The Geneva Risk and Insurance Review, **2015**, **40**, (102–132)

www.palgrave-journals.com/grir/

# Information Problems in Bancassurance: Empirical Evidence Based on a Comparison Between Over-the-Counter and Telephone Marketing Customers

Jin-Lung Peng[a,b,c] and Kili C. Wang[b,c,d]
[a]Department of Risk Management and Insurance, National Chengchi University, Taipei, Taiwan.
[b]Risk and Insurance Research Center, College of Commerce, National Chengchi University, Taipei, Taiwan.
E-mail: kili@mail.tku.edu.tw
[c]Cardif Bancassurance Research Center, College of Commerce, National Chengchi University, Taipei, Taiwan.
E-mail: kili@mail.tku.edu.tw
[d]Department of Insurance, Tamkang University, Taipei, Taiwan.
E-mail: kili@mail.tku.edu.tw

We investigate the problem of asymmetric information among the different distribution channels of bancassurance, namely, over-the-counter (OTC) selling and telephone marketing (TM). We predict that the bank could bring an information advantage to insurance companies through TM by providing a customer list based on the customers' records in the bank. Bancassurance can play a role of integrating the information that is transferred from the bank to the insurance company. Our empirical evidence shows that, in contrast to OTC customers, there is less evidence of adverse selection, and there is also less evidence of moral hazard or fraud among the TM customers that have been sorted by the bank. This phenomenon could be attributed to the valuable private information provided by the credit records of customers kept by the bank.
*The Geneva Risk and Insurance Review* (2015) **40**, 102–132. doi:10.1057/grir.2014.8

**Keywords:** bancassurance; information asymmetry; distribution channel

Article submitted 27 August 2013; revised 12 May 2014; published online 10 March 2015

## Introduction

Banks nowadays are becoming increasingly important distribution channels for life insurance companies. The highest market share for bancassurance is that in Brazil,[1]

[1] Bancassurance results from the integration that takes place between banks and insurance companies. It can exist in one organisation that performs both the business of a bank and the business of an insurance company, or else in two independent organisations in which only some of the business is integrated.

where it reaches more than 80 per cent. In some European countries, the market share of the bank channel exceeds 60 per cent.[2] In Taiwan, it also reaches 60 per cent, while in some neighbouring countries, such as South Korea and Hong Kong, the market shares accounted for by bank channels are more than half. Many studies in the literature indicate that the insurers can benefit from integrating their sales and marketing operations with those of banks from the point of view of geographic diversification and public trust.[3] The banking system not only has a high density of operating bases through its extended branches, but also enjoys the public's higher trust. This enables the insurance company to benefit from distributing insurance contracts through bank channels.

In reality, there are different ways of marketing insurance contracts via bank channels. The problems associated with information asymmetry in relation to those insurance contracts may also be lessened through different approaches to marketing. To the best of our knowledge, no one ever considers the benefits from distributing insurance contracts through bank channels from the point of view of information. In this study, we investigate the problems of asymmetric information among the contracts within one bancassurance company in Taiwan.

Our target bancassurance company is not an organisation with the business of a bank, but is instead an organisation that is highly horizontally integrated with the bank in terms of its distribution. Two approaches to marketing are used by this bancassurance company. One is where the customers take the initiative to purchase the insurance products over the counter of the bank; we refer to these customers as "over-the-counter (OTC) customers". The other is where the bank generates a customer list for our target insurance company to solicit insurance products through telephone marketing; we refer to these customers as telephone marketing (TM) customers. The information seems to be much more asymmetric among OTC customers than among TM customers. As for OTC customers, a higher percentage of insureds purchase contracts with high coverage. Furthermore, those who are covered by high-coverage contracts as OTC customers tend to have a higher probability of filing a claim than those with high-coverage contracts who are TM customers.[4] The different degrees of asymmetric information between these two

[2] Data are obtained from Financial Globe Bancassurance analysis. The percentages in terms of market shares accounted for by for recent years (2010, 2011 or 2012) are available from each respective country.

[3] Hoschka (1994); Genetay and Molyneux (1998); and Staikouras and Nurullah (2005) mentioned that banks own public trust, and this can benefit the insurance companies in selling their products. Rumelt (1974) and Fields *et al* (2007) pointed out that integration between banks and insurance companies can enable them to benefit from the economies of scope in terms of the profile diversification on the balance sheet or geographic diversification.

[4] We can observe these in Tables 3 and 4. We also provide a more detailed description in the section "Data and Methodology".

groups may arise for two reasons. One relates to the information niche provided by the bank. The other is due to the difference, between these two groups, in terms of who initiates the transaction.

Through the information niche provided by the bank, TM customers with better credit scores are selected, while OTC customers cannot be selected. TM customers tend to be more responsible, to exercise more self-restraint, to take less risk, and to face a higher moral cost while under pressure to maintain their good credit record. Accordingly, TM customers are the insureds with a comparatively lower risk-type, more risk aversion, less *ex ante* moral hazard, or less fraud than OTC customers.

As for the initiation of the transaction, it is the insurer who initiates the transaction in the TM group, whereas it is the customer who initiates the transaction in the OTC group. It is often the person who holds private information regarding the high-risk type,[5] or the person who intends to take advantage of the insurance company by means of some kind of manipulation,[6] who has the incentive to go to the bank to purchase insurance, and especially to purchase higher coverage over the counter. Hence, we can believe that there is more private information, and accordingly, a more severe asymmetric information problem among OTC customers than among TM customers.

This paper investigates the asymmetric information problem among the personal accident insurance contracts and the hospital indemnity rider of our target bancassurance company. When we compare the asymmetric information problems between OTC customers and TM customers, the significantly positive conditional correlation between coverage and claims, for the main contract or its rider,[7] could be evidence of adverse selection, *ex ante*/*ex post* moral hazard or fraud. These factors are not easy to distinguish from each other in empirical work. However, we attempt to distinguish them by identifying whether the filed claim is easily diagnosed or not when the hospital expense is indemnified by the rider. Dionne and St-Michel[8] mentioned that, with the help of judging the reason for the claim, researchers could have the opportunity to identify the moral hazard problem. When the coverage rises, if the increment of the indemnity is consistent for different claim reasons, we cannot infer that this increased indemnity is due to moral hazard. However, on the contrary,

[5] For instance, the living/working environment of the individual is private information held by the insured. The insurer does not use this information when underwriting or pricing. However, the individual who lives or works in a highly hazardous environment will be more willing to purchase an insurance contract and will also have a higher probability of having an accident.

[6] For example, there could be an individual who actually gets injured and buys the insurance contract after the injury happens. This is also the insured's private information that cannot be observed by the insurer. This constitutes an extreme example of fraud.

[7] The insured in our target personal accident insurance contract has the option to further purchase different amounts of the hospital indemnity rider or not. We will provide the details in the next section.

[8] Dionne and St-Michel (1991).

if the increment of the indemnity is only particularly high when the claim reason is not easy to diagnose, we can confirm that this increased indemnity is due to moral hazard. Based on a similar concept, we examine the hospital indemnity rider in our contract. If it is adverse selection, a high-risk individual will tend to purchase a high-coverage rider, and his/her claim probability should be higher regardless of whether the claim is easy to diagnose or not. However, if it is *ex post* moral hazard or manipulation, the individuals with higher coverage could have more chances to take advantage of this insurance contract when the claim is not easy to diagnose.

We find that not only is there less positive conditional correlation between coverage and claim in both the main contract and its rider among the insureds solicited through TM, but there is also the phenomenon that the positive conditional correlation between the coverage of the hospital indemnity rider and claim when the claims are not easy to diagnose is less severe among the insureds solicited through TM. These phenomena could be attributed to the fact that the credit record information is also related to the information regarding the individual's risk type, cautious behaviour or moral behaviour; they could also contribute to the differentiated asymmetric information phenomena brought by whether the trade is initiated by the insurer or the insured. Hence, horizontal integration of sales between the bank and insurance company brings an opportunity to reduce the asymmetric information problems confronted by the insurance company. The insurance company can benefit from distributing the insurance contract according to the sorted customers' list provided by the bank.

Following this Introduction, which briefly profiles the motivation, contribution and important outcomes of this paper, the remainder of this paper is organised as follows. In the second part, we introduce the background to our target insurance company and research contract. According to the practice of Bancassurance and the predictions in the literature, we establish our hypotheses. The third part describes the data and methodology. The results of our empirical research are presented in the fourth part. The final part is the conclusion.

## Background and hypotheses

Our target insurance company is the Taiwan branch of a leading bancassurer from Europe.[9] The branch in Taiwan sells the insurance products only through bank channels. Except for those customers who take the initiative to purchase the

[9] Our target insurance company's parent company was founded in the 1970s and established its subsidiaries and branches in 37 countries across Europe, Asia and Latin America. It set up a branch in Taiwan in 1996 and became the first and largest professional bancassurer in this market. In 2012, this Taiwan branch was the largest player providing unit-linked and credit life insurance services in Taiwan's insurance market.

insurance products over the counter of a bank (OTC), banks also generate a customer list for our target insurance company to solicit insurance products through TM.

On the bank side, the information regarding the customer's type is revealed through the practices of the bank, such as deposits, loans, credit cards and different kinds of wealth management.[10] When the insurance company distributes its contracts through bank channels, the insurance company cannot directly access this information related to each individual's credit score or wealth situation without the customer's consent. This is prohibited by the laws and regulations.[11] However, the bank can provide a "suggestion list" for the insurance company to solicit its products. In our research sample, the customer list provided by the bank to solicit insurance contracts through TM is mainly constituted by its credit card holders who have better credit scores.[12] In Taiwan, while the default rate on bank loans is about 0.4 per cent, the default rate on credit cards is only about 0.047 per cent. The credit scores of banks' credit card customers are generally higher on average already, not to mention that the bank has also further sorted out some bad record credit card customers. Hence, we can believe that, compared with the average customers in the bank, these customers who are on the list provided by the bank are customers with better scores. The agent of our target insurance company will call the customers on this list and promote the company's products.

By contrast, OTC customers are randomly selected from the population. They are either not frequent customers and just transact the insurance contracts over the counter of the bank, or they are frequent customers but their credit information on the bank side cannot be used for underwriting or pricing the insurance contracts.[13] Accordingly, bank agents may not have the incentive,[14] and are also not allowed to be asked to select OTC customers. Therefore, we can believe that, compared with OTC customers, the TM customers that are selected have better credit scores.

