Entropy Characterisation of Insurance Demand: Theory and Evidence*

Hiroyuki Nakata†, Yasuyuki Sawada‡ and Mari Tanaka§

June 11, 2010

Abstract

This paper characterises the insurance demand in terms of the entropy of the underlying probability distribution for losses. A characterisation of this nature provides the prediction that insurance for large losses with small probabilities tends to be purchased less frequently than insurance for moderate losses with higher probabilities, without deviating from the standard expected utility framework. The predictions of the theoretical model are tested empirically using household data collected in Vietnam.

Keywords: Entropy, Large Deviation, Relative Entropy, Subjective Probability, Willingness-to-Pay.

*We acknowledge financial support from the Research Institute for Economy, Trade, and Industry (RIETI). We would like to thank Dang Kim Son, Nguyen Ngoc Que, Truong Thi Thu Trang, Phung Duc Tung for their invaluable contributions for the implementation of the household survey. We would also like to thank Sarah Bales, Bob Baulch, Paul Glewwe, and the participants at the conference at RIETI, Masahisa Fujita in particular, for their useful suggestions and comments. The usual disclaimers apply.

†Essex Business School, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom; E-mail: hnakata@stanfordalumni.org or hnakata@essex.ac.uk.

‡Faculty of Economics, University of Tokyo.

§Graduate School of Economics, University of Tokyo.

1 Introduction

How can one be best prepared for large losses (catastrophes) with small probabilities? Should we purchase insurance against such losses, if such a policy is available? A positive answer to this classic question appears plausible, since reserves or savings alone are typically insufficient to cover large losses. However, there is evidence that insurance for catastrophes such as earthquakes or flooding is not very widely purchased, even though many policies against catastrophes are subsidised by governments to keep the premium favourable to the buyers, e.g. earthquake insurance in Japan and the National Flood Insurance Program (NFIP) in the United States.[1] In contrast, it is widely known that commonly sold insurance policies such as travel insurance, home insurance and medical insurance, have a substantial mark-up (loading factor). Yet, many people voluntarily elect to purchase such policies. Why? It is rather tempting to claim that this is yet more evidence of behaviourism prevailing. For instance, one may argue that people assigns lower weights on low probability events by referring to the cumulative prospect theory by Kahneman and Tversky (1979, 1992).[2] Before accepting that claim, however, we make further observations about the demand for catastrophe insurance.

First, as stated above, insurance for catastrophes is purchased with less frequency compared to insurance for moderate risk (e.g. travel insurance), even if the premium is often set more favourably for catastrophe insurance (Stylised Fact 1 hereafter). On top of Stylised Fact 1, the following stylised facts regarding insurance for catastrophes are reported largely based on aggregate data.

- The market penetration is much lower in areas that have historically been less frequently hit by catastrophes, even if the premium is adjusted to reflect the lower frequencies. (Stylised Fact 2)
- The market penetration jumps up immediately after a catastrophe. (Stylised Fact 3)

Dixon et al. (2006) collects a panel data of NFIP, and estimates that market penetration is only one percent outside the Special Flood Hazard

[1]Kunreuther et al. (1978) is one of the pioneering works that reported this 'anomaly'. Moreover, there is long history of studies on the structure of the insurance market and/or insurance demand; earlier studies include Arrow (1963), Borch (1962), Ehrlich and Becker (1972), Mossin (1968), and Smith (1968).

[2]There are some existing studies that are in support of the standard expected utility framework, too—e.g. Brookshire et al. (1985), and Kunreuther and Pauly (2004).

Area (SFHA), where flooding takes place more frequently, while it is about fifty percent within the SFHA. In Japan, the market penetration rate of the earthquake insurance is substantially higher (than the national average) in the three prefectures that are considered most at risk, although the premiums for earthquake insurance there are the most expensive in the country. Specifically in March 2006, 27.9 percent in Tokyo, 26.6 percent in Kanagawa and 24.8 percent in Shizuoka, as opposed to 7.7 percent in Okinawa, 7.2 percent in Nagasaki, and 8.7 percent in Yamagata, where there is no history of severe losses caused by earthquakes (data source: the General Insurance Association of Japan). These are clearly consistent with Stylised Fact 2.

Earthquake insurance in Japan and the flooding insurance in the US are also consistent with Stylised Fact 3. For example, market penetration in Hyogo prefecture in Japan, which was hit by the Hanshin-Awaji earthquake in January 1995, was 2.9 percent in March 1994, rising to 8.4 percent in March 1996. Also, in Miyagi prefecture, which experienced a major earthquake in May 2003 and another one in August 2005, market penetration was 16.7 percent in March 2003, jumping to 20.5 percent in March 2004, and 25.9 percent in March 2006. On the other hand, the number of total flooding insurance policies in the US was 4.458 million in 2001, gradually increasing to 4.667 million by 2004. However, the number jumped to 4.956 million by 2005, and then to 5.394 million by April 2007 (data source: the Federal Emergency Management Agency, FEMA). Note that 2005 was the year Hurricane Katrina caused unprecedented losses, and the recent surge in the number of total policies appears consistent with Stylised Fact 3. Also, using the US data, Browne and Hoyt (2000) found that flood insurance purchases at the state level are highly correlated with the level of the flood losses in the state during the prior year.

This paper attempts to explain the three stylised facts above simultaneously within the standard expected utility framework. In so doing, we allow for heterogeneous beliefs, and show that heterogeneity of beliefs has a major impact on insurance demand. In particular, we characterise insurance demand in terms of entropy of the loss probability distribution. We then test the theoretical predictions empirically by using a unique household survey data from Vietnam, which is a resurvey of subsamples of the Vietnam Household Living Standards Survey (VHLSS) 2006 households. The data set includes subjective probability assessments of losses from Avian Influenza (AI), and from flooding, willingness-to-pay for insurance for AI losses and/or flooding, past loss experience, past experience about insurance, as well as basic household information such as income, wealth, and education level.

The rest of the paper proceeds as follows. In the next section, a simple static model is introduced first to define the willingness-to-pay for insurance, and then basic results concerning willingness-to-pay as well as issues regarding rationality are examined. Section 3 examines the theoretical predictions presented in section 2 empirically by using the household data collected in Vietnam. Section 4 concludes the paper.

2 Theoretical Framework

In this section, we first characterise insurance demand in terms of the entropy of the underlying loss probability, while maintaining the subjective expected utility framework. We then examine the question of rationality of subjective probability using the elementary results of the large deviation theory.

2.1 Insurance Demand

Consider an agent who is facing some uncertainty: he may incur a loss. Let a_s denote the amount of loss the agent will suffer in state s ($s = 1, 2, ..., S$), and let W denote the initial wealth. Assume without loss of generality that $a_s > a_{s'}$ for all $s > s'$ and $a_1 = 0$, i.e. state 1 is the no-loss state. The final wealth in state s then becomes $W - a_s$ when there is no insurance.

Assume that the agent is a risk-averse expected utility maximiser, who makes some probability estimate $q = (q_1, q_2, ..., q_S)$, where q_s denotes the agent's subjective probability (estimate) of state s. Now, assume that any loss up to a_k can be covered by insurance with a premium of ρ_k. With this insurance, the final wealth becomes

$$W - a_s - \rho_k = \begin{cases} W - \rho_k & \text{if } s \leq k; \\ W - a_s + a_k - \rho_k & \text{otherwise.} \end{cases}$$

Agent h purchases the insurance if

$$\sum_{s=1}^{S} q_s u(W - a_s) \leq \sum_{s=1}^{k} q_s u(W - \rho_k) + \sum_{s=k+1}^{S} q_s u(W - a_s + a_k - \rho_k),$$

while the agent is indifferent between purchasing and not purchasing when equality holds. This observation leads us to define the agent's willingness-to-pay for this insurance as $\hat{\rho}_k^q$, satisfying the following equation:

$$\sum_{s=1}^{S} q_s u(W - a_s) = \sum_{s=1}^{k} q_s u(W - \hat{\rho}_k^q) + \sum_{s=k+1}^{S} q_s u(W - a_s + a_k - \hat{\rho}_k^q).$$

It is easy to check that the willingness-to-pay $\hat{\rho}_k^q$ is strictly concave in the probability estimate q as well as in the maximum coverage a_k, as long as the von Neumann-Morgenstern utility function u is strictly concave, i.e. risk averse.

Suppose now that the insurance is offered at an actuarially fair premium with respect to a probability vector $\pi = (\pi_1, \pi_2, ..., \pi_S)$, where π_s is the probability of state s. The fair premium with respect to π is given by

$$\bar{\rho}_k^\pi := \sum_{s=1}^{k} \pi_s a_s + \sum_{s=k+1}^{S} \pi_s a_k.$$

It is straightforward that the agent's willingness-to-pay for the insurance exceeds $\bar{\rho}_k^\pi$ if his subjective probability coincides with π. However, it may be the case that the agent's willingness-to-pay is smaller than $\bar{\rho}_k^\pi$ when his subjective probability does not coincide with π.

To see this, let f_s denote the conditional probability of state s given that it is a loss state (i.e. $s \neq 1$). Hence, $\sum_{s=2}^{S} f_s = 1$. Then, without loss of generality, we can write

$$\pi_s = (1 - \pi_1) f_s, \quad \forall s \neq 1.$$

Now, assume that the agent's subjective probability vector q is given as follows. For some $\varepsilon > 0$,

$$\begin{aligned} q_1 &= \pi_1 + \varepsilon; \\ q_s &= (1 - q_1)\, f_s = (1 - \pi_1 - \varepsilon) f_s, \quad \forall s \neq 1, \end{aligned}$$

Namely, π and q have the same conditional probability of state s given that it is not the loss state.

