ANNALS OF ECONOMICS AND FINANCE **17-1**, 179–206 (2016)

Insurance Contracts with Adverse Selection When the Insurer Has Ambiguity about the Composition of the Consumers

Mingli Zheng*

Department of Economics, University of Macau, Macau, China
E-mail: mlzheng@umac.mo

Chong Wang

Department of Economics, University of Macau, Macau, China

and

Chaozheng Li

Department of Economics, University of Kansas, USA

In this paper, we consider the optimal contract in a monopolistic insurance market when the insurer has ambiguity about the composition of the consumers. When there are only two types of consumers, we find that high-risk consumers are fully insured, whereas low-risk consumers are only partially insured. For an ambiguity averse insurer, as ambiguity increases, the optimal menu of contracts moves toward the one that equalizes the profits earned by the insurer from the two types of consumers. The insurer may offer the same menu of contracts even if her prior belief changes. For an ambiguity seeking insurer, when the degree of ambiguity increases, the optimal menu moves away from the menu that equalizes the profits earned from the two types of consumers.

Key Words: Adverse selection; Monopoly; Insurance; Ambiguity; ε-contaminated prior.

JEL Classification Numbers: D82, D42.

* Corresponding author. Mingli Zheng's research is supported by the research grant of University of Macau (MYRG2014-00032-FSS).

1529-7373/2016

1. INTRODUCTION

Consumer information is important for firms. Firms spend a lot of resources in collecting information about the consumers, and they adjust their strategies through interactions with the consumers (Cao and Sun, 2007). In insurance markets, consumers may have private information, and it is hard for the insurers to get useful information of the consumers.

In exploring the problem of private information in insurance market, Rothschild and Stiglitz (1976) considered the problem of adverse selection in a competitive market and Stiglitz (1977) considered the adverse selection problem in a monopolistic market. From these works, when there exists private information, an insurer does not need to know the type of each consumer. If the distribution of different types of consumers is known, an insurer can offer a menu of contracts from which consumers can self-select. It is possible that the consumers reveal their information truthfully and the profit of the insurer can increase.

In designing the optimal menu of contracts, the information about the proportion of different types of consumers is crucial. However, when an insurer enters a new market or faces a new pool of customers such as new immigrants, they cannot confidently use a best-guess subjective prior about the composition of the consumers to design the optimal menu of contracts, and the insurer has to face ambiguity. For short term or renegotiable contracts, the insurer has opportunity to learn from the choices of the consumers. However, for long term contracts that cannot be renegotiated in the short run, the insurer has no opportunity to learn. The insurer has to design the contract by considering the ambiguity about the consumer composition, and the sub-optimality of the contract can be costly. Even if an insurer has opportunities to learn from the consumers' choices and renegotiate the contracts, there may be bunching in the optimal menu of contracts (though there is no bunching in this paper for two types of buyers), which prevents an insurer from learning effectively the distribution of the types of consumers; see Berg and Ehtamo (2009) for a model of firm's learning for screening the types of consumers. Therefore, ambiguity about the consumer composition matters if we need to know how to design the optimal contract when learning opportunity doesn't exist. If the opportunity exists, we need to know the effect of ambiguity on the learning of the insurer.

In this paper, we consider the optimal insurance contract when a monopolist insurer faces ambiguity with regards to the proportion of different types of consumers. Most of the studies of ambiguity in the insurance market considered the ambiguity about the distribution of loss. For example, it is nature to assume that both insurers and customers may have ambiguity about the accident rate in catastrophic risks, as they do not have chances to learn the accident rate from available information. Because large organizations may have advantage in collecting information, it is usually assumed in the literature that insurers are ambiguity neutral while consumers are ambiguity averse, or insurers are less ambiguous than consumers[1]. In our paper, we assume that the risk is well-understood, and neither the insurer nor the consumers face ambiguity about the distribution of loss. Since the insurer can pool the risk, we assume that consumers are risk averse while the insurer is risk neutral, as in most of the literature.

When ambiguity exists, a decision maker can be ambiguity averse, ambiguity neutral, or ambiguity seeking. Empirical evidences suggested that decision-makers are heterogeneous in their attitude towards ambiguity, and many studies found that most decision-makers are ambiguity averse. For example, from a survey that elicits respondents' ambiguity attitudes using questions based on the classical Ellsberg urn experiment, Dimmock et al. (2013) found that for a representative sample of U.S. households, half of them are ambiguity averse, 12% are ambiguity neutral, and 37% are ambiguity seeking. There is no direct empirical study exploring an insurer's ambiguity attitude when there is ambiguity about the proportion of the types of customers. To study the optimal insurance contract when the insurer has ambiguity about the composition of the customers, we explore the optimal insurance policy for an ambiguity averse insurer, and use the similar method to briefly explore the optimal insurance for an ambiguity seeking insurer.

