Subjective Belief, Risk Information and Earthquake Insurance Purchase: Evidence from Japanese Post-Quake Data*

Michio Naoi†, Takuya Ishino‡ and Miki Seko§

March 1, 2014

Abstract

We develop a simple model of insurance purchase where consumers have imperfect knowledge about the potential risk based on Kunreuther and Pauly (2005) and test its theoretical implications using observed insurance behavior of potential policy holders. Our empirical findings suggest that obtaining the regional disaster hazard information makes the subjective probability of a loss significantly higher, thereby it facilitates insurance purchase among potential policy holders. This result is robust to the possible endogeneity of consumer's information search behavior. Our result implies that appropriate dissemination of objective risk information can enhance the consumer's risk mitigation activities.

JEL Classification: D03, D81, D83, G22.

Keywords: Subjective belief, risk perception, earthquake insurance.

---

* We thank Misuzu Azuma, Hiroshi Fujiu, Keiichiro Kobayashi, Colin R. McKenzie, Masao Ogaki, Rui Ota, Makoto Yano and the participants at the Market Quality Research Project Workshop and at the Western Economic Association International, 10th Biennial Pacific Rim Conference for their helpful comments and suggestion.

† Corresponding Author. Faculty of Economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan. (e-mail: naoi@econ.keio.ac.jp).

‡ Department of Economics, Kanazawa Seiryo University, 10-1 Ushi, Gosho-machi, Kanazawa-shi, Ishikawa 920-8620, Japan. (e-mail: stonefield@gmail.com).

§ Faculty of Political Science and Economics, Musashino University, 3-3-3 Ariake, Koto-ku, Tokyo 135-8181, Japan. (e-mail: seko@econ.keio.ac.jp).

1. Introduction

The Great East Japan Earthquake and subsequent tsunami on March 11, 2011 were a reminder of Japan's huge earthquake risk and need for appropriate measures to mitigate disasters. Obviously, earthquake insurance is one of the most important measures for mitigating earthquake losses. Nonetheless, despite increased disaster awareness and preparedness following the March 11 earthquake, the proportion of Japanese households covered by earthquake insurance is still low.

According to the General Insurance Association of Japan, only 24 percent of Japanese households had purchased earthquake insurance as of 2011. The low purchase rates for earthquake insurance have important policy implications for not only Japan but also many other quake-prone countries. Previous studies have analyzed the incidence of underinsurance against natural disasters, including earthquakes (Kunreuther, 1984; Naoi et al., 2010). Aware of sweeping exclusions, homeowners may be skeptical that their claims will be fully compensated. Also, ex-post government grants for victims create an immense moral hazard for the insurance market. Since inhabitants in quake-prone areas expect to receive grants, compensation and low cost loans from local and central governments after a massive earthquake, they have little incentive to purchase insurance.

In addition to these potential reasons, several previous studies suggest that consumer's misperception of earthquake risk might be another source of underinsurance among potential policy holders. Generally, consumers have difficulty in dealing with low-probability, high-loss events, and earthquake would be plausible candidate of such an event. In fact, Fujimi and Kakimoto (2012) report that the majority of homeowners in Japan considerably underestimate house destruction risks due to earthquake.

Our contribution in this paper is twofold. First, we develop a simple model of insurance purchase where consumers have imperfect knowledge about the potential risk and test its theoretical implications using observed insurance behavior of potential policy holders. Although insurance decision of consumers with imperfect knowledge is studied in several theoretical models in recent years, there are only a few empirical studies that deal explicitly with the concept. As far as we know, this is the first study to empirically test the theoretical predictions of such models using the consumer's insurance behavior as a prominent example.

Second, and related to the first contribution, our theoretical model and empirical evidence can explain why so many people do not insure against natural disasters. Our empirical findings suggest that people obtaining the earthquake-related hazard information are significantly more likely to insure themselves against earthquake, implying that the consumer's underestimation of disaster risk can be a major source of underinsurance against natural disasters, and that detailed hazard information can enhance consumer's ability to assess disaster risks.

Organization of the rest of the paper is as follows. In Section 2, we illustrate a simple model of consumer's subjective belief and insurance decision, based on Kunreuther and Pauly (2004), and provide a set of empirically testable hypotheses. Section 3 describes the dataset and variables. Section 4 sets out our empirical model and provides empirical results. Section 5 concludes the paper.

2. Subjective Belief and Insurance Purchase

As discussed in the previous section, consumer's misperception of potential earthquake risk can be a major source of underinsurance. In the canonical model of insurance demand, consumers are often assumed to have precise information about potential risk. However, psychological literature provides a wide range of experimental evidence that individuals often have systematic biases in evaluating probabilistic events (Slovic et al., 1986). This requires a significant departure from the canonical models. In this section, we present a simple model of insurance purchase where consumers have imperfect knowledge about the potential risk but have an opportunity to obtain better risk information at a fixed search cost. We adopt a similar setting as Kunreuther and Pauly (2004), and draw several empirically testable predictions from the model.

