 1|X_{1i}) = \Phi(X_{1i}\beta_{1ins}), \quad (1)$$

$$\text{Prob}(accid_i = 1|X_{1i}) = \Phi(X_{1i}\beta_{1accid}), \quad (2)$$

and

$$\text{Prob}(sp_i = 1|X_{1i}) = \Phi(X_{1i}\beta_{1sp}), \quad (3)$$

where  $i = 1, \dots, n$  denotes the firm,  $\text{Prob}(\bullet)$  denotes a probability function,  $\beta_{1j}$ ,  $j = ins, accid, sp$ , is the corresponding parameter vector, and  $\Phi$  is the cumulative distribution function of  $N(0, 1)$ . In addition,  $X_{1i}$  is an exogenous variable vector for firm  $i$  which is defined in note 1 of Table 1, and the variables in  $X_{1i}$  are all dummy variables. The three endogenous dummy variables are as follows: (1)  $ins_i = 1$  if firm  $i$  is insured under the basic fire insurance;  $ins_i = 0$  if firm  $i$  is uninsured; (2)  $accid_i = 1$  if firm  $i$  has suffered a fire accident;  $accid_i = 0$  otherwise; and (3)  $sp_i = 1$ <sup>11</sup> if firm  $i$  engages in self-protection;  $sp_i = 0$  otherwise. We record the estimated probabilities of the three endogenous variables,  $\text{Pr}\hat{\text{ob}}(ins_i = 1|X_{1i})$ ,  $\text{Pr}\hat{\text{ob}}(accid_i = 1|X_{1i})$  and  $\text{Pr}\hat{\text{ob}}(sp_i = 1|X_{1i})$ , from probit regressions (1) to (3).

---

<sup>11</sup> To measure a firm's self-protection activities, we use two proxies,  $sp\_qualified$  and  $sp\_org$ , as defined in Table 1.

Then, in the second stage, we run another three probit regressions for each of the three endogenous variables:

$$\begin{aligned} & \text{Prob}(ins_i = 1|X_{2i}, accid_i, sp_i, \text{Pr}\hat{\text{ob}}(accid_i = 1|X_{1i}), \\ & \text{Pr}\hat{\text{ob}}(sp_i = 1|X_{1i})) = \Phi(X_{2i}\beta_{2ins} + b_{ins,accid}accid_i + b_{ins,sp}sp_i \\ & + b_{ins,eaccid} \text{Pr}\hat{\text{ob}}(accid_i = 1|X_{1i}) \\ & + b_{ins,esp} \text{Pr}\hat{\text{ob}}(sp_i = 1|X_{1i})), \end{aligned} \quad (4)$$

$$\begin{aligned} & \text{Prob}(accid_i = 1|X_{2i}, ins_i, sp_i, \text{Pr}\hat{\text{ob}}(ins_i = 1|X_{1i}), \\ & \text{Pr}\hat{\text{ob}}(sp_i = 1|X_{1i})) = \Phi(X_{2i}\beta_{2accid} + b_{accid,ins}ins_i + b_{accid,sp}sp_i \\ & + b_{accid,eins} \text{Pr}\hat{\text{ob}}(ins_i = 1|X_{1i}) \\ & + b_{accid,esp} \text{Pr}\hat{\text{ob}}(sp_i = 1|X_{1i})), \end{aligned} \quad (5)$$

$$\begin{aligned} & \text{Prob}(sp_i = 1|X_{2i}, accid_i, ins_i, \text{Pr}\hat{\text{ob}}(accid_i = 1|X_{1i}), \\ & \text{Pr}\hat{\text{ob}}(ins_i = 1|X_{1i})) = \Phi(X_{2i}\beta_{2sp} + b_{sp,accid}accid_i + b_{sp,ins}ins_i \\ & + b_{sp,eaccid} \text{Pr}\hat{\text{ob}}(accid_i = 1|X_{1i}) + b_{sp,eins} \text{Pr}\hat{\text{ob}}(ins_i = 1|X_{1i})), \end{aligned} \quad (6)$$