We use the Choquet expected utility to incorporate the insurer's ambiguity about the proportion of different types of customers. The study of ambiguity is originated from the work of Knight (1921). The choice of an agent in the famous example of Ellsberg paradox (Ellsberg, 1961) cannot be represented by a probability distribution. Ambiguity can be described by the expected utility with the existence of multiple priors, or Choquet expected utility with respect to a capacity (see Gilboa, 1987; Gilboa and Schmeidler, 1989). The Choquet expected utility theory is widely used to model ambiguity because of its solid axiomatic foundation and tractable preference representation. If there are only two types of consumers, as is assumed in our paper, we can see that Choquet integral with respect to a capacity can be described by the integral with respect to ε-contamination of a probabilistic prior. A monopolistic insurer may hold a subjective prior on the consumer distribution based on her experiences or market researches. However, the insurer is not confident about her subjective prior, believing

[1] In the insurance market, when there is ambiguity about the distribution of loss, empirical evidence suggested the prevalence of ambiguity and market players are ambiguity averse; see Hogarth and Kunreuther (1989), Kunreuther, Hogarth, and Meszaros (1993). Hogarth and Kunreuther (1989) concluded that "firms show greater aversion to ambiguity than consumers".

the prior to be erroneous with probability ε (the true distribution could be any distribution). In this case, the insurer's prior can be represented by ε-contamination of her probabilistic prior. The degree of ambiguity can be represented by value ε, and an ambiguity averse (ambiguity seeking) decision-maker assumes that the worst (best) outcome will occur in the presence of ambiguity.

In this paper, we examine a monopolistic insurance market. The study of a monopolist market can be useful when market power exists. It is known that the model of a monopolistic insurance market with private information is more complicated than the standard principal-agent model with private information as in Maskin and Riley (1984). A monopolistic insurance market is characterized by (1) wealth effects for risk-averse consumers; (2) common valuation problem in which consumer type enters into the principals objective function; and (3) the type-dependent reservation utilities of consumers (see Chade and Schelee, 2012). The solution to the practical problem of the optimal insurance contract when there is ambiguity about the proportion of customers can be more complicated.

We assume that a consumer is either high-type or low-type, with a high or low probability of a given loss. The consumers know their types, but the insurer cannot observe the type of each consumer. Furthermore, the insurer has ambiguity about the distribution of consumer types. Under a principal-agent framework, the lowest profit earned by the insurer can be obtained from the high-type or the low-type consumers under different menus of contracts. An ambiguity averse insurer assumes that the extra ε-proportion of the consumers are the type that gives her the lowest profit. Therefore, an ambiguity-averse insurer may adjust her prior in different ways for different menus of contracts. An ambiguity seeking insurer also adjusts her prior for different menus of contracts.

We find that with ambiguity (no matter the insurer is ambiguity averse or ambiguity seeking), the high-type consumers acquire full coverage, but the low-type consumers obtain less than full coverage. For an ambiguity averse insurer, we find that as the degree of ambiguity increases (i.e., as ε increases), the optimal menu of contracts moves toward a menu (the "attraction" menu) in which the profits that the insurer earns from the two types of consumers are the same. The coverage of the low-type consumer can increase or decrease when ambiguity increases. Once the attraction menu of contracts is achieved, the optimal menu will no longer change even as ε continues to increase. An ambiguity averse insurer may set the same menu of contracts (which is the attraction menu) for a range of prior beliefs. For an ambiguity seeking insurer, when the degree of ambiguity increases, the menu of contract moves away from the menu from which the profits from two types of buyers are equal. The reason for our result is because of the shifting of the insurer's prior for different menu of contracts.

Our result suggests that when there are two types of consumers, an insurer who has chance to learn and renegotiate the contract can learn effectively the distribution of consumers even if ambiguity about the consumer distribution exists. Ambiguity is resolved and the insurer can set the optimal contract using the revealed proportion of consumers in the renegotiation. For a varying population, ambiguity may not be resolved because of the introduction of new ambiguity. For long term contracts that cannot be renegotiated, the ex-ante optimal contracts depends on the insurers prior, the degree of ambiguity and the insurer's attitude towards ambiguity.

As far as we know, there is no discussions exploring the effect of ambiguity in the distribution of the consumers in the literature of the insurance market. However, ambiguity in the distribution of types is already discussed in other fields of the economic studies. For example, Salo and Weber (1995) discussed bidders' ambiguity about the distribution of types of other bidders in the first-price auction. The paper used Choquet expected utility with respect to a form of capacity which is a transformation of an additive probability (Quiggin, 1982), and showed that ambiguity aversion causes the bidders to underestimate their chances of winning the auction. Zheng et al. (2015) discussed a seller's ambiguity of the distribution of consumers in the simple nonlinear pricing problem in the framework of Maskin and Riley (1984). Their paper used Choquet expected utility with respect to ε-contaminated prior, and showed that ambiguity causes bunching for low valuation consumers.

There is a lot of discussion of the effects of ambiguity in the distribution of loss in the insurance market. For example, Gollier (2014) considered the optimal insurance policy between an ambiguity averse policy holder and an ambiguity neutral insurer facing linear transaction costs. The paper used the smooth ambiguity aversion model of Klibanoff et al. (2005), where the distribution function of the loss is parameterized by θ that can take several possible values. The paper implied that under some conditions, the optimal indemnity schedule contains a disappearing deductible. Amarante et al. (2015) discussed the insurance between an insurer who has ambiguous beliefs about the loss and an ambiguity-neutral policy-holder.