As mentioned in the previous section, the asymmetric information problems seem to be more severe in relation to OTC customers than to TM customers. We will start with the adverse selection problem. Two possible reasons bring about the phenomenon that OTC insureds could have a comparatively more adverse selection

[10] Bergendal (1995) mentioned that banks have exclusive access to the payment system and a special right to provide customers with deposits. Hence, the bank exclusively knows its customers' credit scores.

[11] This is regulated by the Personal Information Protection Act and Financial Holding Company Act of Taiwan. Moreover, we can find many detailed regulations to protect financial customers' transaction information and privacy.

[12] About 90 per cent of customers on the list consist of credit card customers in each year.

[13] The Financial Holding Company Act regulates the direct usage of transaction information from the bank side by the allied insurance company.

[14] Bank agents will generally focus on their own business instead of performing underwriting.

problem. Firstly, TM customers are selected with better credit scores, while OTC customers cannot be selected and their credit scores are randomly distributed. The individuals with better credit scores tend to be more responsible, to exercise more self-restraint and to be persons who take less risk.[15] Accordingly, TM customers are lower risk-type insureds or more risk-averse than OTC insureds, while the credit score is strongly related to the individual's risk type and risk-averse. When the credit score is prohibited from being explicitly used in underwriting or pricing, selecting better risk-type or more risk-averse individuals through the private information revealed by credit scores can help the insurance company mitigate the information inferiority problem. Hence, we can believe that OTC insureds give rise to more of a private information problem when they cannot be selected, and they face more of a problem of adverse selection than TM insureds.

Secondly, it is the customer who initiates the transaction in the OTC channel, whereas it is the insurance company that initiates the transaction in the TM channel. Based on what was mentioned in the previous section, the individuals who hold more private information and perceive that they belong to the high-risk type tend to exhibit more urgency when initiating the transaction.[16] Therefore, we believe that there is a more severe adverse selection problem among OTC customers than among TM customers.

For the above two reasons, we can believe and build our first hypothesis as follows:

*H1:* By taking OTC customers as the control group, the customers selected by TM exhibit a weaker problem in adverse selection.

Furthermore, the customers with better credit scores are under pressure to keep their good credit record. They face higher costs if they engage in *ex post* moral hazard or commit fraud. Therefore, they are more careful when dealing with their policies

[15] An individual's credit score can reveal his/her attitude of responsibility. Brockett and Golden (2007) point out that certain biological, psychological and behavioural characteristics of individuals could, in conjunction with each other, influence those individuals' risk-taking behaviours from a financial point of view as well as the degree of their risk-taking in other behaviour such as driving. Lastovicka *et al.* (1987); Arnett (1990); Donovan (1993) and McMillen *et al.* (1992) also argued that individuals with certain psychological characteristics, such as sensation, thrill or adventure seeking, lack of inhibition, boredom susceptibility and egocentrism, tend to exhibit certain problem behaviours, such as taking drugs, or being delinquent, hostile, aggressive or unconventional, and also tend to adopt risky driving habits (drunk driving, speeding, traffic violations, etc.), hence resulting in more accidents and loss. Hence, they infer a relationship between insured auto losses and credit scores.

[16] Of course, we cannot preclude that the individual who owns the private information regarding more risk aversion would also tend to initiate the transaction. Furthermore, those more risk-averse individuals are characterised by a lower *ex post* risk probability. Hence, as to which type of private information is stronger when the individual is travelling to purchase insurance over the counter still awaits practical evidence from our empirical test.

and have no incentive to file unnecessary claims.[17] By contrast, the ones who intend to take advantage of the insurance company by manipulating claims have the incentive to initiate the trading of the insurance contracts over the counter. Hence, TM customers face higher costs associated with engaging in *ex post* moral hazard or fraud than OTC customers. Accordingly, we can establish our second hypothesis as follows:

*H2:* In contrast to the OTC group, TM customers would be less likely to be associated with *ex post* moral hazard and less likely to engage in fraudulent behaviour through such activities as filing an unnecessary claim when the accident actually happened or making false claims.

In this paper, we investigate the policies of the personal accident insurance contract and its hospital indemnity rider based on our target insurance company. The reasons why we focus our attention on this line are two-fold. First, the insurance company only uses the individual's occupational categories when underwriting this contract.[18] However, substantial heterogeneity among insureds still exists and needs to be further explored. Except for the common factors related to the individual's characteristics, the credit score used by TM for our target bancassurer provides a good opportunity and direction to proceed with an investigation.

Second, this line of insurance can also be extended to the coverage that includes hospital expenses. It is well known that wasted medical expenses, including the

[17] Furthermore, we also believe that the TM customers who have better credit scores also have less of a tendency to engage in *ex post* moral hazard or fraudulent behaviours. It is also widely accepted in the literature that, whether or not an individual would act immorally or defraud, mostly depends on whether he/she is responsible or honest. Duffield and Grabosky (2001) pointed out that the psychological explanation of fraud includes greed and dishonesty. Mazar *et al.* (2008); Langer (1989) and Duval and Wicklund (1972) agreed that people who pay more attention to their own moral standards, or are more mindful of themselves, tend to adhere to stricter standards of honesty. Lammers and Shiller (2010) argued that fraudulent individuals are selfish and amoral. Matching the above-mentioned points with the point raised by Brockett and Golden (2007) leads us to believe that the individuals with better credit scores are mostly honest and self-restrained persons. We can also believe that the insured's moral hazard or fraudulent behaviour could be related to his/her credit score.

[18] There are six categories of occupations in the personal accident insurance market. From the first category to the sixth category, the risk associated with an occupation is ranked from lowest to highest. The first category includes in-house operation workers, such as administrators, the professions, agents, etc. The second category includes outdoor workers who are not engaged in highly dangerous work, such as farmers, deliverymen, drivers, porters, etc. The third category includes skilled workers, such as technical personnel, postmen, telecommunications workers, etc. The fourth category includes operators at work sites, such as fishermen, traffic policemen, security service personnel, lifeguards, etc. The fifth category includes workers who are engaged in more dangerous work, such as timber porters, supervisors of lumbermen, captains, etc. The sixth category includes workers who are engaged in more dangerous work than the fifth category, such as lumbermen, sailors, flight crew, pilots, etc.

abuse of unnecessary hospital expenses, is a serious problem in insurance. Although insureds may have less of a tendency to have accidents on purpose, they may be more likely to overuse the insurance or abuse the unnecessary hospital expenses after the accident happened under the extended coverage of the hospital indemnity rider. Even more, there is the possibility of making false claims if it is hard for a diagnosis to be made. This *ex post* moral hazard, or fraud problem, is a broad concern in medical-related insurance. Hence, it is particularly meaningful to try to identify some personality characteristics that are related to the *ex post* moral hazard or fraudulent behaviour of the insured.

The personal accident insurance contract provides coverage only for the event of death or disability. Insureds who take out such personal accident insurance have the option to buy a rider to extend the coverage to the hospital expenses associated with an accident. The options for the amount insured under the personal accident policy include NT$0.5, NT$1, NT$1.5, NT$2, NT$3, NT$4, NT$5 and NT$6m.[19] In addition, the options for the payment of the hospital indemnity rider include: NT$1,000, NT$2,000, or NT$3,000 per day. The policyholder can choose any such combination in terms of the amount insured between this personal accident insurance and its hospital indemnity rider as he or she may desire. In our research data, we find that the coverage choice of the main contract and the coverage choice of the rider are only slightly correlated. The correlation coefficient for choosing high coverage for the main contract and choosing high coverage for the rider is only 0.03 and is insignificantly different from 0.[20] When we further analyse the factors that affect an individual who would further extend his/her main contract to the rider, we find that whether the insured faces more risk in filing the personal accident claim is not a significant factor. Occupation, age, gender and marital status are significant factors: insureds who are engaged in more highly dangerous work, males, the unmarried and elderly are more likely to include the rider in their contracts.[21] In the later analysis, we separately test the asymmetric information problem in relation to the main contract and the rider, since the coverage choices of these two contracts are not only almost uncorrelated and their choices are also uncorrelated with the individual's risk type.

According to the prediction of adverse selection, the high-risk-type insured tends to purchase a high-coverage insurance contract. Hence, there could be a positive

[19] The average NT$:US$ exchange rate during the research period is NT$30 = US$1.

[20] The high coverage of the personal accident insurance contract is defined by an insured amount that exceeds NT$2m. High coverage of the hospital indemnity rider is defined by an insured amount that is above NT$1,000. These will be described in detail later.

[21] This is observed by regressing the coverage extending choice on the personal accident claim, occupation and the characteristics of the insured. These results are also consistent with the observations from the descriptive statistics in the next section. The regression results are available upon request.

correlation emerging from the empirical evidence between the frequency of personal accident claims and the probability of choosing high-coverage personal accident insurance, and a positive correlation between the frequency of hospital expense claims and the probability of choosing a high hospital indemnity rider. According to the prediction of *ex ante* moral hazard, the insured who is covered by a high-coverage insurance contract tends to be less cautious, and the probability of having an accident accordingly rises. Hence, a positive correlation could also exist between the frequency of personal accident claims and the probability of choosing high-coverage personal accident insurance as well as between the frequency of hospital expense claims and the probability of choosing a high hospital indemnity rider. The insured covered by more coverage could also tend to file additional claims when the accident actually happens if there is *ex post* moral hazard. For example, a higher rider could induce the insured who had actually had an accident but did not need to stay in hospital to file the unnecessary hospital expense. Accordingly, the positive correlation between the frequency of hospital expense claims and the probability of choosing a high hospital indemnity rider could also rise where there is *ex post* moral hazard.[22] Moreover, some opportunistic insured who are covered by a high-coverage insurance contract have the incentive to defraud by making a false claim to take advantage of the insurance company.

As mentioned above, for both the main contract and the rider, empirical evidence showing that a high-coverage insurance contract is positively correlated with the high frequency of a claim could not be construed as being due to adverse selection only, but could probably be also due to *ex ante* moral hazard, *ex post* moral hazard or fraud.

However, we have opportunities to identify them in this research. First, when we explore the asymmetric information problem in the main personal accident insurance contract, if we only focus on the claim arising due to the event of death or complete disability,[23] the higher probability of a claim accompanied by a high-coverage contract is less likely to be due to *ex ante*/*ex post* moral hazard or fraud. It is hard to feel convinced, due to the motivation of *ex ante* moral hazard, that people would intend to be less cautious as a result of the fatal hazard under the high-coverage insurance contract, or that, due to the motivation of *ex post* moral hazard or fraud, people would be willing to sacrifice their life or become more seriously physically disabled by manipulating the severity of an accident or fabricating a fatal accident under the

[22] Some studies also discuss the *ex post* moral hazard based on the relationship between the claim amount and contract coverage when the accident actually happened. In this research, we can only discuss the "unnecessary claim" instead of "the amount of the claim" when the accident actually happened. This is because the indemnity of the claim amount (hospital days) is fixed by the contract according to the different kinds of accident, instead of being filed by the insured.