$\beta_{2j}$ ,  $j = ins, accid, sp$ , is the parameter vector of the exogenous vector  $X_{2i}$ , which differs from  $X_{1i}$  and is listed in note 2 of Table 1.  $b_{j,k}$  and  $b_{j,ek}$  are the coefficients of the variables  $k$  and  $\text{Pr}\hat{\text{ob}}(k_i = 1|X_{1i})$  on the dependent variable  $j$ , respectively, where  $j, k = ins, accid, sp, j \neq k$ . We test the dependency among the three endogenous variables through the significance of the coefficients  $b_{ins,accid}$  and  $b_{ins,sp}$  in regression (4),  $b_{accid,ins}$  and  $b_{accid,sp}$  in regression (5), and  $b_{sp,accid}$  and  $b_{sp,ins}$  in regression (6). According to our three hypotheses, in regression (4), the coefficient  $b_{ins,accid}$  should be significantly negative, while the coefficient  $b_{ins,sp}$  should be significantly positive. In regression (5), the coefficient  $b_{accid,ins}$  should be significantly negative, while the coefficient  $b_{accid,sp}$  should be significantly negative. In regression (6), the coefficient  $b_{sp,accid}$  should be significantly negative, while the coefficient  $b_{sp,ins}$  should be significantly positive.

The second empirical method follows a reduced form approach. This method is derived by Gourieroux *et al.* (1987) and proposed by Chiappori and Salanie (2000). We test the conditional dependency between the three endogenous variables through the significance of the residual correlation between each pair of them. Specifically, we establish three probit regressions:

$$\text{Prob}(ins_i = 1|X_{3i}) = \Phi(X_{3i}\beta_{3ins}), \quad (7)$$

$$\text{Prob}(\text{accid}_i = 1 | X_{3i}) = \Phi(X_{3i}\beta_{3\text{accid}}), \quad (8)$$

and

$$\text{Prob}(\text{sp}_i = 1 | X_{3i}) = \Phi(X_{3i}\beta_{3\text{sp}}), \quad (9)$$

where  $\text{Prob}(\bullet)$ ,  $\Phi$ ,  $\text{ins}_i$ ,  $\text{accid}_i$ , and  $\text{sp}_i$  are as defined in the previous model,  $X_{3i}$  is the vector of control variables which is described in note 3 of Table 1, and  $\beta_{3j}$ ,  $j = \text{ins}, \text{accid}, \text{sp}$ , is the corresponding parameter vector.

Let  $\varepsilon_i^j$  be the independent centered normal errors with unit variance for the endogenous variable  $j$  in the above regressions. The estimated residuals  $\hat{\varepsilon}_i^j$  from Eqs. (7) to (9) are defined as

$$\hat{\varepsilon}_i^j = E(\varepsilon_i^j | j_i, X_{3i}) = \frac{\phi(X_{3i}\beta_{3j})}{\Phi(X_{3i}\beta_{3j})} j_i - (1 - j_i) \frac{\phi(-X_{3i}\beta_{3j})}{\Phi(-X_{3i}\beta_{3j})}, \quad (10)$$

where  $j = \text{ins}, \text{accid}, \text{sp}$ , and  $\phi$  represents the density distribution function of  $N(0, 1)$ .

To test the conditional dependence, a statistic  $W_{jk}$ , derived by *Gourieroux et al.* (1987), is obtained as

$$W_{jk} = \frac{(\sum_{i=1}^n \hat{\varepsilon}_i^j \hat{\varepsilon}_i^k)^2}{\sum_{i=1}^n (\hat{\varepsilon}_i^j)^2 (\hat{\varepsilon}_i^k)^2}, \quad j, k = \text{ins}, \text{accid}, \text{sp}, \quad j \neq k. \quad (11)$$

Under the null hypothesis of  $\text{cov}(\varepsilon_i^j, \varepsilon_i^k) = 0$ , the  $W_{jk}$  statistic is distributed asymptotically as  $\chi^2(1)$ . If  $W_{jk}$  is significant, we reject the hypothesis that the probability of the occurrence of the endogenous variables  $j$  and  $k$  is conditionally independent. However, the  $W_{jk}$  statistic reveals nothing about the direction of the dependency. In order to clarify the direction of the dependence, we further calculate the conditional correlation coefficient  $\rho_{jk}$  for  $\hat{\varepsilon}_i^j$  and  $\hat{\varepsilon}_i^k$ . Under the predictions made based on the theories of advantageous selection, the  $W_{jk}$  for all  $j$  and  $k$ ,  $j \neq k$  are hypothesized to be significant and  $\rho_{\text{ins}, \text{sp}} > 0$ ,  $\rho_{\text{ins}, \text{accid}} < 0$ , and  $\rho_{\text{accid}, \text{sp}} < 0$ .