2.1. Setting

Consider an individual who faces a risk of earthquake events causing a substantial loss relative to his wealth. For simplicity, we assume that all individuals suffer the same loss (L) in case of an earthquake.[1] The person is assumed to be risk averse, have wealth (W), and wants to determine how much insurance (I) to purchase.

The individual is assumed to have an imperfect knowledge about the loss probability in the sense that he believes that the probability has n possible values, p_j $(j = 1, \cdots, n)$, with subjective likelihoods or "weights" w_j associated with each p_j where $\sum_{\mathrm{j}} w_j = 1$.

Although the individual initially does not know the exact loss probability, he can obtain some information about the loss probability at a fixed search cost (C). Assume that once information is obtained the most accurate estimate of the loss probability is chosen and the individual determines his insurance amounts accordingly. The consumer's problem is to determine how much insurance to purchase so as to maximize expected utility by either deciding to undertake a search for information or not searching at all. The outcome will either be $\mathrm{EU}(Search)$ or $\mathrm{EU}(No\ Search)$.

2.2. Insurance Decision without Searching

If the individual decides not to seek additional information, the "subjective" expected utility is

$$\mathrm{EU}(No\ Search) = \sum_{j=1}^{n} w_j \left[p_j U(W - L + (1-r)I) + \left(1 - p_j\right) U(W - rI)\right], \tag{1}$$

where r is insurance premium.

Optimal amount of insurance (I^*) is determined by maximizing $\mathrm{EU}(No\ Search)$ subject to $0 \leq I \leq L$. The first-order condition is given by

---

[1] Ehrlich and Becker (1972) examine two alternatives to market insurance: self-insurance—a reduction in the size of a loss—and self-protection—a reduction in the probability of a loss. They show that market insurance and self-insurance are substitutes. As a result, the degree of loss after the particular earthquake would be smaller in riskier region. However, they also show that the incentive for self-insurance, compared with that for market insurance, is smaller for rare losses like earthquake.

$$\hat{p}(1-r)U'(W-L+(1-r)I^*)-(1-\hat{p})rU'(W-rI^*)=0, \tag{2}$$

where $0 \le I^* \le L$ and $\hat{p}=\sum_j w_j p_j$ which is the average subjective probability of a loss. In this case, note that the optimal amount of insurance depends only on the average subjective probability $(\hat{p})$, not on possible probability values (p_j) or their distribution (w_j), because expected utility is linear in probabilities.

The individual will choose not to purchase any insurance if the left-hand-side of equation (2) is negative at $I=0$. Hence the decision whether or not to purchase insurance can be given as

$$a^* = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} \text{ if } \hat{p} \begin{Bmatrix} \le \\ > \end{Bmatrix} \frac{rU'(W)}{(1-r)U'(W-L)+rU'(W)}, \tag{3}$$

where a^* is a binary indicator whether or not purchase insurance.

2.3. Insurance Decision with Searching

Following Kunreuther and Pauly (2004), we assume that, if the individual decides to seek additional information, expected utility is determined by the following procedure. First, the individual chooses the optimal amount of insurance for each subjective estimate of the probability of a loss. Namely, the amount of insurance with the probability estimate p_j is give as

$$I_j^* \equiv \arg\max_I \; p_j U(W-L+(1-r)I)+(1-p_j)U(W-rI), \tag{4}$$

where $0 \le I_j^* \le L$.

Given that search takes place at a fixed cost C, the expected utility when searching for information is, conditional on the optimal levels of insurance demand for each p_j given by equation (4), determined as

$$\mathrm{EU}(Search)=\sum_{j=1}^{n} w_j\left[p_j U\left(W-L+(1-r)I_j^*\right)+(1-p_j)U\left(W-rI_j^*\right)\right]-C. \tag{5}$$

The individual will search for information on the probability of a loss only when $\mathrm{EU}(Search)$ is greater than the maximum value of $\mathrm{EU}(No\ Search)$.

Once information is obtained, the estimate of the loss probability is updated and the individual determines his insurance amounts accordingly. Let $\hat{p}'$ be the *ex-post* subjective probability that takes additional information into account. Then the *ex-post* insurance demand is determined by maximizing expected utility based on $\hat{p}'$. Analogous to equation (3), the *ex-post* decision whether or not to purchase insurance can be given as

$$a^* = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} \text{ if } \hat{p}' \begin{Bmatrix} \le \\ > \end{Bmatrix} \frac{rU'(W)}{(1-r)U'(W-L)+rU'(W)}. \tag{6}$$

where again a^* is a binary indicator whether or not purchase insurance.

If the individual has an overly optimistic estimates of a loss probabilities, additional information is likely to increase consumer's subjective probability $(\hat{p}' > \hat{p})$. In this case, comparing equations (3) and (6) suggests that the individual will be more likely to purchase insurance after obtaining additional information.