[23] Our target personal accident insurance also indemnifies disabilities other than complete disability. However, for the purpose of identifying adverse selection, we only focus on complete disability and drop the observations for which disabilities other than complete disability have been filed.

inducement of high insurance coverage. Hence, this positive correlation is more likely to be evidence of adverse selection than of *ex ante*/*ex post* moral hazard or fraud.

Secondly, when we investigate the hospital indemnity rider, although the *ex post* moral hazard or fraud could be combined with *ex ante* moral hazard or adverse selection, the opportunity to identify them still exists, because the insurance company will also audit the claims to prevent the occurrence of *ex post* moral hazard or fraud. The probability of successfully spotting *ex post* moral hazard or fraud varies with the probability of detecting the truthfulness of an accident. Hence, the probability of *ex post* moral hazard or fraud is higher when the event of a claim is more difficult to verify. If there is *ex post* moral hazard or fraud, the correlation between the probability of choosing a high-coverage rider and the frequency of a hospital expense claim should be positive and particularly higher when the reason for the accident is hard to diagnose. If there is only *ex ante* moral hazard or adverse selection, the positive correlation should be consistent for different reasons for the accident.

There are two main reasons that are given by insureds when they file a claim. One is that benign paroxysmal positional vertigo (BPPV) resulted in the accident, and the other is that the accident was associated with trauma. BPPV is a symptom whereby the patient experiences dizziness, which is not caused by a noxious or virulent disease. By contrast, trauma is very easy to observe based on appearance. The former is apparently more difficult to diagnose than the latter. We can treat the former as a reason that is not easy to diagnose, and we refer to such claims as "ambiguous claims". By identifying a positive correlation between high coverage and ambiguous claims instead of all the reasons for the claims, we can identify *ex post* moral hazard or fraud.

## Data and methodology

Our empirical data come from the target insurance company, the largest life professional bancassurer in Taiwan, and its distribution totally depends on bank channels. We investigate the personal accident insurance policy and its rider, which covers the hospital expenses associated with the accident.

The length of this accident insurance policy is one year for each contract period and the policy can be repeatedly purchased. We collected the policies sold during the period from the year 2000 to the year 2010. The number of contracts in each year is displayed in Table 1. For each contract, we observe the claim record for the full contract year. In other words, our data period is in the form of a policy year instead of a calendar year. Hence, our research sample consists of panel data which are made up of insureds covered by the personal accident insurance contract of this insurance company from the year 2001 to the year 2009. There are 375,399 policies for personal accident insurance during this period. We can see the coverage structure of this database in Table 1. About 71.2 per cent of these personal accident insurance policies are also extended to hospital expense riders. Hence, we have 267,276 policies which are the contracts that are also covered by hospital expense insurance.

**Table 1** The coverage structure and year structure in our research sample

| *Personal accident insurance* | | *Hospital expense rider* | |
|---|---|---|---|
| *Structure of coverage* | | | |
| *cov_hosp* | 0.7120 | 1,000 | 0.8184 |
| 500,000 | 0.0125 | 2,000 | 0.1785 |
| 1,000,000 | 0.3291 | 3,000 | 0.0031 |
| 1,500,000 | 0.0010 | | |
| 2,000,000 | 0.4936 | | |
| 3,000,000 | 0.1606 | | |
| 4,000,000 | 0.0001 | | |
| 5,000,000 | 0.0025 | | |
| 6,000,000 | 0.0006 | | |
| *observations* | 37,5399 | *observations* | 26,7276 |
| *Structure of year* | | | |
| Year 2000 | 0.0265 | Year 2000 | 0.0135 |
| Year 2001 | 0.0864 | Year 2001 | 0.0765 |
| Year 2002 | 0.1139 | Year 2002 | 0.1146 |
| Year 2003 | 0.1561 | Year 2003 | 0.1530 |
| Year 2004 | 0.1620 | Year 2004 | 0.1731 |
| Year 2005 | 0.1257 | Year 2005 | 0.1282 |
| Year 2006 | 0.0950 | Year 2006 | 0.0959 |
| Year 2007 | 0.0775 | Year 2007 | 0.0789 |
| Year 2008 | 0.0622 | Year 2008 | 0.0646 |
| Year 2009 | 0.0509 | Year 2009 | 0.0546 |
| Year 2010 | 0.0438 | Year 2010 | 0.0471 |

*Notes*: (1) *cov_hosp* means the contract also extends to the endorsement of hospital expenses.
(2) There are eight levels of the insurance amount for the personal accident insurance contract. They are NT $0.5, NT$1, NT$1.5, NT$2, NT$3, NT$4, NT$5 and NT$6m.
(3) There are three levels of hospital expense indemnity per day for the hospital expense endorsement. They are NT$1,000, NT$2,000 or NT$3,000 per day.

We can also see that the insurance amount for this personal accident insurance contract is based on choosing an amount between NT$0.5m and NT$6m. Most insureds choose either NT$1m or NT$2m as their insurance amount. These amounts account for 32.91 and 49.36 per cent of our research sample, respectively. In our later empirical work, we define those contracts that have insurance amounts of over NT$2m as high-coverage contracts in personal accident insurance. Hence, for the personal accident insurance in our sample, 16.39 per cent of the contracts are high-coverage contracts.[24]

[24] We also alternatively treat NT$1m as the threshold, and define those contracts for which the amounts exceed NT$1m as high-coverage contracts and perform a robustness test in later empirical work. In this alternative coverage definition, 65.84 per cent of the contracts are high-coverage contracts.

The riders for hospital expenses are chosen based on three levels: 81.84 per cent of the insureds who also purchased this rider chose NT$1,000 of hospital expense indemnity per day, 17.85 per cent of them chose NT$2,000 per day, and only 0.31 per cent of them chose NT$3,000 per day. We define the contracts that include this rider of over NT$1,000 per day as high-coverage contracts in the hospital expense insurance rider. Hence, in our research sample, for these hospital expense riders, 18.16 per cent of these contracts are high coverage.[25]

In Table 2, we list the statistical descriptions for our overall research sample of personal accident insurance contracts. We find that about 76.54 per cent of the policies are sold through TM. This implies that the marketing niche is properly used by this insurance company. The claim rate on this personal accident insurance contract is about 1.04 per cent.

The average age of the insureds in our sample is about 37 years old. Almost half of them are middle-aged. However, the proportion of female and married persons is a little lower in our sample than in the overall population. The insureds are much more concentrated in cities (67.7 per cent), or in northern Taiwan (51.14 per cent) than the general population. The occupations of the insureds in our research sample are highly concentrated in the first and second categories, with the respective proportions being 76.34 and 12.30 per cent. These two categories of occupations are the lowest risk categories in different occupations. This could reflect the fact that this insurance company has a comparatively moderate underwriting strategy.

We also list in Table 2 the statistical descriptions for the sub-sample comprising those insureds who had further purchased the hospital expense rider. In all, 74.53 per cent of these contracts are sold through TM, which is very consistent with the proportion for the overall sample. The probability of filing a hospital expense claim is 5.78 per cent, which is higher than the percentage of claims for personal accident insurance contracts. Furthermore, 12.56 per cent of those claimed riders are ambiguous claims.[26] The structure regarding the insured's living location is similar to those structures for the overall sample. The explicit differences are that the average age of the insureds in the sub-sample of this hospital expense rider is slightly higher than that in the full sample. There are fewer female insureds, and there are more insureds that are not married. They are also more concentrated in higher-risk categories of occupations. In other words, the insureds who also extend their coverage to include the hospital expense rider are more concentrated among the female, unmarried and elderly insured and they are also engaged in occupations with higher risk.

[25] We do not switch the threshold and perform a robustness test for the hospital indemnity rider. This is because, if we switch this threshold to NT$2,000 per day, only 0.31 per cent of the contracts can be classified as high-coverage contracts. This robustness test will not lead to meaningful results.

[26] In the subsample, 5.78 per cent of the hospital indemnity riders had filed a claim, and 0.726 per cent of the hospital indemnity riders had filed ambiguous claims. Hence, about 12.56 per cent of all claim riders are ambiguous.

**Table 2** Variable definitions and statistical descriptions

| *Variable* | *Definition* | *Whole sample (all personal accident insurance contracts)* | | *Sub-sample (the contracts with hospital indemnity riders)* | |
|---|---|---|---|---|---|
| | | *Mean* | *Std. dev.* | *Mean* | *Std. dev.* |
| *Explained variable:* | | | | | |
| *claim* | A dummy variable that equals 1 when an insured has filed the claim during the corresponding contract period, otherwise 0. | 0.0104 | 0.1013 | 0.0578 | 0.2334 |
| *ambg* | A dummy variable that equals 1 when an insured has filed the claim and its reason is not easy to be diagnosed during the corresponding contract period, otherwise 0. | | | 0.0073 | 0.0849 |
| *cov_high* | A dummy variable that equals 1 when the amount of the insurance contract is more than NT$2m, otherwise 0. | 0.1639 | 0.3701 | 0.1816 | 0.3855 |
| *Explanatory variables:* | | | | | |
| *sexf* | A dummy variable that equals 1 when the insured is female, otherwise 0. | 0.3972 | 0.4893 | 0.3712 | 0.4831 |
| *married* | A dummy variable that equals 1 when the insured is married, otherwise 0.[a] | 0.4626 | 0.4986 | 0.4270 | 0.4946 |
| *age* | The variable which represents the insured's age. | 36.9066 | 10.2405 | 38.2260 | 10.8583 |
| *occp_1* | A dummy variable that equals 1 when the occupation of the insured is in the first category, otherwise 0. | 0.7634 | 0.4250 | 0.7462 | 0.4352 |
| *occp_2* | A dummy variable that equals 1 when the occupation of the insured is in the second category, otherwise 0. | 0.1230 | 0.3285 | 0.1283 | 0.3344 |
| *occp_3* | A dummy variable that equals 1 when the occupation of the insured is in the third category, otherwise 0. | 0.0489 | 0.2156 | 0.0535 | 0.2250 |
| *occp_4* | A dummy variable that equals 1 when the occupation of the insured is in the fourth category, otherwise 0. | 0.0500 | 0.2180 | 0.0550 | 0.2281 |
| *occp_5* | A dummy variable that equals 1 when the occupation of the insured is in the fifth category, otherwise 0.[b] | 0.0067 | 0.0816 | 0.0079 | 0.0885 |
| *city* | A dummy variable that equals 1 when the insured lives in a city, otherwise 0[c] | 0.6770 | 0.4676 | 0.6598 | 0.4738 |
| *north* | A dummy variable that equals 1 when the insured lives in the northern area of Taiwan, otherwise 0. | 0.5114 | 0.4999 | 0.5275 | 0.4992 |
| *south* | A dummy variable that equals 1 when the insured lives in the southern area of Taiwan. | 0.2592 | 0.4382 | 0.2331 | 0.4228 |
| *central* | A dummy variable that equals 1 when the insured lives in the central area of Taiwan.[d] | 0.1509 | 0.3579 | 0.1440 | 0.3511 |
| *teleph* | A dummy variable that equals 1 when the insurance contract is sold through a telephone marketing channel, otherwise 0.[e] | 0.7654 | 0.4237 | 0.7453 | 0.4357 |
| *Observations* | | 375,399 | | 267,276 | |

[a]The control group of *married* includes all the insureds who are not in the status of marriage.