Observe that the difference between the agent's subjective probability and π is completely represented by ε. Moreover, q dominates π in the sense of first order stochastic dominance when $\varepsilon > 0$. Note that $\varepsilon \leq 1 - \pi_1$ since

$$q_s = (1 - \pi_1 - \varepsilon) f_s \geq 0.$$

In contrast with $\bar{\rho}_k^\pi$, the subjectively fair premium from the agent's perspective is defined as follows.

$$\bar{\rho}_k^q := \sum_{s=1}^{k} q_s a_s + \sum_{s=k+1}^{S} q_s a_k.$$

It follows that

$$\begin{aligned}\bar{\rho}_k^\pi - \bar{\rho}_k^q &= \sum_{s=1}^{k}(\pi_s - q_s)a_s + \sum_{s=k+1}^{S}(\pi_s - q_s)a_k \\ &= \varepsilon \sum_{s=2}^{k} f_s a_s + \varepsilon a_k \sum_{s=k+1}^{S} f_s, \quad \forall \varepsilon > 0.\end{aligned}$$

Hence, $\bar{\rho}_k^q < \bar{\rho}_k^\pi$ and the difference between $\bar{\rho}_k^q$ and $\bar{\rho}_k^\pi$ is independent of π_1, and is linear in ε. Moreover, it is clear that $\hat{\rho}_k^q > \bar{\rho}_k^q$ since the agent is assumed to be risk averse. Hence, it is possible that the agent does not purchase the insurance at the premium of $\bar{\rho}_k^\pi$, but purchases at $\bar{\rho}_k^q$, i.e. $\bar{\rho}_k^q < \hat{\rho}_k^q < \bar{\rho}_k^\pi$. Clearly, this becomes more likely when the discrepancy between the agent's subjective probability q and π is greater.

Observe that $\hat{\rho}_k^q - \bar{\rho}_k^q$ is the agent's subjective risk premium, which is a function of q_1. Let $r_k^q(q_1) := \hat{\rho}_k^q(q_1) - \bar{\rho}_k^q(q_1)$. Since $u(\cdot)$ is assumed to be strictly concave, it is easy to show that $\frac{d^2 r_k^q}{dq_1^2}(q_1) < 0$ for all q_1, and there is a critical value $\hat{q}_1$ such that $\frac{dr^q}{dq_1}(\hat{q}_1) = 0$, while $r_k^q(1) = 0$ (in which case, $\varepsilon = 0$ must hold, too). Hence, $r_k^q(q_1)$ is decreasing in q_1, when q_1 is sufficiently large. It follows that for $\varepsilon > 0$ given, r_k^q is also decreasing in π_1, when π_1 is sufficiently large.

Recall that $\bar{\rho}_k^\pi - \bar{\rho}_k^q$ is independent of π_1. Notice also that

$$\hat{\rho}_k^q - \bar{\rho}_k^\pi = r_k^q - (\bar{\rho}_k^\pi - \bar{\rho}_k^q).$$

It follows that $\hat{\rho}_k^q - \bar{\rho}_k^\pi$ is decreasing in π_1 when π_1 is sufficiently large. Also, since $r_k^q = 0$ holds when $q_1 = 1$, for sufficiently small $\varepsilon > 0$ and sufficiently large π_1, and that, for every k, there exists a critical value $p_1(k, \varepsilon)$ such that

$$\hat{\rho}_k^q \begin{cases} < \bar{\rho}_k^\pi & \text{if } \pi_1 > p_1(k,\varepsilon), \\ = \bar{\rho}_k^\pi & \text{if } \pi_1 = p_1(k,\varepsilon), \\ > \bar{\rho}_k^\pi & \text{otherwise.} \end{cases}$$

Hence, for a given $\varepsilon > 0$, it is more likely that the agent does not purchase insurance that is actuarially fair with respect to π when π_1 is large. In other words, given $\varepsilon > 0$, the impact of discrepancy between q and π becomes more significant when the probability of the no-loss state is larger and loss events are rarer.

Note that we focus on the case $\varepsilon > 0$, because when $\varepsilon < 0$, it is trivial that $\bar{\rho}_k^{\pi} < \bar{\rho}_k^{q} < \hat{\rho}_k^{q}$ holds. Namely, the agent purchases the insurance irrespective of the discrepancy in the probability estimates.

To further our analysis, we now introduce the notion of entropy, which is useful in characterising probability distributions.[3] The entropy of π is defined as follows.[4]

$$H(\pi) := -\sum_{s=1}^{S} \pi_s \ln \pi_s.$$

In our case, it is

$$\begin{aligned} H(\pi) &= -\pi_1 \ln \pi_1 - (1-\pi_1) \sum_{s=2}^{S} f_s \ln\left[(1-\pi_1) f_s\right] \\ &= -\pi_1 \ln \pi_1 - (1-\pi_1) \sum_{s=2}^{S} f_s \ln f_s - (1-\pi_1)\ln(1-\pi_1) \sum_{s=2}^{S} f_s \\ &= -\pi_1 \ln \pi_1 - (1-\pi_1) \sum_{s=2}^{S} f_s \ln f_s - (1-\pi_1)\ln(1-\pi_1). \end{aligned}$$

Differentiating $H(\pi)$ with respect to π_1, we obtain

$$\begin{aligned} \frac{\partial H(\pi)}{\partial \pi_1} &= -\ln \pi_1 - 1 + \sum_{s=2}^{S} f_s \ln f_s + \ln(1-\pi_1) + 1 \\ &= -\ln \frac{\pi_1}{1-\pi_1} + \sum_{s=2}^{S} f_s \ln f_s. \end{aligned}$$

It is straightforward that $H'(\pi) < 0$ holds when $\pi_1 > 0.5$, while it is the case even when π_1 is smaller than 0.5, depending on the value of $\sum_{s=2}^{S} f_s \ln f_s (\leq 0)$. Moreover, the second derivative of $H(\pi)$ with respect to π_1 is

$$\frac{\partial^2 H(\pi)}{\partial \pi_1^2} = -\frac{1}{\pi_1} - \frac{1}{1-\pi_1} = -\frac{1}{\pi_1(1-\pi_1)} < 0, \quad \text{for all } \pi_1.$$

Hence, it is clear that $H(\pi)$ is decreasing in π_1 when π_1 is sufficiently large.

The above observation leads us to claim the following proposition.

[3]See for example Cover and Thomas (1991) for the definition and properties of entropy as well as those of relative entropy.

[4]In the definitions of entropy and relative entropy below, $0 \ln 0 := 0$ and $0 \ln(0/0) := 0$.

Proposition 1: *For $\varepsilon > 0$ given and sufficiently large π_1, both $\hat{\rho}_k^q - \bar{\rho}_k^\pi$ and $H(\pi)$ are decreasing in π_1. In particular, there exists a critical value $p_1(k, \varepsilon)$ such that $\hat{\rho}_k^q \leq \bar{\rho}_k^\pi$ if and only if $H(\pi) \leq H(p_{k,\varepsilon})$, where $p_{k,\varepsilon}$ is the probability law such that*

$$\text{probability of state } s = \begin{cases} p_1(k, \varepsilon) & \text{if } s = 1, \\ [1 - p_1(k, \varepsilon)]\, f_s & \text{otherwise.} \end{cases} \qquad \blacksquare$$

This result suggests that for a given level of ε, it is more likely that the agent's willingness-to-pay $\hat{\rho}_k^q$ is smaller than the actuarially fair premium $\bar{\rho}_k^\pi$ with respect to π when the entropy of π is smaller. In other words, when the entropy of π is smaller, more agents would not purchase the insurance whose premium is actuarially fair with respect to π. Note that this is consistent with Stylised Fact 1 above.

Moreover, we claim that proposition 1 does not contradict with the classical results of Mossin (1968). In Mossin (1968), the agent is choosing k while taking the premium as given. Hence, when q and π coincide, and the premium is actuarially fair, the agent chooses $k = S$, i.e. a full cover. Or, when q and π coincide, but the premium has a positive loading factor, then the agent chooses $k < S$, i.e. a partial cover.

Before proceeding, we illustrate the impacts of differences in subjective probabilities and/or those in degrees of risk aversion on willingness-to-pay for insurance. The key observation is that the scale of the willingness-to-pay and that of the subjective probability are roughly proportionate even for risk averse agents.[5]

Example 1: Suppose there are only two states, i.e. the loss state and the no-loss state. We assume that the agents' preferences have an expected utility representation with constant relative risk-aversion von Neumann-Morgenstern utilities. We focus on two cases, in which the relative risk aversion (rra) is 1 or 2. The two tables below report the willingness-to-pay for full cover insurance.

Table 1 reports the case in which the final wealth is 20 in the no loss state and 10 in the loss state without insurance. Table 2 reports the case in which the final wealth is 100 in the no loss state and 10 in the loss state without insurance.

The results from the two tables indicate that when a very small change

[5] Willingness-to-pay is concave in subjective probability for risk averse agents, but the scale is roughly proportionate.

Table 1: WTP for full cover insurance: final wealth 20

Loss Probability	WTP (rra = 1)	WTP (rra = 2)
2^{-3}	1.659919136	2.222222222
2^{-7}	0.10801153	0.15503876
2^{-10}	0.01353345	0.019512195
2^{-20}	1.32207E-05	1.90735E-05
2^{-30}	1.29109E-08	1.86265E-08

Table 2: WTP for full cover insurance: final wealth 100

Loss Probability	WTP (rra = 1)	WTP (rra = 2)
2^{-3}	25.01057907	52.94117647
2^{-7}	1.782811081	6.569343066
2^{-10}	0.224609201	0.87124879
2^{-20}	0.000219591	0.0008583
2^{-30}	2.14445E-07	8.3819E-07

in subjective probability in absolute terms is actually a very large change in terms of proportions of probability, this change results in a large proportional change in the willingness-to-pay—e.g. the increase in subjective probability from 2^{-30} to 2^{-20} is miniscule in absolute terms, but it is a huge increase in terms of proportions, i.e. approximately 1000 times. Consequently, willingness-to-pay also increases approximately 1000 times. ■

Note that both 2^{-30} and 2^{-30} are very small probabilities. In the subsequent subsection, we examine under what conditions such small probabilities can be understood to be not contradicting with the empirical data. In so doing, we utilise elementary results from large deviation theory, which is a field in probability theory that studies the properties of very small probabilities. For instance, we know from the strong law of large numbers that the average of a sequence of i.i.d. random variables converges to the expected value of the random variable with probability one.[6] However, it is of interest to study the speed of convergence as well as probability of some finite sequence of realisations. Large deviation theory studies these aspects, and in what follows, we utilise the results in our context.

[6]The strong law of large numbers itself does not require identically distributed random variables, but just independent random variables.

2.2 Rationality

In what follows, we examine the issue of rationality. In the above, we introduced probability laws π and q in a static setup. If all agents and the insurance suppliers hold rational expectations, $\pi = q$ holds, since all agents and insurance suppliers know the true probability. However, there is no reason why the agents know the true probability law π a priori. In this case, all we can hope for is to examine if the subjective probability laws are compatible with the empirical data, i.e. test if the subjective probability cannot be rejected using any econometric or statistical methods.