2.4. Illustrative Example and Prediction of the Model

The individual will search for information on the probability of a loss only when $\mathrm{EU}(Search)$ is greater than the maximum value of $\mathrm{EU}(No\,Search)$. Figure 1 presents the relationship between the average subjective probability $(\hat{p})$ and expected utility levels with or without search.

(Figure 1 around here)

Without searching, the individual with relatively low $\hat{p}$ will optimally choose not to purchase any insurance $(I^* = 0)$. Hence $\mathrm{EU}(No\,Search)$ decreases linearly as $\hat{p}$ rises up to some point.[2] On the other hand, the individual with relatively high $\hat{p}$ will purchase full coverage insurance $(I^* = L)$. Hence, beyond some threshold point, $\mathrm{EU}(No\,Search)$ remains constant whatever the level of the average probability. In between these two situations, consumers purchase partial coverage insurance $(0 < I^* < L)$, and $\mathrm{EU}(No\,Search)$ decreases at a decreasing rate as $\hat{p}$ rises.

When the individual decides to undertake a search for information, the expected utility $\mathrm{EU}(Search)$ is given by equation (5). In general, $\mathrm{EU}(Search)$ depends not only $\hat{p}$ but also on possible probability values (p_j) and their distribution (w_j). For an illustrative purpose, consider the following special case. Let p_k and p_l be two possible subjective probabilities such that $p_k > p_l$. Let us consider the case where w_k, the weight associated with p_k, is increased by some amount, and w_l is decreased by the same amount. Obviously such a change would make $\hat{p}$ higher and, with some straightforward calculation, we can show that

$$\frac{\mathrm{dEU}(Search)}{\mathrm{d}\hat{p}} = \frac{\mathrm{EU}_k - \mathrm{EU}_l}{p_k - p_l}, \tag{7}$$

where $\mathrm{EU}_j = p_j U\big(W - L + (1-r)I_j^* - C\big) + \big(1 - p_j\big)U\big(W - rI_j^* - C\big)$. Except for the case where $I_k^* = I_l^* = L$ (i.e., full coverage), $\mathrm{EU}_k < \mathrm{EU}_l$ holds. As a result, the right-hand-side of equation (7) becomes negative, and we see that $\mathrm{EU}(Search)$ is a decreasing function of $\hat{p}$. Figure 1 shows the special case that the individual has only two possible probability estimates $(n = 2)$ where $p_1 = 0$ and $p_2 = 1$.[3]

From Figure 1, we can see that there are three possible outcomes regarding insurance purchase and information search. First, if the individual has relatively low $\hat{p}$, he/she will neither purchase any insurance $(a = 0)$, nor search for any additional information $(s = 0)$. Second, for relatively high $\hat{p}$, the individual will purchase insurance $(a = 1)$ without any search $(s = 0)$. Finally, in between these two cases, the individual with moderate $\hat{p}$ will choose to undertake a search for information $(s = 1)$. In the latter case, the individual may update their subjective probability based on the information obtained, and the actual insurance decision is made based on this *ex-post* subjective probability. The actual decision whether or not to purchase insurance is based on the condition given by equation (6), which is analogous to equation (3) where *ex-ante* probability $\hat{p}$

[2] If the individual optimally chooses not to purchase any insurance $(I^* = 0)$, equation (1) suggests that $\mathrm{EU}(Search) = \hat{p}U(W - L) + (1 - \hat{p})U(W)$, which is linearly decreasing function of $\hat{p}$.
[3] For other parameter values, we set $W = 100$, $L = 50$, $r = 1/3$, $\mathrm{U}(x) = x^{1-\theta}/(1-\theta)$ and $\theta = 1/2$.

is replaced with updated, *ex-post* probability $\hat{p}'$.

Our model can potentially explain the stylized fact that consumers often fail to purchase insurance even if it is offered at favorable (actuarially fair or even subsidized) premiums. If the search cost C is high and the *ex-ante* subjective probability $\hat{p}$ is perceived to be low, consumer will optimally choose not to purchase insurance without trying to update their (presumably underestimated) probability estimates.

Our model also has some important implications that help to understand the consumer's behavior of insurance purchase and information search. First, consumer's subjective estimate of a loss probability $(\hat{p})$ is an important determinant for not only insurance demand but also whether or not to conduct a search. As discussed earlier, Figure 1 suggests that consumers with relatively high or low $\hat{p}$ tend not to search for additional information, while those with moderate $\hat{p}$ tend to search for information. At the same time, as the subjective probability estimate gets larger, the individual are more likely to purchase insurance (see equation (3)). Note that, for consumers not seeking information, insurance decision is based on their *ex-ante* subjective probability, while for those seeking information actual decision is made based on the *ex-post* probability.

Second, search cost (C) influences the outcome regarding insurance purchase and information search. With higher C, $\mathrm{EU}(Search)$ is shifted downward, while $\mathrm{EU}(No\,Search)$ is held constant since it is independent of C. This is illustrated in Figure 2. As C gets larger, the region of $\hat{p}$ where $s = 1$ is optimally chosen becomes smaller, and the region with $s = 0$ on the both sides becomes larger.