[b]The control group for *occp_1, occp_2,…occp_5* consists of the insureds whose occupation is in the 6$^{th}$ category.

[c]The control group consists of the owners of insured vehicles who do not live in cities in Taiwan.

[d]The insureds in the control group for *north, south* and *central* comprise those insureds who live in east area and some inland areas of Taiwan.

[e]The control group for *teleph* includes all the insureds who are from the OTC channel.

For a further preliminary observation of the insureds who were solicited by the TM channel vs those who were contracted through the OTC channel, we further separately display the descriptive statistics for these two groups of insureds for the whole sample and for the sub-sample in Table 3. We first observe the structure of insureds. The most explicit difference between these two groups of insureds is that, for the whole sample, the average age of OTC insureds is higher than that of TM insureds.

Secondly, the most important observation is focused on the claim rate and the percentage of high-coverage contracts among these two groups of insureds. For the whole sample, of the personal accident insurance contracts, the percentage of high-coverage contracts for OTC insureds is 18.93 per cent, whereas the percentage of high-coverage contracts for TM insureds is 15.6 per cent. It is thus higher in the OTC insured group. The claim rate is 1.05 per cent for OTC insureds, and 1.03 per cent for TM insureds. It is also slightly higher in the OTC insured group. In the sub-sample comprising the hospital indemnity riders, the percentage of high-coverage contracts for OTC insureds is 66.41 per cent, while the percentage of high-coverage contracts for TM insureds is only 1.67 per cent. This shows that it is significantly higher for the OTC insured group. The claim rate is 6.58 per cent for OTC insureds and 5.51 per cent for TM insureds. Furthermore, the ambiguous claim rate is 1.03 per cent for OTC insureds, and 0.63 per cent for TM insureds. The claim rate is higher and the ambiguous claim rate is even higher for the OTC insured group.

These observations seem to imply that OTC insureds tend to have a high claim rate as well as more high-coverage contracts. According to our hypotheses, we also expect that the OTC insureds who purchased high-coverage contracts tend to have a higher probability of filing a claim than the insureds who purchased low coverage contracts. Before we perform a formal test, we further make a preliminary observation on the claim frequencies based on a 2×2 correlation between the TM/OTC marketing status and coverage level in Table 4.[27] These preliminary observations show that, for the personal accident insurance contract, OTC insureds covered by a high (low) coverage contract tend to have a higher (lower) personal accident claim frequency than TM insureds; for the hospital indemnity rider, OTC insureds covered by a high (low) amount of coverage tend to have a higher (lower) ambiguous claim frequency than TM insureds. This implies that there is less severe adverse selection, *ex post* moral hazard, or fraud among TM insureds.

[27] The personal accident claim frequency of the OTC insured who is covered by a high (low) coverage contract is 0.0153 (0.0094), whereas it is 0.0124 (0.0100) for the TM insured who is covered by a high (low) coverage contract. The hospital expense claim frequency of the OTC insured who is covered by a high (low) coverage contract is 0.0721 (0.0642), whereas it is 0.0608 (0.0534) for the TM insured who is covered by a high (low) coverage contract. The frequency of an ambiguous claim involving hospital expenses is 0.0089 (0.0069) for OTC insureds who are covered by a high (low) coverage contract, and it is 0.0079 (0.0070) for TM insureds who are covered by a high-coverage contract.

**Table 3** Statistical descriptions for TM insureds vs OTC insureds

| | *The whole sample (all personal accident insurance contracts)* | | | | *The sub-sample (the contracts with hospital indemnity riders)* | | | |
|---|---|---|---|---|---|---|---|---|
| | *TM insureds* | | *OTC insureds* | | *TM insureds* | | *OTC insureds* | |
| | *Mean* | *Std. dev.* | *Mean* | *Std. dev.* | *Mean* | *Std. dev.* | *Mean* | *Std. dev.* |
| *claim* | 0.0103 | 0.1011 | 0.0105 | 0.1019 | 0.0551 | 0.2281 | 0.0658 | 0.2479 |
| *ambg* | | | | | 0.0063 | 0.0647 | 0.0103 | 0.0954 |
| *cov_high* | 0.1560 | 0.3629 | 0.1893 | 0.3918 | 0.0167 | 0.1282 | 0.6641 | 0.4723 |
| *sexf* | 0.4051 | 0.4909 | 0.3713 | 0.4832 | 0.3675 | 0.4821 | 0.3818 | 0.4858 |
| *married* | 0.3910 | 0.4880 | 0.6960 | 0.4600 | 0.3548 | 0.4785 | 0.6382 | 0.4805 |
| *age* | 36.5360 | 8.8383 | 38.1157 | 13.7933 | 38.0855 | 9.4521 | 38.6373 | 14.1869 |
| *occp_1* | 0.7159 | 0.4510 | 0.9185 | 0.2736 | 0.6849 | 0.4646 | 0.9256 | 0.2624 |
| *occp_2* | 0.1510 | 0.3581 | 0.0316 | 0.1749 | 0.1633 | 0.3697 | 0.0257 | 0.1582 |
| *occp_3* | 0.0586 | 0.2348 | 0.0172 | 0.1298 | 0.0662 | 0.2486 | 0.0163 | 0.1264 |
| *occp_4* | 0.0576 | 0.2329 | 0.0253 | 0.1572 | 0.0655 | 0.2475 | 0.0243 | 0.1541 |
| *occp_5* | 0.0072 | 0.0845 | 0.0051 | 0.0715 | 0.0085 | 0.0920 | 0.0060 | 0.0774 |
| *city* | 0.6813 | 0.4660 | 0.6630 | 0.4727 | 0.6644 | 0.4722 | 0.6465 | 0.4780 |
| *north* | 0.4778 | 0.4995 | 0.6210 | 0.4852 | 0.4966 | 0.5000 | 0.6179 | 0.4859 |
| *south* | 0.3067 | 0.4611 | 0.1041 | 0.3054 | 0.2882 | 0.4529 | 0.0717 | 0.2579 |
| *central* | 0.1748 | 0.3798 | 0.0728 | 0.2597 | 0.1728 | 0.3781 | 0.0596 | 0.2368 |

The formal judgement regarding asymmetric information is based on the conditional correlation between contract coverage and claim. The methodology we adopt to test the conditional correlation is the two-stage method.[28] In the first stage, we estimate the probability of a claim by means of a probit regression:

$$\text{Prob}(claim_{it} = 1 | \alpha_i, y_t, X_{1it}) = \Phi\left(\alpha_i + y_t\beta_y + X_{1it}\beta_1\right), \tag{1}$$

where $claim_{it} = 1$ when the insured has filed the claim in this policy year $t$, otherwise $claim_{it} = 0$. $X_{1it}$ is the vector of control variables. The variables within this vector only include the variables used by the insurance company for underwriting and pricing. In Taiwan, the insurance company only considers the occupation of the insured when it underwrites and prices the personal accident insurance contract and its hospital indemnity rider. Hence, only occupation variables exist in the $X_{1it}$ vector. $\beta_1$ is the vector of corresponding coefficients. Our data comprise the panel data which span the period from the year 2000 to the year 2010, and we use the random effects model to control the unobservable heterogeneity ($\alpha_i$) for each insured individual, and the year dummy variables ($y_t$) are also included in the vector of control variables to control for the year effect.

[28] This method is similar to the methodology in Dionne *et al.* (2001).

**Table 4** The preliminary observation of claim frequency among different contracts between two different channels

| | *cov_high = 1* | *cov_high = 0* |
|---|---|---|
| Panel 1: the personal accident claim frequency in the whole sample | | |
| *TM* | 0.0124 | 0.0100 |
| *OTC* | 0.0153 | 0.0094 |
| Panel 2: the hospital expense claim frequency in the sub-sample | | |
| *TM* | 0.0608 | 0.0534 |
| *OTC* | 0.0721 | 0.0642 |
| Panel 3: the ambiguous claim frequency in the sub-sample | | |
| *TM* | 0.0079 | 0.0070 |
| *OTC* | 0.0089 | 0.0069 |

After the probability of a claim $(\hat{claim}_{it})$ is estimated from the regression in the first stage, we estimate the probability of choosing a high-coverage insurance contract, also by the random effects probit regression, in the second stage:

$$\begin{aligned}&\text{Prob}\left(\text{cov\_}H_{it} = 1 | \alpha_i, y_t, \hat{claim}_{it}, claim_{it}, X_{2it}\right)\\&= \Phi\left(\alpha_i + y_t\beta_y + \beta_{estclm}\hat{claim}_{it} + \beta_{clm}claim_{it} + X_{2it}\beta_2\right),\end{aligned} \tag{2}$$

where $\text{cov\_}H_{it} = 1$ when the insured chooses a high-coverage contract in year $t$, otherwise $\text{cov\_}H_{it} = 0$. $X_{2it}$ is also the vector of occupation variables. With the exception of $X_{2it}$, the estimated probability of a claim calculated from the first stage $(\hat{claim}_{it})$ as well as the dummy variable of the claim $(claim_{it})$ are also included in the explanatory variables. The concept behind this method, raised by Dionne *et al.*,[29] is that if there is "asymmetric information", which makes the decision regarding choosing a high-coverage insurance contract and filing a claim related, the probability of choosing a high-coverage contract should be related to the claim decision, which is beyond what the underwriting and pricing variables can explain. The $\hat{claim}_{it}$ in the second-stage regression controls the explained part by means of the underwriting and pricing variables, and the $claim_{it}$ in the second-stage regression controls the residual part which is not completely explained by these variables. Hence, if the estimated coefficient of $claim_i$, $\hat{\beta}_{clm}$, is significantly different from 0, which means that there is asymmetric information that makes the decision of choosing a contract and filing a claim correlated, we refer to it as being "conditionally correlated".