Let random variable X_t denote the loss in period t, and let $X_1, X_2, ..., X_T$ be an i.i.d. sequence. Also, let $\mathcal{P}(\mathcal{A})$ denote the space of all probability laws on $\mathcal{A} := \{a_1, a_2, ..., a_S\}$. Furthermore, for a finite sequence (of realisations) $\mathbf{x}^T = (x_1, x_2, ..., x_T)$, we define the empirical measure of a_s as follows:

$$m_T^{\mathbf{x}}(a_s) := \frac{1}{T} \sum_{t=1}^{T} \mathbf{1}_{a_s}(x_t), \quad \forall s,$$

where $\mathbf{1}_{a_s}(\cdot)$ is an indicator function, i.e.

$$\mathbf{1}_{a_s}(x_t) = \begin{cases} 1 & \text{if } x_t = a_s, \\ 0 & \text{otherwise.} \end{cases}$$

Then, we define type $m_T^{\mathbf{x}}$ of $\mathbf{x}^T$ as

$$m_T^{\mathbf{x}} := \left(m_T^{\mathbf{x}}(a_1), m_T^{\mathbf{x}}(a_2), ..., m_T^{\mathbf{x}}(a_S) \right).$$

Let $\mathcal{M}_T$ denote the set of all possible types of sequences of length T, i.e.

$$\mathcal{M}_T := \left\{ \nu : \nu = m_T^{\mathbf{x}} \text{ for some } \mathbf{x}^T \right\}.$$

Also, the empirical measure $m_T^{\mathbf{X}}$ associated with a sequence of random variables $\mathbf{X}^T := (X_1, X_2, ..., X_T)$ is a random element of $\mathcal{M}_T$.

Let P_μ denote the probability law associated with an infinite sequence of i.i.d. random variables $\mathbf{X} := (X_1, X_2, \ldots)$ distributed following $\mu \in \mathcal{P}(\mathcal{A})$. Also, the relative entropy of probability vector ν with respect to another probability vector μ is

$$H(\nu|\mu) := \sum_{s=1}^{S} \nu_s \ln \frac{\nu_s}{\mu_s}.$$

Proposition 2 (Lemma 2.1.9; Dembo and Zeitouni [1998]): *For any* $\nu \in \mathcal{M}_T$,

$$(T+1)^{-S} e^{-TH(\nu|\mu)} \leq P_\mu(m_T^{\mathbf{X}} = \nu) \leq e^{-TH(\nu|\mu)}. \tag{1}$$

(Proof) See Dembo and Zeitouni (1998). ■

Proposition 2 states that the probability of observing type ν for a sequence of length T with respect to probability law μ has the lower and upper bounds as specified in (1).[7] Clearly, both the lower and upper bounds are decreasing in $H(\nu|\mu)$. Note that this result (and the results in the literature of large deviations) is very useful, since it may well be rather difficult to compute the precise probability $P_\mu(m_T^{\mathbf{X}} = \nu)$ in many cases. This difficulty arises from the fact that we need to consider all possible paths/sequences that belong to the specified type, which involves combinatorics. Moreover, from this result, we know that the relative entropy $H(\nu|\mu)$ characterises the probability $P_\mu(m_T^{\mathbf{X}} = \nu)$, although the bounds may not be very tight in some cases.[8]

Suppose μ is specified as $\mu_s = (1-\mu_1)f_s$ for all $s \neq 1$ with $\sum_{s=2}^{S} f_s = 1$, i.e. the same structure as π above. Also, suppose the empirical measure is specified as $\nu_1 = \mu_1 + \varepsilon$ and $\nu_s = (1-\nu_1)f_s$ for all $s \neq 1$, i.e. just like q above relative to π, although we allow for $\varepsilon < 0$ here. In this case, the relative entropy $H(\nu|\mu)$ is

$$\begin{aligned} H(\nu|\mu) &= (\mu_1+\varepsilon)\ln\frac{\mu_1+\varepsilon}{\mu_1} + \sum_{s=2}^{S}(1-\mu_1-\varepsilon)\,f_s \ln\frac{(1-\mu_1-\varepsilon)f_s}{(1-\mu_1)f_s} \\ &= (\mu_1+\varepsilon)\ln\frac{\mu_1+\varepsilon}{\mu_1} + (1-\mu_1-\varepsilon)\ln\frac{1-\mu_1-\varepsilon}{1-\mu_1}. \end{aligned} \tag{2}$$

The following proposition examines how the relative entropy (hence, the bounds on the large deviation probability) changes with respect to changes in μ_1 given $\varepsilon \neq 0$.

Proposition 3: *For every $\varepsilon \neq 0$ given, there exists a critical value $\hat{\mu}_1$ such that*

$$\frac{\partial H(\nu|\mu)}{\partial \mu_1} \begin{cases} > 0 & \text{if } \mu_1 > \hat{\mu}_1, \\ = 0 & \text{if } \mu_1 = \hat{\mu}_1, \\ < 0 & \text{otherwise.} \end{cases}$$

[7] $P_\mu(m_T^{\mathbf{X}} = \nu)$ is a likelihood function in the language of Bayesian statistics, in which case an explicit updating of beliefs is modelled. However, we do not assume such an explicit belief updating mechanism in the current paper.

[8] In particular, relative entropy characterises the speed of convergence in the context of the strong law of large numbers.

Note that ε can be either positive or negative, but it is restricted so that ν is a probability vector.

(Proof) See appendix. ■

Proposition 3 shows that the bounds in (1) become smaller when μ_1 is larger for a given level of ε. In other words, for a given level of discrepancy between the probability law in mind and the empirical measure in absolute terms, it is less likely to observe such a discrepancy when μ_1 is greater. Hence, if we regard μ as an agent's subjective probability belief, a small discrepancy in absolute terms of probability between the agent's subjective probability belief and the empirical measure may well result in the rejection of his previous belief when μ_1 is large. In contrast, an agent is less likely to reject his belief when μ_1 is small. Consequently, a larger class of probability laws would be regarded as compatible with an empirical distribution that exhibits a lower frequency of the no loss state a_1.

Example 2: We examine simple examples where $S = 2$ so as to illustrate proposition 3. Table 3 reports the precise probability of observing particular types of ν as well as the upper bound in (1) when $\mu_1 = 0.95$. On the other

Table 3: Probability of observing ν when $\mu_1 = 0.95$ and $T = 100$

ν_1	$P_\mu(m_T^{\mathbf{X}} = \nu)$	Upper bound in (1)
0.96	0.20086009	0.893444928
0.94	0.202529288	0.905551559

hand, table 4 reports the precise probability of observing particular types of ν as well as the upper bound in (1) when $\mu_1 = 0.9$. Namely, a case with a lower μ_1 compared to the one in table 3 is reported here.

Table 4: Probability of observing ν when $\mu_1 = 0.9$ and $T = 100$

ν_1	$P_\mu(m_T^{\mathbf{X}} = \nu)$	Upper bound in (1)
0.91	0.268163789	0.944319750
0.89	0.360862383	0.947443263

It is clear from tables 3 and 4 that for a given level of ε (0.01 or -0.01 in the tables), the probabilities (and the upper bounds) become smaller when

μ_1 is larger. ■

Notice that, for rare events, it is very likely that $m_T^{\mathbf{x}}(a_1) = \nu_1 = 1$ and $m_T^{\mathbf{x}}(a_s) = \nu_s = 0$ for all $s \neq 1$, i.e. no loss is ever realised. In this case, the following holds trivially:

$$P_\mu \left\{ m_T^{\mathbf{X}} = (1, 0, ..., 0) \right\} = \mu_1^T.$$

Note that this coincides with the upper bound in (1), since in this case,

$$H(\nu|\mu) = \nu_1 \ln \frac{\nu_1}{\mu_1} = -\ln \mu_1,$$

which implies that the upper bound is $e^{T \ln \mu_1} = \mu_1^T$. Table 5 reports the probability of observing no losses for different levels of μ_1, when we fix $T = 100$.[9]

Table 5: Probability of observing $\nu_1 = 1$ for $T = 100$

μ_1	$P_\mu(m_T^{\mathbf{X}} = \nu)$
0.999999999	0.999999900
0.999999	0.999900050
0.9999	0.990049339
0.999	0.904792147
0.99	0.366032341

Observe that when μ_1 is as low as 0.9999 (i.e. the probability of the loss state is 1/10000), the upper bound is 0.99. This means that for any $\mu_1 \geq 0.9999$, up to about 99% there will be no loss for a period of length $T = 100$. Hence, the scale of loss probabilities that are compatible with the empirical frequency may vary substantially for rare events, particularly unprecedented events (any probability less than 1/10000 is very plausible when $T = 100$). Note that the results of table 5 holds regardless of the number of states S as long as it is larger than 1.

Table 6 meanwhile reports the upper bound (1) for the probability of observing one loss event in $T = 100$ for different values of μ_1 (assuming that there are only two states). It is obvious that one realisation of state 2 out of 100 periods/samples is not compatible with any μ_1 greater than 0.999. This means that one occurrence of a rare event out of large enough samples for one's lifetime experience is typically incompatible with beliefs that assign

[9] In fact, the upper bounds are the exact probabilities here.

Table 6: Probability of observing $\nu_1 = 0.99$ for $T = 100$

μ_1	Upper bound
0.999999999	2.70468×10^{-7}
0.999999	0.000270441
0.9999	0.026780335
0.999	0.244962197

very low probability to such an event. As a result, one occurrence of a rare event may well result in a substantial revision of beliefs of the agents so that their willingness-to-pay for the insurance rises rapidly, which is consistent with Stylised Fact 2, and particularly, Stylised Fact 3.

However, the above result does not necessarily imply that agents can learn the true probability law quickly. One reason is that the number of empirical data (samples) is typically limited for natural disasters (e.g. 100 samples). This is because different properties in different locations cannot be understood as random samples (i.e. not i.i.d.) when it comes to natural disasters, since losses for different properties in the same region are correlated, and each property has its own characteristics, i.e. losses for different properties are not identically distributed. This is in sharp contrast with more conventional insurance products, such as travel insurance, where the number of empirical data (samples) that one can refer to is far larger (1 million samples, for instance).