(Figure 2 around here)

Third, our model also provides several stylized predictions regarding insurance demand, although most of them are common to other much simpler models and are not specific to ours. For example, the individual is more likely to purchase insurance as L increases. In comparison, if the utility function exhibits decreasing absolute risk aversion (DARA), the individual is less likely to purchase insurance as W increases.[4]

In the following empirical analysis, we test these theoretical predictions using observed insurance behavior of potential policy holders. Specifically, we will examine the relationship between consumer's subjective probability and insurance decision. Our primary interest is about the effect of additional information on consumer's subjective probability, i.e., changes in *ex-ante* and *ex-post* subjective probabilities.

3. Data and Variables

In this paper, we use the 2011 Post-Quake Supplement (PQS) of the Keio Household Panel Survey (KHPS) and the Japan Household Panel Survey (JHPS). KHPS and JHPS are nationally representative panel surveys launched in 2004 and 2009, respectively, with initial panel sizes of approximately 4,000 households each. The two surveys essentially share the same questionnaire and are conducted every January. In 2011, a total of

[4] Differentiating the right-hand side of equation (3) yields

$$\frac{r(1-r)U'(W)U'(W-L)}{[(1-r)U'(W-L)+rU'(W)]^2} \times \left[\frac{U''(W)}{U'(W)} - \frac{U''(W-L)}{U'(W-L)}\right].$$

If the utility function exhibits DARA, the second term becomes positive, indicating that threshold increases as W rises.

6,000 households participated in these surveys (N = 3,030 for the KHPS and 3,160 for the JHPS).

In the wake of the Great East Japan Earthquake, KHPS and JHPS conducted an extensive supplemental survey on the post-quake situations of respondents. The PQS survey was conducted twice in 2011, once in June and again in October as a follow-up to the regular survey. Each 2011 KHPS/JHPS respondent was mailed a self-administered questionnaire. In the first-round PQS, 4,215 households replied for a response rate of 68.1%. In the second-round PQS, which follows up on first-round respondents, 3,591 households replied for a response rate of 85.2%. These supplementary surveys focused on the degree and extent of damage caused by the earthquake and the post-disaster situations of the respondents, including employment, housing, consumption and income. The PQS posted an especially wide range of questions about respondent’s insurance decision and disaster mitigation activities.

As for insurance decision, respondents are asked to select from one of the following options: (1) Already covered by earthquake insurance prior to the Great East Japan Earthquake; (2) Not covered but plan to purchase it in the future, or (3) Not covered and do not plan to purchase it in the future. In the benchmark case, we focus on potential policy holders (i.e., those not insured at the time of the survey), and examine their intention to purchase insurance after the March 11 earthquake. We define a dummy variable taking the value of one if the respondent plans to purchase insurance in the future. Those who were already insured at the time of the survey are excluded from the estimation sample.

The second-round PQS also included questions regarding whether or not the respondent obtained the regional hazard information such as earthquake hazard map provided by the local governments. Our maintaining assumption is that regional hazard information can serve as external risk information for potential insurance customers which may alter their *ex-ante* subjective probability.

An important feature of the PQS data is that it contains a wide range of geographic information about the objective earthquake risk. This information includes objective risk measures such as the likelihood of massive earthquake, distance from the coastline, and site liquefaction index. These geographic risk measures are assigned to PQS respondents based on the exact location of their residence.

Information about the likelihood of a massive earthquake is based on the Probabilistic Seismic Hazard Map (PSHM) provided by the National Research Institute for Earth Science and Disaster Prevention (NIED).[5] The PSHM data provides the probability that earthquakes with a given seismic intensity will take place in the next 30 years. In the following analysis, we use the 30-year probability of earthquakes with ground motions equal to or larger than JMA seismic intensity 6^{-} as our objective probability measure.[6] The original PSHM data is provided with the ESRI Grid format where geographic space is partitioned into an equally sized (approximately $1\text{km} \times 1\text{km}$) square grid cells. Estimated probabilities at each location are matched with the

[5] The original data is available at http://www.j-shis.bosai.go.jp/.

[6] The Japan Meteorological Agency (JMA) seismic intensity scale, which is measured with a seismic intensity meter, and is graded from 0 to 7, provides a measure of the strength of seismic motion. The typical situations and damages caused by the earthquake with JMA seismic intensity 6^{-} are as follows: People have difficulty standing, wooden houses occasionally collapse, and walls and pillars may be damaged even in highly earthquake-resistant houses. For full explanation of the JMA seismic intensity scale, see http://www.jma.go.jp/jma/kishou/know/shindo/explane.html. In general, the relationship between the JMA scale and the Richter scale basically depends on the distance from the epicenter. Even an earthquake with a small intensity on the Richter scale can have a large JMA intensity at locations near the epicenter.

PQS based on the exact location of each respondent.