Conditional correlation refers to the correlation after pricing or underwriting factors have been controlled for, and can only be controlled for. However, in Taiwan, the factors used for underwriting and pricing the personal accident insurance contract and its rider only include the occupational factors. The heteroskedasticity problem

[29] Dionne *et al.* (2001).

could arise due to the diversification among certain characteristics, such as sex, age, marital status and residential area, that individuals face. Chiappori and Salanie[30] mentioned that controlling the heterogeneity of insureds is very important when we test the conditional correlation between the contract coverage and claim. To avoid the estimation bias raised by heterogeneity, we seek to derive robust variance estimates.[31]

As mentioned earlier, when we investigate the asymmetric information problems in regard to personal accident insurance, the positive conditional correlation between coverage and death or a complete disability claim is more likely to be due to the phenomenon of adverse selection than to *ex ante*/*ex post* moral hazard or fraud. Accordingly, based on the personal accident insurance contracts, we build our Model 1 to test our Hypothesis 1. In Model 1, the insureds who are solicited through the TM channel and the insureds who are contracted through the OTC channel are separated into two subsamples and tested separately. In both subsamples, we first drop those observations that have filed claims other than for death or complete disability, and then, in Eq. (1), we define the $claim_{it} = 1$ by the insured who has filed for death or a complete disability claim in the policy year $t$, otherwise $claim_{it} = 0$; In Eq. (2), $\text{cov\_}H_{it} = 1$ when the coverage of the contract in year $t$ is higher than NT$2m, otherwise $\text{cov\_}H_{it} = 0$. The significantly positive conditional correlation between the coverage and claims, when $\hat{\beta}_{clm}$ is positive and significantly different from 0, is evidence of adverse selection. According to our Hypothesis 1, the adverse selection problem should be more severe among OTC customers than among TM customers. Hence, the prediction of Model 1 is that the $\hat{\beta}_{clm}$ in the OTC subsample should be more likely to be significantly different from 0 and should also be larger than $\hat{\beta}_{clm}$ in the TM subsample.

When we focus on the coverage of hospital expense insurance, we have a further chance to identify whether *ex post* moral hazard or fraud exists. Hence, based on the personal accident insurance contracts, we can build our Model 2 to test our Hypothesis 2. In Model 2, we still use Eq. (1) to perform the first-stage estimation. Its $claim_{it} = 1$ when the hospital expense has been claimed in policy year $t$, otherwise $claim_{it} = 0$. For the second-stage regression of this conditional correlation test, it is necessary to make some modifications. The additional term $ambgclaim_{it}$ is included:

$$\begin{aligned}&\Pr ob\left(\text{cov\_}H_{it} = 1 | \alpha_i, y_t, \hat{claim}_{it}, claim_{it}, ambgclaim_{it}, X_{2,it}\right)\\&= \Phi\left(\alpha_i + y_t\beta_y + \beta_{estclm}\hat{claim}_{it} + \beta_{clm}claim_{it} + \beta_{ambclm}ambgclaim_{it} + X_{2,it}\beta_2\right). \quad (3)\end{aligned}$$

We create a dummy variable $ambgclaim_{it}$, $ambgclaim_{it} = 1$ when the insured files a claim in policy year $t$, and this claim is an ambiguous claim, otherwise $ambgclaim_{it} = 0$.

[30] Chiappori and Salanie (2000).

[31] All the estimations in this research are performed by using robust variance estimates in order to avoid the estimation bias arising from heterogeneity.

We define the ambiguous claim in terms of the claim being due to the BPPV reason. Regardless of which kind of claim it is, if the conditional correlation between the coverage of the hospital expense insurance and the hospital expense claim (tested by $\hat{\beta}_{clm}$) is positive and significantly different from 0, there is the phenomenon of asymmetric information, but we cannot identify whether it is based on adverse selection, *ex ante*/*ex post* moral hazard or fraud. However, if this conditional correlation is also positive and significantly different from 0 when it is an ambiguous claim (tested by $\hat{\beta}_{ambclm}$), we can infer that *ex post* moral hazard or fraud also exists.

In Model 2, we also separate insureds into two subsamples of OTC customers and TM customers, before performing this revised conditional correlation test for these two subsamples separately. According to Hypothesis 2, OTC customers will be more likely to be associated with *ex post* moral hazard or fraud. Hence, the prediction of Model 2 is that $\hat{\beta}_{ambclm}$ should be more likely to be significantly different from 0 and larger among OTC customers than among TM customers.

## Empirical results

In Model 1, the personal accident insurance contracts are separated into the subsample that includes TM insureds and the subsample that includes OTC insureds. The empirical results of Model 1 are listed in Table 5.[32] We find that, for the subsample of TM insured, the estimated coefficient $\hat{\beta}_{clm}$ in regression (2) is positive but is not significantly different from 0. There is no significantly conditional correlation between coverage and claim for TM insureds. In the personal accident insurance market, there is no information asymmetry among the insureds from the TM channel.

In the subsample of insureds contracted through the OTC channel, we can also find in Table 5 that the estimated coefficient $\hat{\beta}_{clm}$ in regression (2) is positive and significantly different from 0 at the 1 per cent significance level. The conditional correlation between coverage and claim for OTC insureds is significantly positive, which means that insureds who purchased higher coverage in this group tend to have a higher probability of death or a complete disability claim. Thus it is less likely that insureds covered by a high coverage contract would become so careless as to encounter a fatal accident, or sacrifice their life or suffer serious bodily injury,

[32] Our research sample comprises panel data that cover a period of 11 years from the year 2000 to the year 2010. Hence, all our empirical regressions are not only performed by the random effects model, but also involve the use of ten-year dummy variables to control for the year effect. However, the approach is redundant if we list all these ten-year dummy variables in each regression result. Hence, in all the regressions from Table 5 to Table 8, as well as the tables in the Appendix, we have controlled for the year dummy variables in the regressions, but we do not display them in the tables of empirical results.

**Table 5** The conditional correlation test of Model 1

| | *Telephone channel* | | | | *OTC channel* | | | |
|---|---|---|---|---|---|---|---|---|
| | *Regress on* claim *Eq. (1)* | | *Regress on* cov_high *Eq. (2)* | | *Regress on* claim *Eq. (1)* | | *Regress on* cov_high *Eq. (2)* | |
| | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* |
| *Intercept* | −3.2284 | <0.0001 | −0.6868 | <0.0001 | −2.7253 | <0.0001 | 0.1456 | 0.1590 |
| $\widehat{claim}$ | | | 86.2730 | <0.0001 | | | 91.5434 | 0.0011 |
| *claim* | | | 0.2337 | 0.1711 | | | 1.5050 | <0.0001 |
| *occp_1* | −0.0443 | 0.5233 | −0.3839 | <0.0001 | 0.2682 | 0.4356 | −1.0253 | <0.0001 |
| *occp_2* | −0.0135 | 0.8488 | −0.2280 | <0.0001 | 0.1352 | 0.7018 | −0.8473 | <0.0001 |
| *occp_3* | 0.1064 | 0.1464 | −0.4554 | <0.0001 | 0.4113 | 0.2446 | −1.4382 | <0.0001 |
| *occp_4* | 0.0914 | 0.2134 | −0.4040 | <0.0001 | 0.4466 | 0.2019 | −1.4496 | <0.0001 |
| *occp_5* | 0.1907 | 0.0459 | −0.5783 | <0.0001 | 0.5591 | 0.1287 | −1.7353 | 0.0004 |
| *Log likelihood* | −516.521 | | −681.236 | | −641.211 | | −953.371 | |

regardless of whether they exaggerate a minor accident that actually happened or fabricate an accident, in order to take advantage of the insurance company. It is more likely to be the case that insureds who perceive themselves to be of the high-risk type will tend to purchase a high-coverage contract to avoid the possibility of financial loss caused by death or disability. Accordingly, it is more likely that the evidence will point to adverse selection instead of *ex ante*/*ex post* moral hazard or fraud in the personal accident insurance market.[33]

The comparison between TM insureds and OTC insureds based on the above results supports the inference of Hypothesis 1 in that the customers selected by TM exhibit a weaker problem of adverse selection than the over-the-counter insured. We also provide a formal test of the comparison of the asymmetric information problem in Appendix A, and the empirical results also further confirm our finding for Hypothesis 1.

A direct observation with regard to purchasing behaviour helps us better to understand the adverse selection problem among OTC insureds and TM insureds. According to the theory of adverse selection, we know that, if the adverse selection problem is severe for OTC customers, we would expect the OTC customers to purchase more insurance; on the contrary, TM insureds are characterised by

[33] We also perform sensitivity tests when we switch the threshold of the high vs low coverage to NT $1m. The empirical results are robust. Advantageous selection exists among TM insureds, whereas adverse selection is found to exist among OTC insureds. The conclusions are not affected by the level of the threshold that we use for the criterion of coverage. Owing to space constraints, we do not present the results of this sensitivity analysis. The empirical results are available upon request.

**Table 6** The decision factors for choosing a high-coverage contract in personal accident insurance

| *Variables* | *Est. coeff.* | *P-value* |
|---|---|---|
| *Intercept* | −1.9159 | <0.0001 |
| *teleph* | −0.0109 | 0.1095 |
| *female* | −0.2239 | <0.0001 |
| *married* | 0.4307 | <0.0001 |
| *age* | 0.0314 | <0.0001 |
| *agesq* | −0.0002 | <0.0001 |
| *ocp_1* | −0.2391 | <0.0001 |
| *ocp_2* | −0.3126 | <0.0001 |
| *ocp_3* | −0.2369 | <0.0001 |
| *ocp_4* | −0.0938 | 0.0008 |
| *ocp_5* | −0.0006 | 0.9868 |
| *city* | −0.0310 | <0.0001 |
| *north* | −0.0690 | <0.0001 |
| *south* | 0.0609 | <0.0001 |
| *central* | 0.0163 | 0.1794 |

comparatively fewer high-coverage contracts. We further run a Probit regression of *cov_high* on *teleph* and other control variables. The outcomes are listed in Table 6. We find that the estimated coefficient of *teleph* in this regression is negative, but it is insignificantly different from 0. The *p*-value of this corresponding estimated coefficient is 0.1037. It is close to but does not reach the statistically significant level. It means that the OTC insured has a higher, but not much higher, probability of purchasing a high-coverage contract than the TM insured.