## Empirical results

The empirical results obtained from employing the two-stage method are listed in Tables 3 and 4. In Table 3, the proxy variable for self-protection measures the qualified operation of all four kinds of fire safety equipment (*sp\_qualified*). From Table 3, we find that  $b_{\text{ins}, \text{accid}}$  in regression (4) and  $b_{\text{accid}, \text{ins}}$  in regression (5) are both significantly negative. The purchase of the commercial fire insurance and the occurrence of fire accidents are significantly negatively correlated. This finding is consistent with *Cawley and Philipson* (1999), *Dionne*

**Table 3** The probit regression on the purchase of market insurance, risk probability, and investment in self-protection – using the qualified operation of all four kinds of fire safety equipment (*sp\_qualified*) as the proxy for investment in self-protection

<i>Dependent variable</i> <i>Variable</i>	<i>ins</i> (4)	<i>accid</i> (5)	<i>sp_qualified</i> (6)
Intercept	-0.4917	0.3290	-0.2089
<i>ins</i>		-0.3375***	0.1900***
<i>accid</i>	-0.3662***		-0.6552***
<i>sp_qualified</i>	0.1893**	-0.6127***	
Prôb( <i>ins</i>   $X_2$ )		-0.8330*	2.0908***
Prôb( <i>accid</i>   $X_2$ )	-0.9289**		-4.0407***
Prôb( <i>sp_qualified</i>   $X_2$ )	0.5437*	-1.4835**	
busi_1	-0.0449	-1.2131***	-1.2240***
busi_2	-0.3603	-1.5073***	-1.0186**
busi_3	0.1967	-0.8176**	-1.1646***
busi_4	0.3019	-1.3983***	-1.7457***
busi_5	-0.2013	-1.1320***	-0.9596**
busi_6	0.0825	-1.0608***	-1.2228***
area_100	0.2966	0.0931	-0.8423
area_103	-1.5107***	-0.5394	0.1067
area_104	-0.7530**	-1.4038*	-0.5788
area_105	-0.3535	-0.4581	-0.5970
area_106	-1.1150***	0.2153	0.6798
area_108	-0.0929	-0.4038	-0.5945
area_110	-0.7988***	-0.7194	-0.4925
area_111	-1.3671***	-0.9605	0.0710
area_112	0.0673	-0.5496	-0.5412
area_114	-0.1983	-0.8110	-1.1054**
area_115	-0.3997	-0.3150	-0.4172
area_116	-0.0277	-0.2513	-0.8042
area_800	1.5702***	-0.3148	-1.9960***
area_801	0.7676*	-5.8092	-1.7331***
area_804	-0.3334	-5.2462	4.8407
area_806	-0.0416	-0.2888	-0.6495
area_807	-0.6024*	-0.6476	-0.7037
area_811	-0.1895	-5.2746	-0.2845
area_812	-0.0016	-5.5219	-0.7939
build	-0.3298***	-0.0187	0.4282***
el	-0.0151	-0.2467*	0.6530**
al	0.0451	-0.5717**	0.7401***
ll	0.4410***	0.4364	0.1747
rl	0.1097	0.2346	0.1551

*Note:* The 1 percent significance level is denoted by \*\*\*, the 5 percent significance level is denoted by \*\*, and the 10 percent significance level is denoted by \*.

**Table 4** The probit regression on the purchase of market insurance, the risk probability, and the investment in self-protection – using the operation of the self-defense fire organization (*sp\_org*) as the proxy for investment in self-protection