3 Empirical Analyses

In this section, we shall test the theoretical predictions empirically. In so doing, we use the household data collected in Vietnam. In what follows, we first explain the data, and then report the estimation results.

3.1 Data

We utilise a unique survey data collected jointly by the Research Institute of Economy, Trade and Industry (RIETI) of Japan and the Center for Agricultural Policy in Vietnam (CAP), which we call the RIETI-CAP survey. The data set is a resurvey of subsamples of the Vietnam Household Living Standards Survey (VHLSS) 2006 households.[10]

[10]VHLSS is a biennial nationally representative rotating-panel household survey conducted by the General Statistics Office (GSO) with technical assistance from UNDP and

Table 7: Natural Disasters and Epidemics in the past five years

Province	Floods	Typhoons	Droughts	Natural Disasters	Epidemics
Ha Tay	0.042	0.042	0.000	0.083	*0.917*
Lao Cai	0.111	0.333	0.000	0.444	0.333
Nghe An	*0.533*	0.111	0.378	*1.022*	0.444
Quang Nam	*0.500*	0.143	0.393	*1.036*	*0.714*
Nationwide	0.375	0.292	0.235	0.902	0.656

Data: VHLSS 2004

Since the RIETI-CAP survey aims to collect data to facilitate the design of an insurance scheme against avian influenza (AI, hereafter) and flooding, sub-samples of VHLSS 2006 are chosen from four provinces: (1) Ha Tay (hit only by AI); (2) Nghe An (hit only by flooding); (3) Quang Nam (hit both by AI and flooding); and (4) Lao Cai (hit neither by AI nor by flooding). The selection of these four provinces was made using commune questionnaire data in VHLSS 2004. Table 7 reports the average numbers of natural disasters and animal epidemics per commune for the five years to 2004 in the above four provinces.

Table 8: RIETI-CAP Survey: Basic Information

Province	Training	End of Survey	Communes	Households
Ha Tay	20-21 Mar	3rd week April	22	508
Lao Cai	21-23 Feb	3rd week April	18	450
Nghe An	09-11 Mar	2nd week April	23	550
Qaung Nam	13-15 Mar	1st week April	19	510

The RIETI-CAP survey was conducted from late February 2008 until April 2008 (see Table 8 for basic information). The households covered in the REITI-CAP data include both those with and without the expenditure module in VHLSS 2006. The data covers approximately 500 households from each province, of which 100 households are with both income and expenditure data and 400 households have income data only.

The data represents extensive information, such as current and retro-

the World Bank, and is a multi-purpose household survey covering household characteristics, expenditures, income, health and education. Enumeration areas of VHLSS data are chosen randomly from the 1999 Population Census enumeration areas and households are selected randomly in each enumeration area. In VHLSS 2006, surveys with 30,000 households were conducted, providing representative statistics at the provincial level.

Table 9: Past Loss Experience Data

Causes of Losses	Number of Loss Events							Total	Mean	Std Dev
	0	1	2	3	4	5	6			
AI	1827	161	26	4	0	0	0	2018	0.1115	0.3699
Flood	1553	356	83	20	4	2	0	2018	0.3013	0.6293
Typhoon	1575	401	35	7	0	0	0	2018	0.2438	0.4899
Drought	1903	97	4	14	0	0	0	2018	0.0728	0.3364
Hail	1963	51	3	1	0	0	0	2018	0.0297	0.1866
Landslide	2001	14	3	0	0	0	0	2018	0.0099	0.1131
Other Epidemics	1557	306	83	20	17	34	1	2018	0.3845	0.9120
Other Disasters	1732	218	52	14	2	0	0	2018	0.1843	0.5055

spective income and expenditure information, asset information, subjective questions on insurance subscriptions, borrowing, past loss experiences of natural disasters, subjective probability assessments of AI and/or flooding, the maximum willingness-to-pay for various hypothetical insurance schemes, and time preference. In particular, table 9 reports the summary statistics and the distributions of past loss experience.

Since it is likely to be rather difficult for the respondents to provide a probability assessment of AI outbreak or flooding, the survey asked questions in a somewhat special way. More specifically, for household/personal AI risk:

Suppose you are observing fair coin flips. Please indicate the numbers of consecutive 'Heads' that are more likely to be observed than your household's experiencing losses from AI within one year from today. (Please circle all applicable entries and enter the highest value in to the box)

(a) 3 (b) 7 (c) 10 (d) 20 (e) 30 ■

For example, if none of the five entries is circled, then the respondent assesses that the respondent's household would suffer from AI with a higher probability than $2^{-3} = 0.125$. If only (a) is circled, then the probability assessment is strictly higher than $2^{-7} = 0.0078125$ and lower or equal to 0.125. If all five entries are circled, then the probability assessment is lower or equal to 2^{-30}, which is approximately one in one billion. We ask a similar question for losses from AI at the village level as well as those for losses from flooding at household and/or village levels. Unfortunately, the range of the scale of probability can be very large, in particular when all entries are circled.

3.2 Estimation Results

In what follows, we shall test the theoretical predictions of the subjective expected utility framework by using the RIETI-CAP survey data. The first prediction follows directly from proposition 3.

Prediction 1: *The subjective loss probability is increasing and concave in the frequency of past losses. This applies to flooding more than to AI.*

The following three predictions follow from the strict concavity of von Neumann-Morgenstern utility function.

Prediction 2: *The willingness-to-pay for insurance (both index and indemnity-based) is increasing and concave in subjective loss probability.*

Prediction 3: *The willingness-to-pay for index insurance is increasing and concave in the insurance payment.*

Prediction 4: *The willingness-to-pay for indemnity-based insurance is increasing and concave in the maximum indemnity of the insurance.*

All estimation results as well as the definitions of variables are reported in the appendix.[11] Since the index insurance's payment is contingent on the occurrence of AI outbreak or flooding at the village level, the appropriate subjective probability for index insurance is that for occurrence at the village level. On the other hand, the conventional indemnity-based insurance's payment is the indemnity for the household's losses, the subjective probability is that for the occurrence of losses at the household level.

For estimation results, we report results based on straight OLS estimations as well as instrumental variable (IV) estimations. We report IV estimation results that are more acceptable with respect to Hansen J test for over-identification, albeit with some exceptions. Also, we report F statistics for excluded instruments in the first stage estimation results for IV specifications as weak instrument test statistics.

[11] We also estimated the impact of subjective loss probability on willingness-to-pay semi-parametrically. We do not report the results here, since the qualitative results are similar to the ones reported in the paper. However, the results are available upon request.

3.2.1 Impacts of Past Loss Experience on Subjective Loss Probability

In what follows, we examine the impacts of past loss experience on subjective loss probability. First, table 10 reports the distribution of subjective loss probability for AI at the household level conditional on the past loss experience. Note that '-30' refers to subjective loss probability that is smaller or equal to 2^{-30}, i.e. households who circled all entries (a)–(e), while '-20' refers to subjective loss probability that is lower or equal to 2^{-20} and higher than 2^{-30}. It is clear that subjective loss probability tends to be higher when the number of past loss experience is larger. Also, subjective loss probability is very diverse amongst households with no past loss experience. These are consistent with the theoretical predictions above. However, subjective loss probability remains to vary widely even amongst households with some loss experience in the past.

Table 10: Distribution of Subjective Loss Probability (AI)

Loss Events	-30	-20	-10	-7	-3	Total
0	542	534	413	230	108	1827
1	29	39	42	37	14	161
2	4	0	8	9	5	26
3	0	0	2	2	0	4

Next, table 11 reports the distribution of subjective loss probability for flooding at the household level conditional on the past loss experience. It is clear that the theoretical predictions above are by and large supported here. Namely, (a) subjective loss probability tends to be higher when the number of past loss experience is larger, (b) subjective loss probability is very diverse amongst households with no past loss experience, and (c) subjective loss probability becomes less diverse more in agreement amongst households with some loss experience in the past.

Table 12 in the appendix reports the regression results of the subjective loss probability at the household level on the household's past loss experience in the last 5 years. It is clear that the past loss experience is statistically significant for both the subjective probability of AI outbreak and that of flooding at the village level. For the AI, the subjective probability of AI outbreak becomes $2^{3.160} \approx 8.94$ times higher with one additional AI loss experience on average, while the probability of flooding becomes $2^{6.121} \approx 69.60$ times higher with one additional flooding loss experience on average.

Table 11: Distribution of Subjective Loss Probability (Flooding)

Loss Events	-30	-20	-10	-7	-3	Total
0	582	413	285	202	67	1549
1	15	81	61	73	126	356
2	1	1	20	24	37	83
3	0	1	3	4	12	20
4	0	0	0	1	3	4
5	0	0	1	0	1	2

Hence, one additional loss experience has a very significant impact on the subjective probability for flooding losses, while the impact is more limited for AI losses. Moreover, the estimate for flooding has a much lower standard error (0.148) compared to that for AI (0.248).

Also, an additional loss experience has a diminishing effects on the subjective loss probability, especially for flooding: from no loss experience to one experience, the subjective loss probability becomes $2^{8.285} \approx 311.91$ times higher, from one loss experience to two experiences, it becomes $2^{4.072} \approx 16.82$ times higher, from two experiences to three experiences, it becomes $2^{1.952} \approx 3.87$ times higher, and so on. This is clearly consistent with Prediction 1 above. However, it is less obvious if this is the case for AI.

A caveat to this interpretation is that the impacts of loss experience are measured by comparing different agents (i.e. cross section), but are not based on the changes of the subjective probabilities of the same household over time. Still, the estimation results by and large support Prediction 1. Moreover, the weaker results for AI may be reflecting the aspect that the AI involves some unforeseen contingencies, since the mutation of viruses is highly unpredictable. In other words, mutations cannot be captured as a 'small world' event, and so are not compatible with the subjective expected utility framework of Savage (1954).

3.2.2 Regressions of Willingness-To-Pay

The first two columns of table 13 report the OLS regression results of willingness-to-pay for AI index insurance. Both the subjective probability (*lprob*) and the insurance payment (*lpay*) are statistically significant in both regressions. Note that these are consistent with Predictions 2 and 3 above. In particular, the elasticity of the willingness-to-pay with respect to insurance payments is relatively high, if not unity.[12] Moreover, in the second column, the maximum AI loss the household could incur is significant at the 1% level (and is positive). However, income (*lincome*) is statistically not significant.