What, then, is the reason for the tendency to choose a lower coverage contract to be insignificant in the TM group? This is not quite consistent with the inference of adverse selection that is judged by the conditional correlation test. The alternative approach to testing adverse selection, raised by Finkelstein and Poterba,[34] could help us to better understand this puzzle. Finkelstein and Poterba[34] use the unused characteristic, which is known to the insured, but is unknown to or unused by the insurance company, to test adverse selection. When the unused characteristic is positively (negatively) correlated with the coverage choice and is also positively (negatively) correlated with the claim probability, they refer to this as the evidence of adverse selection.

By following the idea of Finkelstein and Poterba,[34] we also collect several unused characteristics of the insured, such as their age, sex, marital status and residential area. We perform the test of Model 1 again by adding these unused characteristics to the control variables within the two regressions. The outcomes are listed in Appendix B.

[34] Finkelstein and Poterba (2014).

We not only find that the adverse selection phenomenon among OTC insureds is further sustained by this approach, but we are also unable to reject the existence of adverse selection among TM insureds. The most important finding is that, in the TM group, after these unused factors are controlled for, the conditional correlation between coverage and claims becomes negative and is significantly different from 0.

This empirical observation within the group of TM insureds is not surprising. We know that this group is solicited by the suggestion list, which is made up of those customers with better credit scores. They may be more cautious. Hence, even when the adverse selection phenomenon is relatively less severe in this group, we can expect these risk-averse insureds to also choose to purchase high-coverage contracts because of their exercising caution.[35] This further explains the puzzle that the rate at which higher-coverage contracts are purchased is not significantly lower in the TM group than in the OTC group.

In Model 2, we also separate the empirical tests for our two subsamples, namely, TM insureds and OTC insureds. The results are listed in Table 7. In the subsample of TM insureds, the estimated coefficients of $\hat{\beta}_{clm}$ and $\hat{\beta}_{ambclm}$ are not significantly different from 0. Accordingly, we cannot find any problem of asymmetric information among this group of insured. By contrast, both of the estimated coefficients $\hat{\beta}_{clm}$ and $\hat{\beta}_{ambclm}$ are positive and significantly different from 0 for the group of OTC insureds. This implies that there is significantly positive conditional correlation between the claim and the coverage of hospital expense insurance. This conditional correlation is also significantly higher when the claim is not easy to diagnose. There is thus evidence that not only does the phenomenon of adverse selection or *ex ante* moral hazard exist, but also the phenomenon of *ex post* moral hazard or fraud exists among this group of insureds.

These separate analyses tell us that only the insureds who are covered by high-coverage contracts and who are from the OTC channel will tend to have a higher probability of making a claim, and the claim probability will be even higher for ambiguous claims. Both adverse selection or *ex ante* moral hazard, and *ex post* moral hazard or fraud could exist among OTC insureds. However, these asymmetric information phenomena do not exist among TM insureds.[36] From this empirical work

[35] We find that adverse selection and advantageous selection may coexist in the TM group. In this paper, we have not discussed this coexistence phenomenon because of data and space constraints. This is an interesting issue that deserves further study in the future. It also explains the inference from the conditional correlation that the customers selected by telephone marketing are less affected by adverse selection.

[36] There are very few TM insureds who purchase the high amount of the hospital expense rider. Only 1.67 per cent of TM insureds are observed to have purchased high coverage for the rider when 66.41 per cent of OTC insureds have purchased high coverage for the rider. To prevent the statistical power from being a reason for there being no evidence of conditional correlation between the coverage and claim, we match the sample for this group of TM insureds. We select 0.86 per cent of subsample from TM insureds who purchased a low coverage rider. We keep the structure of the claim rate, occupation

**Table 7** The conditional correlation test of Model 2

| | *Telephone channel* | | | | *OTC channel* | | | |
|---|---|---|---|---|---|---|---|---|
| | *Regress on* claim *Eq. (1)* | | *Regress on* cov_high *Eq. (3)* | | *Regress on* claim *Eq. (1)* | | *Regress on* cov_high *Eq. (3)* | |
| | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* |
| *Intercept* | −4.6511 | 0.7463 | 2.0657 | 0.5425 | −2.0045 | <0.0001 | 4.4048 | 0.9742 |
| *claim* | | | −54.8219 | 0.2765 | | | −43.1452 | 0.1746 |
| *claim* | | | 0.0480 | 0.5716 | | | 0.8542 | 0.0287 |
| *ambgclaim* | | | 0.0670 | 0.6764 | | | 1.7741 | <0.0001 |
| *occp_1* | 0.2184 | <0.0001 | 1.3170 | 0.3618 | 0.3442 | 0.0886 | −7.3350 | 0.9570 |
| *occp_2* | 0.1141 | 0.0179 | 0.6829 | 0.4435 | 0.1599 | 0.4425 | −5.2206 | 0.9694 |
| *occp_3* | 0.2625 | <0.0001 | 1.5594 | 0.3633 | 0.3007 | 0.1521 | −7.4686 | 0.9562 |
| *occp_4* | 0.1975 | <0.0001 | 0.8019 | 0.5531 | 0.3454 | 0.0953 | −8.1173 | 0.9524 |
| *occp_5* | 0.2751 | <0.0001 | −1.6914 | 0.9915 | 0.0563 | 0.8064 | −5.2993 | 0.9689 |
| *Log likelihood* | −663.409 | | −500.666 | | −516.004 | | −573.279 | |

based on hospital indemnity riders, not only is the inference of our Hypothesis 2 supported, but the inference of our Hypothesis 1 is also sustained once again.[37]

Since the adverse selection phenomenon inferred by Hypothesis 1 has been sustained again, we also provide a comparison of the purchasing behaviour of TM insureds and OTC insureds. The outcomes of the probit regression of *cov_high* on *teleph* and other control variables are listed in Table 8. We find that the estimated coefficient of *teleph* in this regression is negative and significantly different from 0 at the 1 per cent significance level. This result shows that TM insureds have a significantly lower probability of purchasing high-coverage contracts than OTC insureds, and also further helps support the view that the adverse selection problem is

the same as that in the original sample of TM insureds who purchased a low coverage rider. This selected subsample together with TM insureds who purchase a high coverage riders constitutes the matching sample. There are only 5,011 observations in this matching sample. The percentage of high coverage riders in this matching sample now reaches 66.37 per cent. It approximately matches the percentage of high coverage contracts in the OTC insured subsample. By using this matching sample, we test Model 2 again. In this test we find that the estimated coefficients of $\hat{\beta}_{clm}$ and $\hat{\beta}_{ambclm}$ are still not significantly different from 0. The sign of $\hat{\beta}_{ambclm}$ and $\hat{\beta}_{clm}$is still positive. In other words, the evidence still shows that there is neither adverse selection, nor *ex post* moral hazard or fraud existing among TM insureds. The statistical power is not an issue here in changing our inference from the empirical tests. Owing to space constraints, the empirical results of this matching sample are only available from the authors upon request.

[37] We also perform a formal test of the comparison of the asymmetric information problem in Appendix C. The empirical results further confirm our findings for Hypotheses 1 and 2.

**Table 8** The decision factors for choosing a high-coverage contract with a hospital indemnity rider

| *Variables* | *Est. coeff.* | *P-value* |
|---|---|---|
| *Intercept* | 0.3283 | 0.0001 |
| *teleph* | −2.6523 | <0.0001 |
| *female* | −0.0779 | <0.0001 |
| *married* | 0.3806 | <0.0001 |
| *age* | −0.0255 | <0.0001 |
| *agesq* | 0.0004 | <0.0001 |
| *ocp_1* | 0.0557 | 0.3896 |
| *ocp_2* | −0.0162 | 0.8071 |
| *ocp_3* | −0.1704 | 0.0146 |
| *ocp_4* | −0.1820 | 0.0084 |
| *ocp_5* | −0.6957 | <0.0001 |
| *city* | 0.4576 | <0.0001 |
| *north* | 1.1897 | <0.0001 |
| *south* | 0.5932 | <0.0001 |
| *central* | 0.6488 | <0.0001 |

less serious for TM insureds than OTC insureds with regard to the hospital indemnity riders.[38]

In other words, customers with better credit scores that are sorted out and solicited through TM, as well as the transactions initiated by OTC customers, are two factors that account for the fact that there are fewer problems associated with adverse selection, *ex ante*/*ex post* moral hazard or fraud within the TM insured subsample.

## Conclusion

We explore two different marketing approaches when the insurance company distributes its contracts through the bank channel. One involves OTC selling. The other involves soliciting the contract through telephone-based interviews using a list provided by the bank side, with the list being sorted based on the better credit score customers in the bank.

[38] To better understand the adverse selection behind this conditional correlation test, we also perform an alternative test based on the unused characteristics of the insureds in Appendix D. In this test, we can only find that adverse selection exists in the OTC group, but we cannot find evidence of adverse selection in the TM group. The conditional correlation is also not significantly different from 0 after such unused private information is controlled for. These empirical outcomes together explain the reason why the percentage of those insureds choosing a high-coverage contract is significantly lower for the TM group than the OTC group.

The insurance company is characterised by less asymmetric information among TM insureds than among OTC insureds through the information revealed by the credit score. The individuals with better credit scores also tend to be more responsible, to exercise more self-restraint and to be persons who take less risk. TM insureds also incur higher costs if they engage in *ex post* moral hazard or commit fraud.

By contrast, the customers who purchase insurance contracts over the counter do so randomly. Accordingly, their type of risk, risk attitude and level of morality are average. Even though the agents at the bank counter may also hold information related to the credit score of these OTC customers, this information cannot be used for underwriting or pricing according to law. Hence, from the point of view that the insurance company can benefit from the information niche provided by the bank side among TM insureds, the asymmetric information problems, which include adverse selection, *ex ante*/*ex post* moral hazard or fraud, are less severe among these TM insureds than among OTC insureds.

From the point of view of those who initiate the transactions, it is the insurer in the TM group and the customer in the OTC group. It is then often the case that those who hold private information regarding the higher risk type or some hidden action information will be more willing to travel to the bank, buy the insurance and buy more coverage over the counter. Hence, we can believe that there is more private information or hidden action, and accordingly more severe adverse selection, *ex ante*/*ex post* moral hazard or fraud problems among OTC customers than among TM customers.

We therefore investigate the asymmetric information problems through the personal accident insurance contract as well as its extended coverage in the form of a hospital indemnity rider. We find that it is more likely for there to be an adverse selection problem existing in the personal accident insurance contract. Furthermore, the adverse selection problem is comparatively less common among TM insureds. There are also the problems of adverse selection, *ex ante* moral hazard, *ex post* moral hazard or fraud co-existing in relation to the hospital indemnity rider. Meanwhile, these problems are also comparatively less severe among TM insureds.