In addition, the distance from the coastline (less than 250m and 250 to 500m) is used to capture the potential tsunami risk. The site liquefaction index is used to capture the potential risk of site liquefaction due to earthquake and it takes four distinct values with one the safest and three the riskiest.

Annual insurance premiums (per 1,000 JPY of coverage) are calculated based on the "standard rates" for wooden housing without any anti-seismic construction method, which is obtained from the website of the Non-Life Insurance Rating Organization of Japan (NILRO). In Japan, the earthquake insurance rates are determined by the government regulation, and are same across insurance companies. These are determined based on the estimates of the likelihood of occurrence and the expected damages of an earthquake. The likelihood of occurrence is estimated using the PSHM data discussed above. The insurance rates are set at the geographical level of prefectures, and each prefecture is classified into one of the four rating zones (rank 1 (safest) to rank 4 (riskiest)). As a result, for typical housing, the insurance premiums per 100,000 JPY of coverage range from a low of 100 JPY in the safest zone to a high of 313 JPY in the riskiest zone.

For respondent/household characteristics, we control for respondent's age, sex (1 if female), marital status (1 if married), household size, annual income, household wealth (sum of savings and financial assets), self-reported housing value, whether respondent's household repays any mortgages, floor level of the respondent's dwelling, and fear/anxiety toward possible aftershocks. The last variable (fear/anxiety toward aftershocks) is measured in self-rating scales (from 0 to 100). These respondent/household characteristics other than the last one are extracted from the main body of the survey conducted January 2011, thereby are measured prior to March 11 earthquake.

Since earthquake insurance is not common among renters, we restrict our sample to existing homeowners. Descriptive statistics are summarized in Table 1.

(Table 1 around here)

4. Empirical Analysis

4.1. Empirical Model

Let a be the dummy variable that takes the value of one if the respondent plans to purchase insurance in the future. From equations (3) and (6), the consumer's decision of purchase insurance, for those with or without search, can be described as:

$$a = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} \text{ if } \hat{p}(s) \begin{Bmatrix} < \\ \geq \end{Bmatrix} g(r, W, L), \tag{7}$$

where s is the dummy variable that takes the value of one if the respondent obtained the regional hazard information and zero otherwise, r is insurance premium, W is household wealth, and L is the expected loss proxied by the housing value. As for $\hat{p}(s)$, $\hat{p}(s=0)$ corresponds to the *ex-ante* average subjective probability of consumers without information search $(\hat{p})$, while $\hat{p}(s=1)$ corresponds to the *ex-post* subjective probability of consumers obtaining additional information $(\hat{p}')$.

If the subjective probability were known, the above specification would imply a standard probit model. However, the subjective probability is not known and needs to be modeled in some way. We assume that $\hat{p}(s)$ takes the following form:

$$\hat{p}(s) = f(p^o, x) + \beta s + \varepsilon, \tag{8}$$

where p^o is the objective probability measure of massive earthquake obtained from the PSHM data, and x is a set of respondent's characteristics. Plugging equation (8) into (7), together with linearizing g and f, yields our empirical model.

Our primary interest is to estimate the causal effect of obtaining additional information on the consumer's *ex-post* subjective probability (compared with the *ex-ante* value), which is represented by β in equation (8). However, as discussed in Section 2, searching for risk information itself depends on $\hat{p}$. Hence if there are omitted variables in equation (8), s would correlate with ε, leading to inconsistent coefficient estimates. To cope with this problem, we adopt a standard instrumental variables approach. In constructing instrumental variables for s, we utilize information whether paper- or web-based earthquake hazard information is available in the respondent's municipality.[7] Specifically, we create two dummy variables indicating that hazard information is available on paper/web basis, and the latter variable is interacted with another dummy variable which takes the value of one if respondent family has internet access at home. Since these variables are associated with consumer's search cost (C), they are likely to be correlated with s while independent of subjective probability $\hat{p}$.

4.2. Empirical Results

Before presenting the main empirical results, we briefly discuss some descriptive statistics to see whether these observations are consistent with our theoretical predictions. The upper panel of Table 2 presents sample means for selected objective risk measures by the observed combination of insurance and search outcomes. As shown in Figure 1, compared with consumers searching for additional information, those who purchase insurance without incurring any search would on average have higher $\hat{p}$, and those who do not purchase insurance without search would have the lower $\hat{p}$. Hence, provided that consumer's subjective probability is positively associated with objective measures, consumers who purchase insurance without incurring any search are likely to be in riskier regions. Summary statistics presented in Table 2 are basically consistent with this prediction. For example, the average PSHM probability is highest for respondents who plan to purchase insurance without searching regional hazard information. Other risk measures show similar patterns.

The bottom panel of Table 2 shows the relationship between search cost and observed insurance and search behavior. Consistent with results in Figure 2, in municipalities where web-based risk information is available (i.e., lower search cost), respondents are more likely to search for and obtain the regional hazard information. Furthermore, the proportion of respondents who plan to purchase insurance without a search is also smaller in such municipalities.