<i>Dependent variable</i> <i>Variable</i>	<i>ins</i> (4)	<i>accid</i> (5)	<i>sp_org</i> (6)
Intercept	-0.0487	0.6922	-0.0806
<i>ins</i>		-0.2492**	0.3001***
<i>accid</i>	-0.3041**		-0.4529***
<i>sp_org</i>	0.3560***	-0.2957**	
Prôb( <i>ins</i>   $X_2$ )		-1.9162*	2.6897***
Prôb( <i>accid</i>   $X_2$ )	-1.9289*		-2.9100***
Prôb( <i>sp_org</i>   $X_2$ )	-0.4989	-0.4753	
busi_1	-0.3080	-1.0552***	-0.6647
busi_2	-0.6529	-1.5775***	-0.9647***
busi_3	-0.0162	-0.6353*	-0.2963
busi_4	-0.0276	-1.0115***	-1.0326**
busi_5	-0.4488	-1.1644***	-1.0127**
busi_6	-0.1570	-0.9326***	-0.4348
area_100	0.2636	0.5582	1.4097***
area_103	-1.6743***	-0.6290	1.6468***
area_104	-0.9501***	-1.3520**	0.4930
area_105	-0.4807*	-0.2487	0.3109
area_106	-1.0918***	-0.0478	1.7784***
area_108	-0.1297	-0.2067	1.6525***
area_110	-0.9889***	-0.5542	0.9043**
area_111	-1.4840***	-1.2360*	1.4654***
area_112	0.0561	-0.4703	0.5794**
area_114	-0.4113	-0.3405	0.5393*
area_115	-0.4529	-0.1387	1.6346***
area_116	-0.0895	0.1035	1.8665***
area_800	1.4203***	0.6032	-0.0136
area_801	0.6493	-5.0899	5.6678
area_804	-0.3399	-5.4452	0.5630
area_806	-0.1123	-0.1805	0.0679
area_807	-0.7836***	-0.1878	2.1844***
area_811	-0.2440	-5.2843	0.1119
area_812	-0.1149	-5.3646	-0.2888
build	-0.3035***	-0.2845**	-0.1059
el	0.2104	-0.3066*	-0.4781
al	0.2855**	-0.1182*	0.2874**
ll	0.5857***	0.3031	-0.3222**
rl	0.1640*	0.2454	0.2071**

*Note:* The 1 percent significance level is denoted by \*\*\*, the 5 percent significance level is denoted by \*\*, and the 10 percent significance level is denoted by \*.

*et al.* (2001), and Finkelstein and McGarry (2006), who also find a negative correlation between insurance coverage and the occurrence of losses.

The  $b_{ins,sp}$  in regression (4) and  $b_{sp,ins}$  in regression (6) are both significantly positive, and a positive conditional correlation is found between the purchase of commercial fire insurance and self-protection. The findings support the view that firms that purchase such insurance in the market indeed make more efforts to engage in self-protection in accordance with the theory of advantageous selection.<sup>12</sup>

The  $b_{accid,sp}$  in regression (5) and  $b_{sp,accid}$  in regression (6) are both significantly negative which means that the conditional correlation between the occurrence of a loss and self-protection is significantly negative, which confirms that a firm that never experiences a fire loss has a greater likelihood of protecting itself, and thus self-protection activities could reduce the likelihood of a fire occurring. Our finding suggests that both the quantity and quality of a firm's self-protection activities are essential in order to reduce the probability of fire accidents.

In Table 4, the proxy variable for self-protection is the operation of a self-defense fire organization ( $sp\_org$ ). The empirical results in Table 4 are generally consistent with those in Table 3. In Table 4, we still find that the purchase of the insurance and the occurrence of the accident are negatively correlated, the purchase of the insurance and the efforts at self-protection are positively correlated, and the occurrence of the accident and the efforts at self-protection are negatively correlated.

Table 5 reports the pair-wise residual correlations of the three endogenous variables obtained using the second method. In Panel A of Table 5, the proxy for self-protection is the qualified operation of the fire safety equipment ( $sp\_qualified$ ). In Panel B of Table 5, the proxy for self-protection is the self-defense fire organization ( $sp\_org$ ). Note that, in Table 5, the  $W$  statistic and the conditional correlation coefficient of market insurance and the occurrence of a fire accident in both panels are the same because the use of the proxy for self-protection does not change either the endogenous or the exogenous variables of these two probit regressions. Thus, no matter what the proxy of self-protection is, Eqs. (7) and (8) remain the same.

---

<sup>12</sup> In equilibrium, when both the insurance coverage level and self-protection are endogenous, the prediction for whether a risk-averse individual will invest in more self-protection depends critically on the level of insurance chosen. A firm with low risk aversion might choose low coverage and a high level of self-protection, while a firm with high risk aversion might choose high coverage and a lower level of self-protection. If the above statement is true, then we will find a negative correlation between the investment of self-protection and insurance coverage. However, our empirical results support a positive correlation between these two endogenous variables, which is consistent with the prediction of advantageous selection.