These findings provide us with several important policy implications. First, the bank can provide an information niche for the insurance company so that by making good use of the exclusive valuable private information collected by banks, the problem of information asymmetry can be mitigated. Second, the one who initiates the transaction also matters when the asymmetric information problems exist. We show that the one who initiates the transaction is the side that holds more hidden information and hidden action, which would give rise to asymmetric information problems.

Third, this study also provides a very important hint when the private information is prevented from being directly collected or explicitly used across different businesses, especially for some industries in which asymmetric information issues

are very important, such as the insurance industry. Implicitly using this information by initiating the trade and soliciting the product according to the "suggestion list" collected through this information could be one solution. The empirical results for TM insureds in this paper provide an excellent example.

## Acknowledgements

The authors would like to express their deep gratitude to Alexander Muermann, Larry Tzeng and the two referees who made stimulating comments on the paper.

## References

Arnett, J. (1990) Drunk driving, sensation, and egocentrism among adolescents, *Personality and Individual Differences* 11(6): 541–546.

Bergendal, G. (1995) The profitability of bancassurance for European banks, *International Journal of Bank Marketing* 13(1): 17–28.

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

Chiappori, P. and Salanie, B. (2000) Testing for asymmetric information in insurance markets, *Journal of Political Economy* 108(1): 56–78.

Dionne, G. and St-Michel, P. (1991) Workers' compensation and moral hazard, *Review of Economics and Statistics.* 73(2): 236–244.

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.

Donovan, J. E. (1993) Young adult drinking-driving: Behavioral and psychosocial correlates, *Journal of Studies on Alcohol* 54(5): 600–613.

Duffield, G. M. and Grabosky, P. N. (2001) *The Psychology of Fraud*. Department of the Premier and Cabinet, Queensland Government, http://www.aic.gov.au.

Duval, T. and Wicklund, R. A. (1972) *A Theory of Objective Self Awareness*, New York: Academic Press.

Fields, L. P., Fraser, D. R. and Kolari, J. W. (2007) Is bancassurance a viable model for financial firms? *The Journal of Risk and Insurance* 74(4): 777–794.

Finkelstein, A. and Poterba, J. (2014) Testing for asymmetric information using "unused observables" in insurance markets: Evidence from the U.K. annuity market, *The Journal of Risk and Insurance*, article first published online: 18 February 2014, doi:10.111/jori.12030.

Genetay, N. and Molyneux, P. (1998) *Bancassurance*, London: Macmillan Press.

Hoschka, T. C. (1994) *Bancassurance in Europe*, London: Palgrave Macmillan.

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

Langer, E. J. (1989) Minding matters: The consequences of mindlessness-mindfulness, in L. Berkowitz (ed.) *Advances in Experimental Social Psychology.* San Diego, CA: Academic Press, pp. 137–173.

Lastovicka, J. L., Murry, J. P., Joachimsthaler, E. A., Bhalla, G. and Scheurich, J. (1987) A lifestyle typology to model young male drinking and driving, *The Journal of Consumer Research* 14(2): 257–263.

Mazar, N., Amir, O. and Ariely, D. (2008) The dishonesty of honest people: A theory of self-concept maintenance, *Journal of Marketing Research* 45(6): 633–644.

McMillen, D. L., Pang, M. G., Wells-Parker, E. and Anderson, B. J. (1992) Alcohol, personality traits, and high risk driving: A comparison of young, drinking driver groups, *Addictive Behaviors* 17: 525–532.
Rumelt, R. P. (1974) *Strategy, Structure, Economic Performance*, Boston, MA: Division of Research, Harvard Business School.
Staikouras, S. K. and Nurullah, M. (2005) *The separation of banking from insurance: Evidence from Europe*, working paper, Cass Business School, City University and Glasgow Caledonian University.

## Appendix A

*The comparative conditional correlation test for Model 1*

A formal test of the conditional correlation among two subsamples can be performed by adding one more interaction term of the TM dummy variable and claim dummy variable ($teleph_{it} \times claim_{it}$) in the second stage of regression (2):

$$\begin{aligned}&\Pr ob\left(\text{cov\_}H_{it} = 1 | \alpha_i, y_t, \hat{claim}_{it}, claim_{it}, teleph_{it}, teleph_{it} \times claim_{it}, X_{2it}\right)\\&= \Phi\left(\alpha_i + y_t\beta_y + \beta_{estclm}\hat{claim}_{it} + \beta_{clm}claim_{it} + \beta_{telph}teleph_{it} + \beta_{telphclm}teleph_{it} \times claim_{it} + X_{2it}\beta_2\right).\end{aligned} \tag{A.1}$$

The key estimated coefficient used for testing our first hypothesis is $\hat{\beta}_{telphclm}$, which is the estimated coefficient of $teleph_{it} \times claim_{it}$. When there is a significantly positive conditional correlation between claims and coverage, according to our first hypothesis, $\hat{\beta}_{telphclm}$ should be negative and significantly different from 0. The empirical inference is that it is most likely to be characterised by the phenomenon of adverse selection, and this phenomenon is less severe for the TM channel than for the OTC channel. The empirical outcomes are presented in Table A1.

**Table A1** A formal comparison of the asymmetric information problem in Model 1 between TM insureds and OTC insureds

| *Variables* | *Regress on* claim *Eq (1)* | | *Regress on* cov_high *Eq (2)* | | *Regress on* cov_high *Eq (A.1)* | |
|---|---|---|---|---|---|---|
| | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* |
| *Intercept* | −2.5915 | <0.0001 | −0.5308 | <0.0001 | −0.4998 | <0.0001 |
| *claim̂* | | | −52.4823 | 0.0013 | −47.6436 | 0.0036 |
| *claim* | | | 1.5032 | <0.0001 | 1.5816 | <0.0001 |
| *teleph* | | | | | −0.1081 | 0.0269 |
| *claim×teleph* | | | | | −1.2733 | <0.0001 |
| *occp_1* | −0.0156 | 0.8184 | −0.2327 | <0.0001 | −0.2534 | <0.0001 |
| *occp_2* | −0.0027 | 0.9687 | −0.3460 | <0.0001 | −0.3456 | <0.0001 |
| *occp_3* | 0.1274 | 0.0748 | −0.0195 | 0.7872 | −0.0417 | 0.5632 |
| *occp_4* | 0.1196 | 0.0943 | 0.1101 | 0.1058 | 0.0849 | 0.2127 |
| *occp_5* | 0.2287 | 0.0107 | 0.4000 | 0.0041 | 0.3454 | 0.0133 |
| *Log likelihood* | −622.86 | | −676.827 | | −676.827 | |

The negative and statistically significant estimated coefficient ($\hat{\beta}_{telphclm}$) has further confirmed the comparatively more severe adverse selection problem among OTC insureds.

## Appendix B

*The alternative approach to test for adverse selection for Model 1*

An alternative approach to test for adverse selection is to add the insured's unused characteristics vector ($Z_{it}$) to the regressions of Model 1. The insured's unused characteristics ($Z_{it}$) include the sex variable (*female*), marital status variable (*married*), age variables (*age* and $age^2$) and residential area variables (*city, north, south, central*). Hence, the alternative test for Model 1 is:

$$\Pr ob(claim_{it} = 1|\alpha_i, y_t, X_{1it}, Z_{it}) = \Phi\left(\alpha_i + y_t\beta_y + X_{1it}\beta_1 + Z_{it}\beta_{z1}\right), \tag{B.1}$$

$$\begin{aligned}&\Pr ob\left(\operatorname{cov} H\_{it} = 1|\alpha_i, y_t, \hat{claim}_{it}, claim_{it}, X_{2it}, Z_{it}\right)\\ &= \Phi\left(\alpha_i + y_t\beta_y + \beta_{estclm}\hat{claim}_{it} + \beta_{clm}claim_{it} + X_{2it}\beta_2 + Z_{it}\beta_{z2}\right).\end{aligned} \tag{B.2}$$

**Table B1** Alternative test for asymmetric information in Model 1—by using unused characteristics

| | *TM insureds* | | | | *OTC insureds* | | | |
|---|---|---|---|---|---|---|---|---|
| | *Regress on* claim | | *Regress on* cov_high | | *Regress on* claim | | *Regress on* cov_high | |
| | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* |
| *Intercept* | −2.9879 | <.0001 | −4.9714 | 0.4918 | −2.4906 | <0.0001 | −2.1283 | <0.0001 |
| $\hat{claim}$ | | | 9.1303 | <0.0001 | | | 23.1223 | 0.0004 |
| *claim* | | | −0.2496 | <0.0001 | | | 1.1282 | <0.0001 |
| *female* | −0.0859 | <0.0001 | −0.2248 | <0.0001 | −0.1245 | <0.0001 | −0.1631 | <0.0001 |
| *married* | −0.1532 | <0.0001 | 0.4360 | <0.0001 | −0.1006 | 0.0010 | 0.4004 | <0.0001 |
| *age* | −0.0114 | 0.0221 | −0.0306 | <0.0001 | −0.0071 | 0.0629 | −0.0684 | <0.0001 |
| $age^2$ | 0.0003 | <0.0001 | 0.0007 | <0.0001 | 0.0002 | 0.0011 | −0.0008 | <0.0001 |
| *occp_1* | 0.0161 | 0.8178 | −0.2606 | <0.0001 | 0.2629 | 0.4432 | −0.3616 | 0.0003 |
| *occp_2* | 0.0245 | 0.7310 | −0.3223 | <0.0001 | 0.1429 | 0.6844 | −0.4273 | <0.0001 |
| *occp_3* | 0.1344 | 0.0671 | −0.2138 | <0.0001 | 0.3836 | 0.2762 | −0.4019 | 0.0007 |
| *occp_4* | 0.0960 | 0.1918 | −0.0873 | 0.0041 | 0.3791 | 0.2770 | −0.2105 | 0.0717 |
| *occp_5* | 0.1861 | 0.0519 | −0.1046 | 0.0236 | 0.5163 | 0.1597 | 0.1324 | 0.3813 |
| *city* | −0.0307 | 0.0659 | −0.0418 | <0.0001 | −0.3279 | 0.0003 | −0.0901 | <0.0001 |
| *north* | −0.1020 | 0.0040 | −0.0985 | <0.0001 | −0.1304 | 0.0065 | −0.0408 | <0.0001 |
| *south* | 0.0108 | 0.7730 | 0.0877 | <0.0001 | 0.0536 | 0.1156 | 0.1132 | 0.0140 |
| *central* | −0.0859 | <0.0001 | −4.9714 | 0.4918 | −0.1245 | <0.0001 | −2.1283 | <0.0001 |
| *Log likelihood* | −516.521 | | −667.566 | | −641.211 | | −921.263 | |

The key estimated coefficients used for testing adverse selection are $\hat{\beta}_{z1}$ and $\hat{\beta}_{z2}$. If there are some coefficients in $\hat{\beta}_{z1}$ and $\hat{\beta}_{z2}$ that display the same sign and are both significantly different from 0, they are the private information regarding risk type and indicate that adverse selection exists.