---

7 The original information is available at the website of the Ministry of Land, Infrastructure, Transport and Tourism (http://disapotal.gsi.go.jp/index.html).

(Table 2 around here)

Our benchmark regression results are summarized in Table 3. The sample is existing homeowners without earthquake insurance at the time of the survey. The dependent variable takes the value of one if the respondent plans to purchase insurance in the future, and zero otherwise. In this benchmark case, we assume that the information search behavior (s) is exogenous, and we use a standard probit model.

(Tables 3 and 4 around here)

From Table 3, it is found that homeowners who obtained the regional hazard information are significantly more likely to purchase insurance. Interpreting our coefficient estimate (0.1998) in terms of average marginal effect shows that homeowners with regional hazard information are about 7.4% more likely to purchase insurance (average marginal effect = 0.0744 with $p = 0.023$). Given that the sample mean of the dependent variable is about 0.44, this is fairly large effect.

However, as discussed earlier, a standard probit model is subject to bias stemming from the endogeneity of information search. In order to address the endogeneity problems, we estimate the model based on probit model with binary endogenous regressor (Table 4). First-stage results in Table 5 show that our instruments are significantly strongly correlated with observed search behavior with the expected sign.

(Tables 4 and 5 around here)

Estimation results in Table 4 show that the coefficient of obtaining regional hazard information becomes considerably smaller, but still significantly positive at 5% level. The positive coefficient estimate of 0.078 implies that obtaining additional risk information, on average, increases the consumer's subjective probability of a loss by about 8 percentage points. As a result, homeowners with regional hazard information are about 2.9% more likely to purchase insurance (average marginal effect = 0.0291 with $p = 0.038$).

For other explanatory variables, our empirical findings are as follows. The level of 30-year earthquake probability has insignificant coefficient. This result seems somewhat puzzling since, given the specification of $\hat{p}$ in equation (8), insignificant coefficient suggests that $\hat{p}$ is not associated with objective probability. The potential explanation for this result is given as follows. As discussed in Naoi et al. (2010), insurance rates in Japan are determined based on extremely rough and crude geographical rating categories. As a result, even if two potential policy holders are facing the same objective risk, insurance rates can be different depending on their location of residence. In this case, aggregate bias in insurance pricing may influence the coefficient estimate of the 30-year probability.

Household with larger wealth is less likely to have insurance purchase plan, which is plausible if the absolute risk aversion decreases as wealth rises. In comparison, self-reported housing value is positively associated with insurance purchase.

Higher insurance premiums (per 1,000 JPY of coverage) are shown to be negatively associated with the insurance purchase (especially for homeowners of condominium unit). This result seems to be reasonable at first glance. However, if insurance premiums are actuarially fair, premium levels will not have any direct impact on insurance decision. Hence this suggests that there are some biases in insurance pricing.

Estimation results for site liquefaction and distance from the coastline suggest that people in risky location tend to plan to purchase insurance. But the estimated impacts are in general weak and insignificant. Self-reports of fear and anxiety are strongly and positively associated with the intention to purchase insurance.

5. Conclusion

In this paper, we develop a simple model of insurance purchase where consumers have imperfect knowledge about the potential risk based on Kunreuther and Pauly (2005) and test its theoretical implications using observed insurance behavior of potential policy holders. Our empirical analysis indicates that obtaining the regional disaster hazard information makes the consumer's subjective probability of a loss significantly higher, thereby facilitating insurance demand. This result is robust to the possible endogeneity of consumer's information search behavior.

References

Ehrlich, I. and G.S. Becker (1972) "Market Insurance, Self-Insurance, and Self-Protection," *Journal of Political Economy*, 80(4), 623–648.

Fujimi, T. and R. Kakimoto (2012) "Investigation on Factors to Bias Perceived Risk of House Destruction." Proceedings of the 2012 Annual International Conference of the International Institute for Infrastructure, Renewal and Reconstruction (IIIRR). Available at http://iiirr.ucalgary.ca/files/iiirr/B6-5_.pdf.

Kunreuther, H. (1984) "Causes of Underinsurance Against Natural Disasters," *Geneva Papers on Risk and Insurance*, 9(31), 206–220.

Kunreuther, H. and M. Pauly (2004) "Neglecting Disaster: Why Don't People Insure Against Large Losses?," *Journal of Risk and Uncertainty*, 28(1), 5–21.

Naoi, M., M. Seko and K. Sumita (2010) "Community Rating, Cross Subsidies and Underinsurance: Why So Many Households in Japan Do Not Purchase Earthquake Insurance," *Journal of Real Estate Finance and Economics*, 40(4), 544–561.

Slovic, P., B. Fischhoff and S. Lichtenstein (1986) "Rating the Risks: The Structure of Expert and Lay Perceptions," in Covello, V.T. (eds.), *Environmental Impact Assessment, Technology Assessment, and Risk Analysis*, New York: Springer-Verlag.