**Table 5** Conditional correlations between endogenous variables – based on the residual correlation test method

<i>Items</i>	$W_{jk}$	$\rho_{jk}$
<i>Panel A: Qualified operation of fire safety equipment (<math>sp\_qualified</math>) as a proxy for self-protection</i>		
Market insurance vs. loss occurrence	122.035***	-0.1252***
Self-protection vs. loss occurrence	460.523***	-0.3481***
Self-protection vs. market insurance	10.1094***	0.2042***
<i>Panel B: Self-defense fire organization (<math>sp\_org</math>) as a proxy for self-protection</i>		
Market insurance vs. loss occurrence	122.035***	-0.1252***
Self-protection vs. loss occurrence	398.908***	-0.0475**
Self-protection vs. market insurance	30.5324***	0.4426***

*Notes:*  $\rho$  is the conditional correlation coefficient of coverage and accidents.  $W$  is the  $W$ -statistic of the conditional correlation coefficient of coverage and claims. The 1 percent significance level is denoted by \*\*\*, the 5 percent significance level is denoted by \*\*, and the 10 percent significance level is denoted by \*.

From Table 5, regardless of whether we use the qualified operation of the fire safety equipment or the self-defense fire organization as a proxy variable for self-protection, we find that the results sustain the three hypotheses we put forward. In both panels, we find that the purchase of the commercial fire insurance and the occurrence of the fire accidents are significantly negatively correlated. The conditional correlation between self-protection and loss occurrence is also significantly negative.<sup>13</sup> The correlation coefficients between market insurance and self-protection are, however, significantly positive.

The evidence from our tests as a whole shows that advantageous selection exists in the market.<sup>14</sup> Firms purchasing market insurance also engage in more self-protection activities. Self-protection activities indeed reduce the chances of

<sup>13</sup> In our sample, there are 246 firms purchasing commercial fire insurance due to the banks' loan requirements. The loan sample provides another chance to focus on the relationship between the investment in self-protection and the loss probability, since the insurance purchasing decision in the loan sample is required by banks. Consistent with the findings in Tables 4 and 5, we also find that self-protection is negatively correlated with the probability of a fire accident in the loan sample, although not significantly at 10 percent confident level.

<sup>14</sup> It is worth noting that we cannot identify whether the purchase of insurance or the investment in self-protection is in response to fire accidents in our sample. If both decisions are in response to fire accidents, we should predict that a firm experiencing a fire accident might learn that it is a high risk type and further purchase more insurance or invest in more self-protection. In this case, we should observe a positive relationship between self-protection and loss probability and between insurance coverage and loss probability. However, our results are not consistent with the above prediction. In other words, the bias of the timing of variables may not be serious enough to cloud our main findings.

fire accidents. Thus, we also observe that the occurrence of fire accidents and the purchase of fire insurance counter the predictions of well-known adverse selection theories.

## Conclusions and discussion

With the help of the respective fire departments of the cities of Taipei and Kaohsiung, we have hand-collected firm data which include information on market insurance and fire accidents as well as on self-protection activities. By employing these unique data, we find that: (1) firms that purchase market insurance have a greater tendency to make an effort to engage in self-protection; (2) firms that make an effort to engage in self-protection are less likely to suffer a fire accident; and (3) firms with commercial fire insurance have a lower chance of suffering a fire accident than those without such insurance. The above three findings jointly support the view that advantageous selection exists in the commercial fire insurance market in Taiwan.

Our paper contributes to the literature by providing more direct evidence to support the existence of the advantageous selection theory derived by De Meza and Webb (2001), Koufopoulos (2003), and Huang *et al.* (2006). Our findings are consistent with those of recent empirical studies that confirm the existence of advantageous selection in different lines of insurance.

For future studies examining advantageous selection in commercial insurance, several issues deserve more attention and could provide fruitful results. First, a multiple year data set will help to decrease the noise while estimating loss probabilities and strengthening the robustness of the empirical findings. However, when a dynamic data set is employed, the learning effect should be taken into consideration. Firms might learn about their true risk types from accidents and further change the investment in self-protection or the purchase of insurance in response. Second, the amount of variation in insurance coverage should be documented to completely control for insurance. This paper models the insurance coverage as a binary choice. We focus on whether the firms purchase fire insurance or not. Collecting information on the amount of insurance coverage could help to document whether the amount of insurance coverage varies importantly across firm or risk types. Third, firms are sophisticated in their risk management and their risk management behavior, such as issuing cat bonds, could affect the risk types. Most of our sample consists of small businesses. Small businesses may not be able to engage in such risk management. However, if researchers are interested in large businesses, it is important to collect the information regarding each firm's risk management behavior and to further identify whether this information is hidden to insurance companies.