The third to fifth columns of table 13 report the instrumental variable (IV) estimation results for AI index insurance, while table 14 report the first stage estimation results for subjective loss probability *lprob* (and in addition, the maximum possible AI loss, *lmaxloss*, for the IV3 specification). It is clear that insurance payment (*lpay*) is always statistically significant with a point estimate of around 0.52 and a small standard error (0.013-0.014) in all specifications. Thus, we can conclude that Prediction 3 is comfortably supported here. Moreover, *lprob* is always statistically significant and is positive, and so Prediction 2 is supported here, too.

Tables 15 and 16 report the estimation results for flooding index insurance. Again, the insurance payment (*lpay*) is always significant, and the point estimate is relatively high (0.58-0.59). Thus, Prediction 3 is once again supported. Also, the maximum loss (*lmaxloss*) is always significant. However, income (*lincome*) is not always statistically significant. Moreover, *lprob* is almost always significant for flooding index insurance. Hence, Prediction 2 is supported here. Note that the IV point estimates are far larger than the OLS estimators, although the point estimates indicate that an increase in the willingness-to-pay for the flooding index insurance is not really proportionate with an increase in subjective loss probability: on average, it is one digit less than an increase in subjective loss probability. For instance, in IV 4, the estimated elasticity of willingness-to-pay with respect to the subjective loss probability is 0.163, while the point estimate of $exd1$ in the first stage is 2.928. Hence, willingness-to-pay becomes $(2^{2.928})^{0.163} \approx 1.39$ times higher with an additional flooding loss experience (compared to no loss).[13] This may not provide very strong support for Prediction 2, especially in light of Stylised Fact 1. Nevertheless, we cannot claim that this contradicts with the prediction. Recall that the subjective loss probability data does not necessarily reflect the scales of the actual subjective probabilities very accurately, e.g. respondents that circled all entries up to (d) range from 2^{-20} (approximately 1 in 1 million) to 2^{-30} (approximately 1 in 1 billion). This may well have impacts on the estimation results, and so the estimation results are not necessarily too out of line of Prediction 2 even in light of Stylised Fact 1.

[12] The elasticity is unity if the willingness-to-pay is linear in the insurance payment.

[13] $wtp = a \cdot prob^c$ holds under the specification of the econometric model, where wtp is willingness-to-pay, $prob$ the subjective probability, while a and c are some constants.

We now turn our attention to regression results for indemnity-based insurance. Tables 17 and 18 report the estimation results for AI indemnity-based insurance, while tables 19 and 20 report those for flooding indemnity-based insurance. The payment of the indemnity-based insurance is an indemnity, i.e. $\min\{\text{maximum loss}, \text{maximum insurance payment}\}$. Hence, *lindem* is included as a regressor instead of *lpay*. The estimation results by and large resemble those for index insurance. However, the point estimates of the coefficients of subjective probability (*lprob*) for AI losses are rather small, although they are still statistically significant. Hence, Prediction 2 is supported here, albeit not very strongly. A similar claim may hold for the the point estimates of the coefficients of subjective probability (*lprob*) for flooding losses. Still, the IV estimators reported in the fourth column of table 19 is similar to the ones found in table 15. Hence, Prediction 2 is once again supported. Moreover, Prediction 4 is supported for both AI and flooding indemnity-based insurance, since indemnity (*lindem*) is always significant with a point estimate of about 0.2 to 0.3 with a small standard error (approximately 0.01).

Apart from results concerning the four predictions, there are some interesting results. First, it is clear that bad past experience concerning insurance has a negative effect, since households who couldn't receive the insurance payment in full (*partialpay* dummy) have a lower willingness-to-pay. Moreover, the results here indicate that insurance is not a luxury good. This is in a sharp contrast to some existing studies using macro data (e.g. Beenstock et al., 1988; Outreville, 1990; and Enz, 2000). Our results are more in line with the predictions of the expected utility framework, and thus, our results suggest that the previous findings may well be a fallacy.

3.2.3 Policy Implications

We have found that a first loss experience tends to have a large impact on the subjective loss probability, and consequently on the willingness-to-pay for insurance, especially for flooding insurance (both index and indemnity-based insurance). This indicates that it would be less likely for a household with no past loss experience to purchase flooding insurance even if the insurance premium is actuarially fair in accord with the loss probability model of the insurance supplier. Hence, a paternalistic policy that makes subscription of insurance mandatory may well be more effective in ensuring that the flooding insurance mechanism functions properly, even if the agents behave in accord with the subjective expected utility framework. Note that a mandatory subscription may be fully or partially subsidised by the government to minimise the burden on households.

Meanwhile, we have found that agents may not behave in accord with the subjective expected utility framework as far as AI insurance is concerned. In other words, it is less clear if it is the subjective loss probability that drives the behaviour of the agents concerning AI insurance. This is not very surprising, since AI involves mutations of viruses, and so, there are possible unforeseen contingencies. This makes it harder to have some agreement on the terms and conditions of insurance.

In recent years, the World Bank has implemented a pilot project of rainfall index insurance in India. Giné, Townsend and Vickery (2008) studies the structure of demand for the rain fall index insurance, but it is not very clear if introduction of formal insurance would indeed be successful since the structure of (potential) insurance demand in developing countries is essentially unknown. One advantage of our data based on hypothetical questions is that it includes household that may not purchase insurance, while the actual insurance subscription data does not include them systematically. However, as far as households that purchase insurance are concerned, clearly the actual data is superior to data based on hypothetical questions, since the quality of the response to the hypothetical questions is not as good as the actual subscription data. Since there is unfortunately no immediate way to take advantage of the two methodologies, it is important to gain as much insight as possible from both.

4 Conclusion

We have shown theoretically that under the subjective expected utility framework, more agents would not opt to purchase actuarially fair insurance with respect to the insurer's probability estimate π when the entropy of π is smaller. This is because even a small discrepancy in the probability estimate would have a major impact on the willingness-to-pay for the insurance when the entropy of the loss probability distribution is small. This prediction is compatible with the casual observation that insurance for large losses (catastrophes) with small probabilities is purchased with less frequency than is more conventional insurance products that cover more frequent but moderate losses, even though the premium for catastrophe insurance is more favourable for the potential buyers.

Moreover, we have seen that the class of probability beliefs that is compatible with the empirical data is smaller when the entropy of π is smaller. This appears to suggest that agents can learn and adjust their probability estimates for catastrophes or losses caused by natural disasters more quickly. However, this may well not be the case, since losses for different properties (in the same region) are highly correlated and are not identically distributed, and so we cannot regard them as random samples. Moreover, the scale of loss probabilities may vary substantially for rare events—in particular, unprecedented events, and so heterogeneity in the willingness-to-pay for the insurance would have a major impact in such cases.

We tested these theoretical predictions empirically, using the RIETI-CAP data collected in Vietnam. The tests show that flooding insurance, both index insurance and indemnity-based insurance, is more in line with the theoretical predictions above based on the subjective expected utility framework than AI insurance is. This may be a reflection of the fact that AI involves unforeseen contingencies that are incompatible with the subjective expected utility framework.

References

[1] Arrow, K. J.: 'Uncertainty and the Welfare Economics of Medical Care', *American Economic Review*, **53**, 941-973 (1963).

[2] Beenstock, M., G. Dickinson and S. Khajuria: 'The Relationship between Property-Liability Insurance Premiums and Income: An International Analysis', *Journal of Risk and Insurance*, **55**, 259-272 (1988)

[3] Borch, K.: 'Equilibrium in a Reinsurance Market', *Econometrica*, **30**, 424-444 (1962).

[4] Brookshire, D. S., M. A. Thayer, J. Tschirhart and W. D. Schulze: 'A Test of the Expected Utility Model: Evidence from Earthquake Risks', *Journal of Political Economy*, **93**, 369-389 (1985).

[5] Browne, M. J. and R. E. Hoyt: 'The Demand for Flood Insurance: Empirical Evidence', *Journal of Risk and Uncertainty*, **20**, 291-306 (2000).

[6] Cover, T. M. and J. A. Thomas: *Elements of Information Theory*, Wiley, New York, 1991.

[7] Dembo, A. and O. Zeitouni: *Large Deviations Techiniques and Applications*, 2nd ed., Springer-Verlag, New York, 1998.

[8] Dixon, L., N. Clancy, S. A. Seabury and A. Overton: 'The National Flood Insurance Program's Market Penetration Rate–Estimates and Policy Implications', Technical Report, Rand Corporation (2006).

[9] Ehrlich, I. and G. S. Becker: 'Market Insurance, Self-Insurance, and Self-Protection', *Journal of Political Economy*, **80**, 623-648 (1972).

[10] Enz, R.: 'The S-Curve Relation between Per-Capita Income and Insurance Penetration', *Geneva Papers on Risk and Insurance*, **25**, 396-406 (2000).

[11] Giné, X., R. Townsend, and J. Vickery: 'Patterns of Rainfall Insurance Participation in India', *The World Bank Economic Review*, **22**, 539-566 (2008).

[12] Kahneman, D., and A. Tversky: 'Prospect theory: An analysis of decision under risk', *Econometrica*, **47**, 263-293 (1979).

[13] Kahneman, D., and A. Tversky: 'Advances in prospect theory: Cumulative representation of uncertainty', *Journal of Risk and Uncertainty*, **5**, 297-323 (1992).

[14] Kunreuther, H., R. Ginsberg, L. Miller, P. Sagi, P. Slovic, B. Borkan, and N. Katz:*Disaster Insurance Protection: Public Policy Lessons*, Wiley, New York, 1978.

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

[16] Mossin, J.: 'Aspects of Rational Insurance Purchasing', *Journal of Political Economy*, **76**, 553-568 (1968).

[17] Outreville, J. F.: 'The Economic Significance of Insurance Markets in Developing Countries', *Journal of Risk and Insurance*, **57**, 487-498 (1990).

[18] Savage, R: *The Foundations of Statistics*, New York, NY: Wiley, 1954.

[19] Smith, V. L.: 'Optimal Insurance Coverage', *Journal of Political Economy*, **76**, 68-77 (1968).