Table 1: Descriptive Statistics

	Mean	(S.D.)
Plan to purchase insurance in the future	0.4365	(0.4962)
Obtained the regional hazard information (yes = 1)	0.6391	(0.4805)
30-year probability of massive earthquake (PSHM, JMA 6-)	0.2727	(0.2621)
Age	54.204	(13.439)
Sex (female = 1)	0.4950	(0.5002)
Marital status (married = 1)	0.8236	(0.3814)
Household size	3.2006	(1.3323)
Annual income (in 10,000 JPY)	691.14	(455.49)
Household wealth (in 10,000 JPY)	1461.50	(2133.40)
Self-reported housing value (in 10,000 JPY)	893.06	(905.32)
Repay any mortgages (yes = 1)	0.6522	(0.4765)
Floor level	2.5716	(1.9982)
Insurance premium (per 1,000 JPY of coverage)	1.9709	(0.8447)
Risk assessment of site liquefaction		
Risk assessment = 0 (safest)	(Omitted category)	
Risk assessment = 1	0.1220	(0.3274)
Risk assessment = 2	0.4012	(0.4904)
Risk assessment = 3 (riskiest)	0.0111	(0.1048)
Distance from the coastline		
less than 250m	0.0181	(0.1335)
250 to 500m	0.0232	(0.1506)
Fear/anxiety toward possible aftershocks (self-rating, 0-100)	64.516	(29.533)
N	992	

Table 2: Descriptive Statistics by Insurance and Search Outcomes

		Not purchase insurance without search ($a = 0, s = 0$)	Search for risk information ($s = 1$)	Purchase insurance without search ($a = 1, s = 0$)
30-year probability of massive earthquake (PSHM, JMA 6-)		0.2651	0.2705	0.2970
Site liquefaction index (risk assesment = 2 or 3)		0.0196	0.0196	0.0226
Distance from the coastline (less than 250m)		0.4118	0.4099	0.4323
Floor level		2.4830	2.6061	2.6207
Any web-based hazard information available?				
Yes	N	231	525	168
	(%)	(25.00)	(56.82)	(18.18)
No	N	126	241	98
	(%)	(27.10)	(51.83)	(21.08)

Table 3: Estimation Results for Standard Probit Model

Dependent variable: Plan to purchase insurance in the future (yes = 1)	[1] Coef.	(S.E.)		AME
Obtained the regional hazard information (yes = 1)	0.1998	(0.0885)	*	0.0744
30-year probability of massive earthquake (*p*)	-0.0669	(0.1769)		-0.0249
Age/10	-0.0476	(0.0398)		-0.0177
Sex (female = 1)	0.0953	(0.0854)		0.0355
Marital status (married = 1)	-0.1617	(0.1177)		-0.0602
Household size	-0.0123	(0.0372)		-0.0046
Annual income (in 10 million JPY)	0.0416	(0.0979)		0.0155
Household wealth (in 10 million JPY)	-0.1048	(0.0493)	*	-0.0299
Household wealth^2	0.0086	(0.0045)	+	
Self-reported housing value (in 10 million JPY)	0.1777	(0.0971)	+	0.0407
Self-reported housing value^2	-0.0386	(0.0205)	+	
Repay any mortgages (yes = 1)	-0.1568	(0.1050)		-0.0591
Floor level	0.2863	(0.1049)	**	0.0777
Floor level^2	-0.0152	(0.0058)	**	
Insurance premium (per 1,000 JPY of coverage)				
Single-family detached house	-0.0348	(0.0546)		-0.0129
Apartment with reinforced concrete structure	-0.4052	(0.1697)	*	-0.1211
Risk assessment of site liquefaction				
Risk assessment = 0 (safest)				
Risk assessment = 1	0.1221	(0.1348)		0.0455
Risk assessment = 2	0.0177	(0.0991)		0.0066
Risk assessment = 3 (riskiest)	0.4303	(0.4503)		0.1603
Distance from the coastline				
less than 250m	0.3456	(0.3049)		0.1288
250 to 500m	0.3351	(0.3084)		0.1249
Fear/anxiety toward possible aftershocks (self-rating, 0-100)	0.0051	(0.0015)	**	0.0019
Log likelihood		-645.691		
N		992		

Notes: **, *, and + indicate that the estimated coefficient is significant at the 0.01, 0.05 and 0.10 levels, respectively. Dummy variables for housing types (single-family detached/condo.), city-sizes, region and type of survey (JHPS = 1) are also controlled but are omitted from the results. Robust standard errors are reported in the parentheses.