## References

- Cardon, J.H. and Hendel, I. (2001) 'Asymmetric information in health insurance: Evidence from the national medical expenditure survey', *Rand Journal of Economics* **32**(3): 408–427.
- Cawley, J. and Philipson, T.J. (1999) 'An empirical examination of information barriers to trade in insurance', *American Economic Review* **89**(4): 827–846.
- Chiappori, P.A. and Salanie, B. (1997) 'Empirical contract theory: The case of insurance data', *European Economic Review* **41**(3): 943–950.
- Chiappori, P.A. and Salanie, B. (2000) 'Testing for asymmetric information in insurance markets', *Journal of Political Economy* **108**(1): 56–78.
- Chiappori, P.A., Jullien, B., Salanie, B. and Salanie, F. (2006) 'Asymmetric information in insurance: Some testable implications', *Rand Journal of Economics* **37**(4): 783–798.
- Cohen, A. and Einav, L. (2007) 'Estimating risk preferences from deductible choice', *American Economic Review* **97**(3): 745–788.
- Davidoff, T. and Welke, G. (2007) *Selection and moral hazard in the reverse mortgage market*, Working paper.
- De Meza, D. and Webb, D.C. (2001) 'Advantageous selection in insurance markets', *Rand Journal of Economics* **32**(2): 249–262.
- 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.
- Ehrlich, I. and Becker, G.S. (1972) 'Market insurance, self-insurance, and self-protection', *Journal of Political Economy* **80**(4): 623–648.
- Fang, H., Keane, M. and Silverman, D. (2008) 'Sources of advantageous selection: Evidence from the Medigap insurance market', *Journal of Political Economy* **116**(2): 303–350.
- Finkelstein, A. and McGarry, K. (2006) 'Multiple dimensions of private information: Evidence from the long-term care insurance market', *American Economic Review* **96**(4): 938–958.
- Finkelstein, A. and Poterba, J. (2004) 'Adverse selection in insurance markets: Policyholder evidence from the U.K. annuity market', *Journal of Political Economy* **112**(1): 183–208.
- Gourieroux, C., Monfort, A., Renault, E. and Trognon, A. (1987) 'Generalized residuals', *Journal of Econometrics* **34**(1): 5–32.
- Gronqvist, E. (2004) *Does adverse selection matter? Evidence from a natural experiment*, SSE/EFI Working Paper Series in Economics and Finance No. 575.
- Hemenway, D. (1990) 'Propitious selection', *Quarterly Journal of Economics* **105**(4): 1063–1069.
- Hemenway, D. (1992) 'Propitious selection in insurance', *Journal of Risk and Uncertainty* **5**(3): 247–251.
- Huang, R.J., Liu, Y. and Tzeng, L.Y. (2007) *Hidden overconfidence and advantageous selection*, Working paper.
- Huang, R.J., Muermann, A. and Tzeng, L.Y. (2008) *Hidden regret in insurance markets: Adverse and advantageous selection*, Working paper.
- Jullien, B., Salanie, B. and Salanie, F. (1999) 'Should more risk-averse agents exert more effort?' *Geneva Papers on Risk and Insurance Theory* **24**(1): 19–28.
- Koufopoulos, K. (2005) *Asymmetric information, heterogeneity in risk perceptions and insurance: An explanation to a puzzle*, Working paper.
- Rothschild, M. and Stiglitz, J.E. (1976) 'Equilibrium in competitive insurance markets: An essay on the economics of imperfect information', *Quarterly Journal of Economics* **90**(4): 629–649.
- Sonnenholzner, M. and Wambach, A. (2008) *On the role of patience in an insurance market with asymmetric information*, Working paper.

## About the Authors

**Kili C. Wang** Associate Professor at the Department of Insurance of Tamkang University, Taiwan

**Rachel J. Huang** is Associate Professor at the Department of Finance of Yuan University, Taiwan.

**Larry Y. Tzeng** is Professor at the Department of Finance of National Taiwan University, Taiwan.