A Proof of Proposition 3

From equation (2), i.e.

$$H(\nu|\mu) = (\mu_1 + \varepsilon) \ln \frac{\mu_1 + \varepsilon}{\mu_1} + (1 - \mu_1 - \varepsilon) \ln \frac{1 - \mu_1 - \varepsilon}{1 - \mu_1},$$

we can derive first derivative of $H(\nu|\mu)$ with respect to μ_1 as follows.

$$\frac{\partial H(\nu|\mu)}{\partial \mu_1} = \ln \left(\frac{1 - \mu_1}{\mu_1} \cdot \frac{\mu_1 + \varepsilon}{1 - \mu_1 - \varepsilon} \right) - \frac{\varepsilon}{\mu_1(1 - \mu_1)}.$$

Meanwhile, the second derivative is

$$\begin{aligned} \frac{\partial^2 H(\nu|\mu)}{\partial \mu_1^2} &= \frac{-1}{\mu_1(1 - \mu_1)} + \frac{1}{(\mu_1 + \varepsilon)(1 - \mu_1 - \varepsilon)} + \frac{1 - 2\mu_1}{\mu_1^2(1 - \mu_1)^2}\varepsilon \\ &= \frac{\varepsilon + 2\mu_1 - 1}{\mu_1(\mu_1 + \varepsilon)(1 - \mu_1 - \varepsilon)(1 - \mu_1)}\varepsilon + \frac{1 - 2\mu_1}{\mu_1^2(1 - \mu_1)^2}\varepsilon. \end{aligned}$$

In what follows, we show that $\partial^2 H(\nu|\mu)/\partial \mu_1^2 > 0$ for all ε and μ_1.
(i) For $\varepsilon > 0$ and $\mu_1 \geq (1 - \varepsilon)/2$,

$$\begin{aligned} \frac{\partial^2 H(\nu|\mu)}{\partial \mu_1^2} &= \frac{\varepsilon + 2\mu_1 - 1}{\mu_1(\mu_1 + \varepsilon)(1 - \mu_1 - \varepsilon)(1 - \mu_1)}\varepsilon + \frac{1 - 2\mu_1}{\mu_1^2(1 - \mu_1)^2}\varepsilon \\ &> \frac{\varepsilon + 2\mu_1 - 1}{\mu_1^2(1 - \mu_1)^2}\varepsilon + \frac{1 - 2\mu_1}{\mu_1^2(1 - \mu_1)^2}\varepsilon \\ &= \frac{\varepsilon^2}{\mu_1^2(1 - \mu_1)^2} \\ &> 0. \end{aligned}$$

(ii) For $\varepsilon > 0$ and $\mu_1 < (1 - \varepsilon)/2$,

$$\begin{aligned} \frac{\partial^2 H(\nu|\mu)}{\partial \mu_1^2} &= \frac{\varepsilon + 2\mu_1 - 1}{\mu_1(\mu_1 + \varepsilon)(1 - \mu_1 - \varepsilon)(1 - \mu_1)}\varepsilon + \frac{1 - 2\mu_1}{\mu_1^2(1 - \mu_1)^2}\varepsilon \\ &\geq \frac{\varepsilon + 2\mu_1 - 1}{\mu_1(\mu_1 + \varepsilon)(1 - \mu_1 - \varepsilon)(1 - \mu_1)}\varepsilon + \frac{1 - 2\mu_1}{\mu_1(\mu_1 + \varepsilon)(1 - \mu_1 - \varepsilon)(1 - \mu_1)}\varepsilon \\ &= \frac{\varepsilon^2}{\mu_1(\mu_1 + \varepsilon)(1 - \mu_1 - \varepsilon)(1 - \mu_1)} \\ &> 0. \end{aligned}$$

(iii) For $\varepsilon < 0$ and $\mu_1 \geq 0.5$,

$$\begin{aligned}\frac{\partial^2 H(\nu|\mu)}{\partial \mu_1^2} &= \frac{\varepsilon + 2\mu_1 - 1}{\mu_1(\mu_1+\varepsilon)(1-\mu_1-\varepsilon)(1-\mu_1)}\varepsilon + \frac{1-2\mu_1}{\mu_1^2(1-\mu_1)^2}\varepsilon \\ &\geq \frac{\varepsilon + 2\mu_1 - 1}{\mu_1(\mu_1+\varepsilon)(1-\mu_1-\varepsilon)(1-\mu_1)}\varepsilon + \frac{1-2\mu_1}{\mu_1(\mu_1+\varepsilon)(1-\mu_1-\varepsilon)(1-\mu_1)}\varepsilon \\ &= \frac{\varepsilon^2}{\mu_1(\mu_1+\varepsilon)(1-\mu_1-\varepsilon)(1-\mu_1)} \\ &> 0.\end{aligned}$$

(iv) For $\varepsilon < 0$ and $\mu_1 < 0.5$,

$$\begin{aligned}\frac{\partial^2 H(\nu|\mu)}{\partial \mu_1^2} &= \frac{\varepsilon + 2\mu_1 - 1}{\mu_1(\mu_1+\varepsilon)(1-\mu_1-\varepsilon)(1-\mu_1)}\varepsilon + \frac{1-2\mu_1}{\mu_1^2(1-\mu_1)^2}\varepsilon \\ &\geq \frac{\varepsilon + 2\mu_1 - 1}{\mu_1^2(1-\mu_1)^2}\varepsilon + \frac{1-2\mu_1}{\mu_1^2(1-\mu_1)^2}\varepsilon \\ &= \frac{\varepsilon^2}{\mu_1^2(1-\mu_1)^2} \\ &> 0.\end{aligned}$$

Hence, $\partial^2 H(\nu|\mu)/\partial \mu_1^2 > 0$ for all ε and μ_1.

Next, we show the existence of a critical value $\hat{\mu}_1 \in (0.5-\varepsilon, 0.5)$ for $\varepsilon > 0$ and $\hat{\mu}_1 \in (0.5, 0.5-\varepsilon)$ for $\varepsilon < 0$. To do so, observe that when $\mu_1 = 0.5$,

$$\frac{\partial H(\nu|\mu)}{\partial \mu_1} = \ln\left(\frac{0.5+\varepsilon}{0.5-\varepsilon}\right) - 4\varepsilon \begin{cases} > 0, & \forall \varepsilon > 0, \\ < 0, & \forall \varepsilon < 0. \end{cases}$$

Also, when $\mu_1 = 0.5 - \varepsilon$,

$$\frac{\partial H(\nu|\mu)}{\partial \mu_1} = \ln\left(\frac{0.5+\varepsilon}{0.5-\varepsilon}\right) - \frac{\varepsilon}{0.25-\varepsilon^2} \begin{cases} < 0, & \forall \varepsilon > 0, \\ > 0, & \forall \varepsilon < 0. \end{cases}$$

It follows that $\partial H(\nu|\mu)/\partial \mu_1 = 0$ at $\hat{\mu}_1 \in (0.5-\varepsilon, 0.5)$ for $\varepsilon > 0$, and at $\hat{\mu}_1 \in (0.5, 0.5-\varepsilon)$ for $\varepsilon < 0$, since $\partial^2 H(\nu|\mu)/\partial \mu_1^2 > 0$ for all ε and μ_1. The desired result follows immediately. ■

B RIETI-CAP: Willingness-to-pay questions

AI insurance (indemnity-based)

Consider insurance that pays you for losses on poultry from AI (officially verified) for one year from today. You have to pay the full amount in cash today to purchase the insurance.

The amount of insurance payment	WTP (Thousand VND)
The full amount (i.e. no limit)	
Up to 50 million VND	
Up to 20 million VND	
Up to 5 million VND	
Up to 1 million VND	

AI index insurance (village level)

Consider insurance that pays you a certain amount of money whenever AI hits your village (regardless of your actual loss) for one year from today. On the other hand, the insurance pays you nothing, when there is no AI outbreak in your village.

The amount of insurance payment	WTP (Thousand VND)
50 million VND	
20 million VND	
5 million VND	
1 million VND	

C Estimation Results

List of Variables (Note: The base of log is 2.)

- *wtp*: willingness-to-pay (thousand VND);
- *lwtp*: $\log(1 + wtp)$;
- *lprob*: log(subjective probability of loss event);
- *lmaxloss*: $\log(1 + \text{maximum loss})$;
- *ex*: number of times of losses (from AI for AI insurance, from flooding for flooding insurance) the household experienced in 5 years;
- *exwater*: number of times of losses (from flooding, typhoons and drought for AI insurance, from typhoons and drought for flooding insurance) the household experienced in 5 years;
- *exepi*: number of times of losses from epidemics (other than AI for AI insurance, including AI for flooding insurance) the household experienced in 5 years;
- *exother*: number of times of other losses the household experienced in 5 years;
- *exd*n: dummy with 1 if the household experienced losses (from AI for AI insurance, from flooding for flooding insurance) at least n times in 5 years.
- *lpay*: $\log(1 + \text{insurance payment})$;
- *lindem*: log of $\min\{\text{maximum loss}, \text{maximum insurance payment}\}$;
- *lasset*: log of total value of assets;
- *lincome*: log of annual income;
- *rural*: rural dummy (1 if the household is living in a rural area);
- *edusec*: dummy with 1 if a household member has finished the secondary school;
- *province*1: dummy with 1 if Ha Tay;
- *province*2: dummy with 1 if Lao Cai;
- *province*3: dummy with 1 if Nghe An;
- *partialpay*: dummy with 1 if the household did not receive insurance payment in full amount in the past;
- *borrowing*: dummy with 1 if the household has borrowing;
- *hhsize*: the number of household members;
- *agehead*: the age of the household head;
- *ageheadsq*: $agehead^2$;
- *wife*: dummy with 1 if the respondent is the household head's wife;
- *husband*: dummy with 1 if the respondent is the household head's husband;
- *son*: dummy with 1 if the respondent is the household head's son;
- *daughter*: dummy with 1 if the respondent is the household head's daughter;

- $floodloss1$: dummy with 1 if the most significant loss from flooding was physical assets;
- $floodloss2$: dummy with 1 if the most significant loss from flooding was physical loss of livestock;
- $floodloss3$: dummy with 1 if the most significant loss from flooding was economic loss of livestock;
- $floodloss4$: dummy with 1 if the most significant loss from flooding was loss of crops;

Table 12: Regressions of Loss Probabilities at Household Level (*lprob*)

Regressors	AI		Flooding	
ex	3.160***		6.121***	
	[0.248]		[0.148]	
*exd*1		3.745***		8.285***
		[0.361]		[0.234]
*exd*2		1.22		4.072***
		[0.885]		[0.358]
*exd*3		3.024**		1.952***
		[1.388]		[0.609]
*exd*4				0.556***
				[0.169]
*exd*5				-0.864***
				[0.310]
exwater	4.012***	4.032***	4.792***	4.454***
	[0.097]	[0.098]	[0.153]	[0.155]
exepi	0.410***	0.418***	0.478***	0.565***
	[0.128]	[0.128]	[0.114]	[0.115]
exother	0.105	0.115	1.148***	1.306***
	[0.194]	[0.194]	[0.185]	[0.183]
constant	-20.697***	-20.741***	-20.956***	-21.240***
	[0.136]	[0.138]	[0.141]	[0.143]
Observations	2018	2018	2014	2014
R-squared	0.17	0.17	0.25	0.26

Note 1: Standard errors in brackets (the same applies to all tables hereafter).