Table 4: Estimation Results for Probit Model with Binary Endogenous Regressor

Dependent variable: Plan to purchase insurance in the future (yes = 1)	[2]			
	Coef.	(S.E.)		AME
Obtained the regional hazard information (yes = 1)	0.0780	(0.0379)	*	0.0291
30-year probability of massive earthquake (*p*)	-0.0686	(0.1811)		-0.0256
Age/10	-0.4694	(0.4043)		-0.0175
Sex (female = 1)	0.1071	(0.0868)		0.0400
Marital status (married = 1)	-0.1641	(0.1188)		-0.0613
Household size	-0.0150	(0.0375)		-0.0056
Annual income (in 10 million JPY)	0.0372	(0.0985)		0.0139
Household wealth (in 10 million JPY)	-0.0909	(0.0503)	+	-0.0261
Household wealth^2	0.0074	(0.0045)		
Self-reported housing value (in 10 million JPY)	0.1545	(0.0995)		0.0328
Self-reported housing value^2	-0.0378	(0.0209)	+	
Repay any mortgages (yes = 1)	-0.1546	(0.1069)		-0.0584
Floor level	0.3139	(0.1068)	**	0.0857
Floor level^2	-0.0165	(0.0059)	**	
Insurance premium (per 1,000 JPY of coverage)				
Single-family detached house	-0.0274	(0.0556)		-0.0101
Apartment with reinforced concrete structure	-0.4112	(0.1709)	*	-0.1216
Risk assessment of site liquefaction				
Risk assessment = 0 (safest)				
Risk assessment = 1	0.1015	(0.1375)		0.0379
Risk assessment = 2	0.0068	(0.1007)		0.0025
Risk assessment = 3 (riskiest)	0.2711	(0.4760)		0.1013
Distance from the coastline				
less than 250m	0.3297	(0.3051)		0.1232
250 to 500m	0.4192	(0.3216)		0.1566
Fear/anxiety toward possible aftershocks (self-rating, 0-100)	0.0050	(0.0015)	**	0.0019
Log likelihood		-621.702		
N		953		

Notes: **, *, and + indicate that the estimated coefficient is significant at the 0.01, 0.05 and 0.10 levels, respectively. Dummy variables for housing types (single-family detached/condo.), city-sizes, region and type of survey (JHPS = 1) are also controlled but are omitted from the results. Robust standard errors are reported in the parentheses.

Table 5: Estimation Results for Probit Model with Binary Endogenous Regressor (First-Stage)

Dependent variable: Obtained the regional hazard information (yes = 1)	[3] Coef.	(S.E.)		AME
Paper-based hazard information available (yes = 1)	0.3666	(0.1618)	*	0.1286
Web-based hazard information available (yes = 1) × Internet access available (yes = 1)				
Web-based hazard info. = 0, Internet = 0	(Omitted Category)			
Web-based hazard info. = 0, Internet = 1	0.1400	(0.1816)		0.0513
Web-based hazard info. = 1, Internet = 0	0.2178	(0.1727)		0.0791
Web-based hazard info. = 1, Internet = 1	0.2928	(0.1714)	+	0.1052
Log likelihood	-586.066			
N	953			

Notes: * and + indicate that the estimated coefficient is significant at the 0.05 and 0.10 levels, respectively. The same set of exogenous variables as in Tables 3 and 4 are also controlled but are omitted from the results. Robust standard errors are reported in the parentheses.

Figure 1: Optimal Insurance and Search Decision

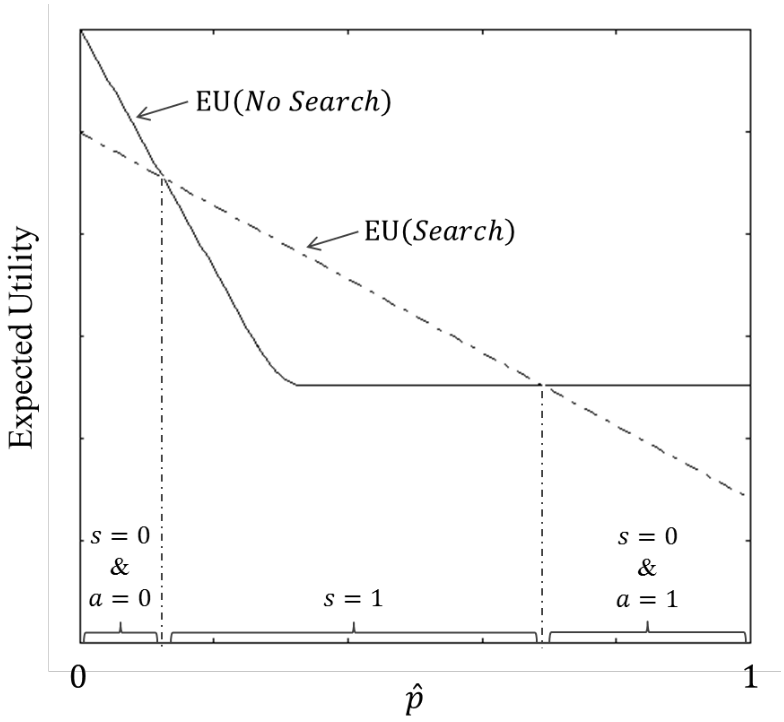

Figure 2: Effect of Search Cost on Insurance and Search Decision