In both groups of insureds, we can see that *female, age, city* and *north* are unused and valuable private information which reveal the risk type of the insured. Female, aged insureds and insureds who live in the city or the northern part of Taiwan are comparatively lower risk-type insureds. They both choose a low coverage contract and face a lower probability of risk at the same time. This result not only further sustains the adverse selection inference from the conditional correlation test, but it also does not reject the idea that adverse selection also exists. Furthermore, after such unused information is controlled for, the significantly negative conditional correlation between coverage and claims emerges in the TM group. SeeTable B1.

## Appendix C

*The comparative conditional correlation test for Model 2*

The formal comparison of these two groups is further performed by adding $teleph_{it}$ and two interaction terms of $claim_{it} \times teleph_{it}$ and $ambgclaim_{it}$v$teleph_{it}$, in the second stage regression of Eq. (3):

$$\begin{aligned}
&\Pr ob\left(\text{cov_}H_{it} = 1 | \alpha_i, y_t, \hat{claim}_{it}, claim_{it}, teleph_{it}, ambgclaim_{it}, \right. \\
&\left. claim_{it} \times teleph_{it}, ambgclaim_{it} \times teleph_{it}, X_{2,it}\right) \\
&= \Phi\left(\alpha_i + y_t\beta_y + \beta_{estclm}\hat{claim}_{it} + \beta_{clm}claim_{it} + \beta_{teleph}teleph_{it} + \beta_{ambclm}ambgclaim_{it}\right. \\
&\quad \left. + \beta_{clmtelph}claim_{it} \times teleph_{it} + \beta_{ambclmtelph}ambgclaim_{it} \times teleph_{it} + X_{2,it}\beta_2\right). \qquad \text{(C.1)}
\end{aligned}$$

When there is positive conditional correlation between coverage and claims, if the estimated coefficient of $claim_{it} \times teleph_{it}$ $(\hat{\beta}_{clmtelph})$ is negative and significantly different from 0, this means that the asymmetric information problem is comparatively less severe in the TM channel. The key variable used to test our second hypothesis is the estimated coefficient of $ambg_{it} \times claim_{it} \times teleph_{it}(\hat{\beta}_{ambclmtelph})$. When there is a positive conditional correlation between coverage and the ambiguous claim, according to the inference of our second hypothesis, $\hat{\beta}_{ambclmtelph}$ should be negative and significantly different from 0. This means that the especially high conditional correlation between the coverage and claims when the reason is not easy to diagnose is less pronounced for those insureds who are solicited through the TM channel. The *ex post* moral hazard or fraud problem is less severe in the case of this TM channel than in the case of the OTC channel. The empirical results are shown in Table C1.

**Table C1** A formal comparison of the asymmetric information problem in Model 2 between TM insureds and OTC insureds

| *Variables* | *Regress on* claim *Eq (1)* | | *Regress on* cov_high *Eq (3)* | | *Regress on* cov_high *Eq (C.1)* | |
|---|---|---|---|---|---|---|
| | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* |
| *Intercept* | −1.9062 | <0.0001 | 1.0344 | 0.2205 | 3.7047 | <0.0001 |
| *claim* | | | −13.3143 | 0.6650 | −67.5461 | 0.0672 |
| *claim* | | | 0.8185 | 0.0309 | 0.8653 | 0.0298 |
| *teleph* | | | | | −2.5240 | <0.0001 |
| *ambgclaim* | | | 1.8217 | <0.0001 | 1.8219 | <0.0001 |
| *claim×teleph* | | | | | −0.8017 | 0.0089 |
| *ambgclaim×teleph* | | | | | −1.7604 | <0.0001 |
| *occp_1* | 0.2381 | <0.0001 | 0.3912 | 0.6668 | 1.5988 | 0.1520 |
| *occp_2* | 0.1111 | 0.0179 | −0.2593 | 0.6287 | 0.6568 | 0.3185 |
| *occp_3* | 0.2571 | <0.0001 | −0.4477 | 0.6505 | 1.4399 | 0.2358 |
| *occp_4* | 0.2046 | <0.0001 | −0.3059 | 0.7031 | 0.7318 | 0.4594 |
| *occp_5* | 0.2227 | 0.0003 | −0.7824 | 0.4111 | −0.0082 | 0.9945 |
| *Log likelihood* | −461.114 | | −477.739 | | −477.739 | |

Before we control for the three additional terms, the $\hat{\beta}_{clm}$ is positive and significantly different from 0 at the 5 per cent significance level, and the $\hat{\beta}_{ambclm}$ is also positive and significantly different from 0 at the 1 per cent significance level. This means that the asymmetric information problem in this market is characterised by adverse selection, and *ex post* moral hazard and/or fraud also co-exists in this market.

When we control the three additional terms of $teleph_{it}$, $claim_{it} \times teleph_{it}$ and $ambg_{it} \times claim_{it} \times teleph_{it}$ in the regression, $\hat{\beta}_{clm}$ and $\hat{\beta}_{ambclm}$ are still positive and significantly different from 0 at the 5 and 1 per cent significance levels. By contrast, $\hat{\beta}_{ambclmtelph}$ and $\hat{\beta}_{clmtelph}$ are negative and significantly different from 0 at the 1 and 5 per cent significance levels. We can also observe that the absolute value of this estimated coefficient $\hat{\beta}_{ambclmtelph}$ is smaller than the value of the estimated coefficient $\hat{\beta}_{ambclm}$. Their sums are now positive. This means that there could be a positive correlation between the claims and the coverage of hospital expense insurance for the TM insured group. This observation is consistent with the separate analysis of the TM insured group in Table 7, although the result is not statistically significant.

The key estimated coefficient for our second hypothesis is $\hat{\beta}_{ambclmtelph}$. When it is negative and significantly different from 0, this finding implies that the *ex post* moral hazard and/or fraud problem is comparatively less severe for the TM channel. Our second hypothesis is sustained. Furthermore, the negative and significantly different form 0 of $\hat{\beta}_{clmtelph}$ also means that the adverse selection phenomenon could be comparatively less severe for this TM channel. Hence, both of our hypotheses are supported by such evidence in the hospital expense insurance market.

## Appendix D

*The alternative approach to test for adverse selection for Model 2*

We also use an alternative approach to test for adverse selection by adding the insured's unused characteristics vector ($Z_{it}$), which was referred to in Appendix B, to the regressions of Model 2. The outcomes are shown in Table D1

We find that there is no unused variable that reveals the private information of the risk type in the TM group. Hence, there is no adverse selection phenomenon within this group that is further sustained. However, in the OTC group, the unused variables *female, north* and *central* are still valuable private information which reveal the risk type of these OTC insureds: the insureds who are married or who live in the northern or central part of Taiwan tend to purchase higher coverage for the hospital expense rider and also tend to have a higher probability of filing a claim. The evidence of adverse selection is further sustained in this group.

**Table D1** Alternative test for asymmetric information in Model 2—by using unused characteristics

| | *TM insureds* | | | | *OTC insureds* | | | |
|---|---|---|---|---|---|---|---|---|
| | *Regress on* claim | | *Regress on* cov_high | | *Regress on* claim | | *Regress on* cov_high | |
| | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* | *Est. coeff.* | *P-value* |
| *Intercept* | −4.6018 | 0.7457 | −6.0331 | 0.9746 | −2.2737 | <0.0001 | 2.1569 | 0.9873 |
| *claim* | | | 10.7027 | 0.3866 | | | 35.1375 | 0.0031 |
| *claim* | | | −0.0555 | 0.6781 | | | 0.7504 | 0.0410 |
| *ambg×claim* | | | 0.0375 | 0.7916 | | | 1.6120 | <0.0001 |
| *female* | −0.0814 | <0.0001 | −0.1495 | 0.3802 | −0.2185 | <0.0001 | −1.0915 | 0.0027 |
| *married* | 0.0341 | 0.0017 | −0.4487 | 0.0062 | 0.0804 | <0.0001 | 1.3192 | <0.0001 |
| *age* | −0.0053 | 0.1060 | 0.0061 | 0.8833 | −0.0025 | 0.2722 | −0.0437 | 0.0619 |
| $age^2$ | 0.0002 | <0.0001 | −0.0003 | 0.6034 | 0.0001 | <0.0001 | 0.0010 | 0.0079 |
| *occp_1* | 0.2388 | <0.0001 | −0.2365 | 0.6824 | 0.3708 | 0.0636 | −1.1984 | 0.9930 |
| *occp_2* | 0.1472 | 0.0024 | −0.1743 | 0.7399 | 0.2021 | 0.3267 | −1.8993 | 0.9889 |
| *occp_3* | 0.2810 | <0.0001 | −0.4837 | 0.4522 | 0.3545 | 0.0881 | −1.7195 | 0.9899 |
| *occp_4* | 0.2086 | <0.0001 | −0.7342 | 0.2299 | 0.3579 | 0.0807 | −1.6134 | 0.9905 |
| *occp_5* | 0.2803 | <0.0001 | −4.3286 | 0.9910 | 0.1041 | 0.6476 | −3.2298 | 0.9810 |
| *city* | −0.0139 | 0.2112 | 0.2901 | 0.0380 | −0.0023 | 0.9275 | 0.6083 | 0.0004 |
| *north* | −0.1886 | <0.0001 | 3.9434 | 0.9834 | 0.1143 | 0.0002 | 1.5142 | <0.0001 |
| *south* | 0.0008 | 0.9736 | 3.6880 | 0.9844 | 0.1029 | 0.0098 | 0.1427 | 0.6155 |
| *central* | −0.0814 | <0.0001 | −0.4837 | 0.4522 | 0.2503 | <0.0001 | 1.0856 | 0.0188 |
| *Log likelihood* | −663.409 | | −500.666 | | −516.004 | | −573.279 | |

## About the Authors

**Jin-Lung Peng** is PhD from National Chengchi University and an assistant professor in the Department of Risk Management and Insurance of National Chengchi University. He is also a research fellow in the Risk and Insurance Research Center of College of Commerce in National Chengchi University.

**Kili C. Wang** is PhD from National Taiwan University. She is a Professor in the Department of Insurance of Tamkang University. She is also a research fellow in the Risk and Insurance Research Center of College of Commerce in National Chengchi University.