Note 2: * significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level (the same applies to all tables hereafter).

Note 3: *lprob* here is the subjective probability of AI outbreak/flooding loss event at the household level.

Table 13: AI index insurance (WTP) regressions (*lwtp*)

Regressors	OLS	OLS	IV 1	IV 2	IV 3
lprob	0.041***	0.018***	0.088***	0.073***	0.158***
	[0.003]	[0.004]	[0.017]	[0.006]	[0.059]
lpay	0.522***	0.520***	0.522***	0.523***	0.521***
	[0.013]	[0.012]	[0.013]	[0.013]	[0.014]
lmaxloss		0.162***			-0.11
		[0.006]			[0.110]
lasset		0.023			0.109***
		[0.018]			[0.041]
lincome		0.016			-0.152**
		[0.029]			[0.076]
rural		-0.220***			0.281
		[0.077]			[0.221]
edusec		0.120*			0.045
		[0.065]			[0.081]
*province*1		-1.002***			1.055
		[0.089]			[0.867]
*province*2		-0.808***			1.129
		[0.087]			[0.827]
*province*3		-1.972***			-0.900*
		[0.081]			[0.465]
partialpay		-0.801***			-1.266***
		[0.133]			[0.232]
borrowing		0.032			0.063
		[0.063]			[0.072]
hhsize		-0.082***			0.022
		[0.018]			[0.046]
agehead		0.060***			0.068***
		[0.013]			[0.016]
ageheadsq		-0.00067***			-0.001***
		[0.00012]			[0.000]
wife		0.356***			0.466***
		[0.072]			[0.096]
husband		0.388*			0.398
		[0.224]			[0.249]
son		0.843***			0.765***
		[0.149]			[0.169]
daughter		0.624***			0.224
		[0.228]			[0.282]
constant	-1.123***	-2.864***	-0.286	-0.551**	0.091
	[0.173]	[0.403]	[0.348]	[0.199]	[1.251]
Observations	7979	7855	7979	7979	7855
R-squared	0.18	0.32			
Hansen J test			0.6277	0.5628	0.2126

Note 1: *lprob* here is the subjective probability of AI outbreak at the village level.
Note 2: *p*-values for Hansen J test are reported.

Table 14: AI index insurance first stage estimation results

Regressors	IV 1	IV 2	IV 3	
	lprob	*lprob*	*lprob*	*lmaxloss*
*exd*1	3.319***	3.649***	3.151***	1.821***
	[0.370]	[0.362]	[0.341]	[0.128]
*exd*2	4.327***	1.657**	-0.089	0.219
	[0.815]	[0.806]	[0.648]	[0.201]
exwater		4.009***	0.733***	0.043
		[0.097]	[0.103]	[0.063]
lpay	-0.004	-0.004	-0.002	0.003
	[0.049]	[0.045]	[0.039]	[0.023]
lasset			-0.231***	0.195***
			[0.060]	[0.036]
lincome			0.453***	-0.388***
			[0.095]	[0.057]
rural			-0.566*	1.587***
			[0.302]	[0.208]
edusec			0.555**	-0.001
			[0.230]	[0.128]
*province*1			-12.819***	0.527***
			[0.256]	[0.163]
*province*2			-12.440***	0.301*
			[0.302]	[0.168]
*province*3			-6.843***	0.202
			[0.241]	[0.143]
partialpay			1.387***	-0.892***
			[0.349]	[0.212]
borrowing			-0.256	0.009
			[0.205]	[0.112]
hhsize			-0.160***	0.302***
			[0.059]	[0.033]
agehead			0.017	0.044*
			[0.046]	[0.025]
ageheadsq			0.000	-0.001**
			[0.000]	[0.000]
wife			-0.025	0.420***
			[0.223]	[0.119]
husband			0.79	0.438
			[0.495]	[0.407]
son			1.649***	0.565**
			[0.421]	[0.241]
daughter			0.474	-1.211**
			[0.753]	[0.493]
constant	-17.998***	-20.460***	-11.006***	5.440***
	[0.644]	[0.597]	[1.362]	[0.755]
Obs	7979	7979	7871	7855
R-squared	0.02	0.17	0.38	0.07
F stats	89.43	659.18	51.26	103.38

Note 1: *lprob* here is the subjective probability of AI outbreak at the village level.

Note 2: The F statistics are for the excluded instruments.

Table 15: Flooding index insurance (WTP) regressions (*lwtp*)

Regressors	OLS	OLS	IV1	IV2	IV3	IV4
lprob	0.027***	0.005	0.089***	0.089***	0.156***	0.163***
	[0.002]	[0.004]	[0.005]	[0.005]	[0.020]	[0.020]
lpay	0.584***	0.583***	0.584***	0.591***	0.580***	0.580***
	[0.012]	[0.012]	[0.013]	[0.013]	[0.014]	[0.014]
lmaxloss		0.090***			0.324***	0.243***
		[0.007]			[0.064]	[0.043]
lasset		0.097***			0.069***	0.076***
		[0.017]			[0.021]	[0.020]
lincome		-0.083***			-0.077**	-0.067*
		[0.028]			[0.036]	[0.034]
rural		-0.309***			-0.775***	-0.618***
		[0.071]			[0.159]	[0.128]
edusec		0.256***			0.347***	0.296***
		[0.063]			[0.089]	[0.083]
*province*1		-1.001***			0.845*	1.195***
		[0.094]			[0.432]	[0.380]
*province*2		-0.849***			2.099***	2.027***
		[0.089]			[0.354]	[0.341]
*province*3		-1.704***			-0.432**	-0.382**
		[0.080]			[0.191]	[0.184]
partialpay		-0.721***			-0.851***	-0.803***
		[0.133]			[0.140]	[0.137]
borrowing		0.218***			0.012	0.053
		[0.060]			[0.078]	[0.073]
hhsize		-0.055***			-0.072***	-0.061***
		[0.016]			[0.021]	[0.020]
agehead		0.024*			-0.043**	-0.038**
		[0.013]			[0.017]	[0.016]
ageheadsq		-0.00032***			0.000**	0.000*
		[0.00012]			[0.000]	[0.000]
wife		0.289***			0.144*	0.178**
		[0.070]			[0.087]	[0.083]
husband		0.102			0.185	0.149
		[0.219]			[0.269]	[0.259]
son		0.764***			0.752***	0.815***
		[0.141]			[0.174]	[0.169]
daughter		0.308			0.027	0.041
		[0.226]			[0.299]	[0.285]
*floodloss*1		-0.032			0.106	0.202
		[0.239]			[0.285]	[0.294]
*floodloss*2		0.482***			0.336***	0.325***
		[0.090]			[0.112]	[0.105]
*floodloss*3		-0.173			-0.555**	-0.533*
		[0.236]			[0.281]	[0.293]
*floodloss*4		-5.677***			-5.283***	-5.045***
		[0.629]			[0.665]	[0.651]
constant	-1.828***	-2.612***	-0.784***	0.839***	-2.331***	-1.580**
	[0.160]	[0.392]	[0.183]	[0.184]	[0.766]	[0.637]
Observations	7940	7828	7940	7940	7828	7828
R-squared	0.23	0.31				
Hansen J test			0.0057	0	0.6419	0.1827

Note: *lprob* here is the subjective probability of flooding at the village level.

Table 16: Flooding index insurance first stage estimation results

Regressors	IV1	IV2	IV3		IV4	
	lprob	*lprob*	*lprob*	*lmaxloss*	*lprob*	*lmaxloss*
*exd*1	9.555***	9.617***	2.914***	1.051***	2.928***	1.190***
	[0.237]	[0.239]	[0.206]	[0.110]	[0.210]	[0.112]
*exd*2	3.763***	3.708***	2.057***	-1.410***	2.075***	-1.366***
	[0.284]	[0.286]	[0.273]	[0.157]	[0.273]	[0.157]
exlandslide	0.331	-0.15	1.964**	1.169***	1.975**	1.232***
	[0.943]	[0.980]	[0.924]	[0.349]	[0.924]	[0.353]
exAI		1.056***			-0.452	0.560***
		[0.248]			[0.324]	[0.242]
exepi		0.416***			0.064	0.519***
		[0.133]			[0.096]	[0.045]
lpay	-0.002	-0.002	0.003	0.008	0.003	0.008
	[0.047]	[0.047]	[0.037]	[0.019]	[0.037]	[0.019]
lasset			0.015	0.093***	0.013	0.075**
			[0.051]	[0.031]	[0.052]	[0.031]
lincome			-0.155*	0.074	-0.149*	0.109**
			[0.087]	[0.050]	[0.087]	[0.049]
rural			0.034	1.894***	0.001	1.698***
			[0.287]	[0.202]	[0.290]	[0.201]
edusec			0.146	-0.643***	0.15	-0.532***
			[0.225]	[0.109]	[0.227]	[0.108]
*province*1			-15.544***	3.106***	-15.537***	3.322***
			[0.219]	[0.121]	[0.223]	[0.125]
*province*2			14.704***	-1.940***	-14.729***	-2.045***
			[0.244]	[0.164]	[0.243]	[0.164]
*province*3			-7.454***	-0.041	-7.439***	0.129
			[0.215]	[0.127]	[0.217]	[0.128]

Note: *lprob* here is the subjective probability of flooding at the village level.

(Table 16 continued)

Regressors	IV1	IV2	IV3		IV4	
	lprob	*lprob*	*lprob*	*lmaxloss*	*lprob*	*lmaxloss*
hhsize			-0.079	0.129***	-0.083	0.098***
			[0.056]	[0.026]	[0.057]	[0.026]
agehead			0.261***	0.109***	0.261***	0.103***
			[0.042]	[0.021]	[0.042]	[0.021]
ageheadsq			-0.002***	-0.001***	-0.002***	-0.001***
			[0.000]	[0.000]	[0.000]	[0.000]
wife			0.152	0.492***	0.133	0.397***
			[0.221]	[0.100]	[0.222]	[0.100]
husband			-0.209	-0.238	-0.215	-0.17
			[0.550]	[0.381]	[0.550]	[0.382]
son			-1.005***	0.501**	-1.001**	0.508**
			[0.402]	[0.200]	[0.402]	[0.200]
daughter			1.550**	0.613*	1.521**	0.434
			[0.697]	[0.361]	[0.699]	[0.361]
*floodloss*1			-8.876***	-1.038	-8.128***	-2.621***
			[2.088]	[0.762]	[2.157]	[0.862]
*floodloss*2			0.279	0.351**	0.812	-0.21
			[0.330]	[0.137]	[0.576]	[0.327]
*floodloss*3			1.247	0.789***	1.758*	0.124
			[0.910]	[0.258]	[1.030]	[0.367]
*floodloss*4			-6.247***	2.574***	-5.782***	2.058***
			[0.267]	[0.108]	[0.418]	[0.278]
constant	-19.459***	-19.745***	-14.746***	7.420***	-14.743***	7.380***
	[0.629]	[0.630]	[1.241]	[0.652]	[1.241]	[0.642]
Observations	7940	7940	7832	7840	7832	7840
R-squared	0.2	0.2	0.51	0.23	0.51	0.24
F stats	1318.18	805.25	133.53	46.82	80.71	52.55

Note 1: *lprob* here is the subjective probability of flooding at the village level.

Note 2: The F statistics are for the excluded instruments.

Table 17: AI indemnity-based insurance (WTP) regressions (*lwtp*)

Regressors	OLS *lwtp*	OLS *lwtp*	IV1 *lwtp*	IV2 *lwtp*	IV3 *lwtp*
lprob	0.015***	0.007**	0.055***	0.058***	0.058***
	[0.003]	[0.003]	[0.007]	[0.020]	[0.020]
lindem	0.220***	0.228***	0.212***	0.215***	0.214***
	[0.008]	[0.008]	[0.008]	[0.009]	[0.009]
lasset		0.027		0.025	0.026
		[0.019]		[0.019]	[0.019]
lincome		0.016		0.000	-0.001
		[0.033]		[0.034]	[0.034]
rural		-0.351***		-0.382***	-0.368***
		[0.091]		[0.094]	[0.094]
edusec		0.136*		0.152**	0.167**
		[0.070]		[0.071]	[0.071]
*province*1		-0.831***		-0.266	-0.256
		[0.083]		[0.229]	[0.231]
*province*2		-0.476***		0.084	0.096
		[0.082]		[0.230]	[0.231]
*province*3		-1.483***		-1.218***	-1.223***
		[0.080]		[0.132]	[0.134]
partialpay		-0.366***		-0.431***	-0.441***
		[0.133]		[0.134]	[0.135]
borrowing		0.099		0.104	0.119*
		[0.065]		[0.066]	[0.066]
hhsize		-0.083***		-0.075***	-0.075***
		[0.019]		[0.019]	[0.019]
agehead		0.060***		0.056***	0.055***
		[0.014]		[0.014]	[0.014]
ageheadsq		-0.00066***		-0.00062***	-0.00057***
		[0.00012]		[0.00013]	[0.0004]
wife		0.233***		0.242***	0.237***
		[0.076]		[0.078]	[0.078]
husband		0.473**		0.581**	0.562**
		[0.237]		[0.246]	[0.247]
son		0.903***		0.937***	0.946***
		[0.155]		[0.154]	[0.154]
daughter		0.522**		0.509**	0.518**
		[0.241]		[0.244]	[0.246]
constant	3.330***	2.883***	4.228***	3.876***	3.850***
	[0.097]	[0.395]	[0.174]	[0.542]	[0.546]
Observations	7935	7827	7935	7827	7827
R-squared	0.13	0.19			
Hansen J test			0.199	0.1458	0.003

Note: *lprob* here is the subjective probability of AI loss at the household level.

Table 18: AI indemnity-based insurance first stage estimation results

Regressors	IV1 *lprob*	IV2 *lprob*	IV3 *lprob*
ex		3.319***	
		[0.249]	
*exd*1	4.788***		4.164***
	[0.367]		[0.358]
*exd*2	2.618***		1.355
	[0.715]		[0.729]
exwater	3.285***	0.546***	0.565***
	[0.112]	[0.124]	[0.125]
lpay	0.157***	0.216***	0.213***
	[0.023]	[0.023]	[0.023]
lasset		0.025	0.021
		[0.062]	[0.062]
lincome		0.316***	0.317***
		[0.104]	[0.104]
rural		0.381	0.32
		[0.339]	[0.339]
edusec		-0.302	-0.331
		[0.248]	[0.248]
*province*1		-10.132***	-10.186***
		[0.287]	[0.286]
*province*2		-10.186***	-10.199***
		[0.324]	[0.324]
*province*3		-4.665***	-4.718***
		[0.282]	[0.282]
partialpay		0.938**	0.878**
		[0.418]	[0.417]
borrowing		0.066	0.061
		[0.215]	[0.215]
hhsize		-0.138**	-0.138**
		[0.064]	[0.064]
agehead		0.069	0.073
		[0.047]	[0.047]
ageheadsq		-0.001	-0.00056
		[0.000]	[0.0004]
wife		-0.101	-0.127
		[0.240]	[0.241]
husband		-1.780***	-1.813***
		[0.579]	[0.574]
son		-0.67	-0.705
		[0.460]	[0.457]
daughter		0.701	0.709
		[0.706]	[0.704]
constant	-24.161***	-19.633***	-19.628***
	[0.220]	[1.323]	[1.323]
Observations	7935	7827	7827
R-squared	0.14	0.28	0.28
F stats	448.85	103.05	74.66

Note 1: *lprob* here is the subjective probability of AI loss at the household level.

Note 2: The F statistics are for the excluded instruments.

Table 19: Flooding indemnity-based insurance (WTP) regressions (*lwtp*)

Regressors	OLS	OLS	IV1	IV2
	lwtp	*lwtp*	*lwtp*	*lwtp*
lprob	0.009***	-0.007*	0.077***	0.140***
	[0.003]	[0.004]	[0.006]	[0.019]
lindem	0.262***	0.298***	0.243***	0.252***
	[0.009]	[0.009]	[0.010]	[0.011]
lasset		0.102***		0.073***
		[0.018]		[0.021]
lincome		-0.026		0.037
		[0.030]		[0.034]
rural		-0.481***		-0.547***
		[0.086]		[0.100]
edusec		0.268***		0.267***
		[0.066]		[0.075]
*province*1		-1.233***		1.028***
		[0.080]		[0.305]
*province*2		-0.573***		1.498***
		[0.087]		[0.287]
*province*3		-1.622***		-0.645***
		[0.080]		[0.158]
hhsize		-0.091***		-0.080***
		[0.017]		[0.020]
partialpay		-0.681***		-0.797***
		[0.135]		[0.139]
borrowing		0.162***		0.037
		[0.061]		[0.070]
agehead		0.012		-0.021
		[0.013]		[0.015]
ageheadsq		-0.00018		0.000
		[0.00012]		[0.000]
wife		0.292***		0.328***
		[0.071]		[0.078]
husband		0.042		0.359
		[0.217]		[0.241]
son		0.633***		0.783***
		[0.143]		[0.151]
daughter		0.299		0.086
		[0.238]		[0.264]
*floodloss*1		0.710*		2.417***
		[0.373]		[0.472]
*floodloss*2		0.442***		0.404***
		[0.097]		[0.107]
*floodloss*3		-0.092		-0.214
		[0.207]		[0.265]
*floodloss*4		1.587***		2.358***
		[0.199]		[0.264]
constant	2.505	2.777***	4.038***	5.569***
	[0.125]	[0.387]	[0.170]	[0.542]
Observations	7935	7827	7935	7827
R-squared	0.14	0.22		
Hansen J test			0.0003	0.2263

Note: *lprob* here is the subjective probability of flooding loss at the household level.

Table 20: Flooding indemnity-based insurance first stage regressions (*lprob*)

Regressors	IV1 *lprob*	IV2 *lprob*	Regressors	IV1 *lprob*	IV2 *lprob*
*exd*1	9.800***	3.923***	*floodloss*1		-17.901***
	[0.256]	[0.254]			[2.184]
*exd*2	2.366***	1.112***	*floodloss*2		0.163
	[0.356]	[0.368]			[0.682]
exlandslide	-0.835	1.724**	*floodloss*3		0.655
	[0.994]	[0.849]			[1.240]
exAI	0.851***	-0.01	*floodloss*4		-4.927***
	[0.231]	[0.476]			[0.624]
exepi	0.627***	0.334***	*constant*	-24.344***	-19.714***
	[0.133]	[0.099]		[0.329]	[1.28]
lindem	0.224***	0.268***			
	[0.027]	[0.026]			
lasset		0.145**			
		[0.058]			
lincome		-0.389***			
		[0.094]			
rural		0.282			
		[0.327]			
edusec		-0.133			
		[0.236]			
*province*1		-13.213***			
		[0.261]			
*province*2		-12.579***			
		[0.278]			
*province*3		-5.678***			
		[0.252]			
partialpay		0.709*			
		[0.385]			
borrowing		0.595***			
		[0.201]			
hhsize		-0.099			
		[0.061]			
agehead		0.205***			
		[0.043]			
ageheadsq		-0.002***			
		[0.000]			
wife		-0.392*			
		[0.230]			
husband		-2.236***			
		[0.579]			
son		-1.229***			
		[0.387]			
daughter		1.718**			
		[0.719]			
Observations	7935	7827			
R-squared	0.2	0.44			
F stats	536.95	93.61			

Note 1: *lprob* here is the subjective probability of flooding loss at the household level.
Note 2: The F statistics are for the excluded instruments.