For a competitive insurance market, Anwar and Zheng (2013) considered a competitive insurance market with identical insurers and identical customers, but both insurer and customers have ambiguity about the probability of loss and ambiguity is described by ε-contaminated prior. They showed that in the equilibrium the customers can be full insured or under-insured. Koufopoulos and Kozhan (2015) introduced ambiguity in the consumers' perception about loss in a competitive insurance model with asymmetric information (with two types of insures) and adverse selection, and the consumers are ambiguity averse while the insurers are ambiguity neutral. In the paper, ambiguity was described by a set of priors about the true probability of accident. It was shown that unique equilibrium exists and the equilibrium can be pooling or separating, and the increase of ambiguity can lead to a strict Pareto improvement. In a slightly different framework, Huang, Snow, and Tzeng (2015) also assumed that there is ambiguity in the consumers' perception about the general loss in a competitive insurance model. The ambiguity reduces willingness to bear wealth risk for ambiguity-averse consumers. In the equilibrium, there could exist adverse selection (with a positive correlation between insurance coverage and expected claim frequency) or advantageous selection.

For a monopolistic insurance market, Jeleva and Villeneuve (2004) explored the adverse selection problem where two types of consumers know their types but they have ambiguity about the distribution of loss, while the monopolistic insurer is ambiguity neutral but cannot observe the type of the customers. Ambiguity of loss was described by the rank dependent expected utility, which is a Choquet integral with respect to a capacity described by a transformation of an additive probability proposed by Quiggin (1982). As the marginal values for a better coverage across types may be ranked by types and by the insurer in opposite ways, the optimal menu of contracts can be pooling over a continuum of parameters even though the single-crossing property is satisfied. Vergote (2010) made similar assumptions as in Jeleva and Villeneuve (2004), but interpreted one-sided ambiguity as a signaling game in which the principal has better quality of information, in which a contractual offer leaves room for the transmission of information to the qualitatively least informed party. The paper showed that both pooling and separating equilibria exist but only the full insurance pooling contract is a Perfect sequential equilibrium. We assume that there is no ambiguity about the risk (the risk is well-understood), but the insurer has ambiguity about the composition of the customers. In our model, ambiguity leads to an adjustment of the insurers belief about the composition of consumers for different menu of contracts.

The rest of the paper is organized as follows. Section 2 is a brief introduction of decision-making in the presence of ambiguity. Section 3 introduces the adverse selection model of an insurance market without ambiguity. Section 4 explores the optimal menu of contracts with ambiguity, and Section 5 concludes the paper.

2. DECISION MAKING UNDER AMBIGUITY

Decision making under ambiguity was first discussed in Knight (1921). The famous Ellsberg paradox (Ellsberg, 1961) illustrates a situation in which a decision-maker's choice cannot be explained by a subjective additive probability. For example, consider an urn containing 100 balls. Out of these balls, 30 are red, and the rest are either black or yellow, but no further information is available. A subject is asked to bet on the color of a ball drawn from the urn. To "bet on red" means that the subject will receive a prize of \$10 if the ball drawn from the urn is red and \$0 if the ball is not red. A test subject is given the following 4 options: (I) "a bet on red," (II) "a bet on black," (III) "a bet on red or yellow," (IV) "a bet on black or yellow." The subject is asked to choose between bets I and II, and between bets III and IV. Most subjects prefer bet I over bet II, and bet IV over bet III. This preference violates the sure-thing principle, which requires that bet III is preferred over bet IV once bet I is preferred over bet II (because these two pairs differ only in the payoff when a yellow ball is drawn). In this case, the choice of a decision-maker cannot be described by any subjective probability.

To solve the problem of ambiguity as in the Ellsberg paradox, researchers have attempted to extend the expected utility theory by using a Choquet expected utility with respect to a capacity (a non-additive probability) or by using an expected utility with the existence of multi-priors. When there are multi-priors, the set of probability is sometimes called the set of expert opinions. The existence of multi-priors was used to explain the puzzles in asset return (Chen and Epstein, 2002) and in robust control and model uncertainty (Hansen and Sargent, 2001). The Choquet expected utility is more widely used in applications for its solid axiomatic foundation and its tractability. Gilboa (1987) justified the Choquet utility approach by using consistency conditions on a decision-makers preferences without the sure-thing principle. The Choquet integral is the integral with respect to a non-additive probability (or capacity), and it can be written as a rank dependent weighted sum.

A widely used special form of Choquet integral in describing ambiguity in the literature is the Choquet integral with respect to ε-contamination of a given prior (or a simple capacity as in Eichberger and Kelsey 1999). If P is an additive probability distribution over state space S, then the ε-contamination of P is defined by a capacity ν such that

$$\nu(A) = (1-\varepsilon)P(A) \text{ for any measurable set } A \neq S, \text{ and } \nu(S) = 1.$$

The Choquet expected utility of a decision maker with the above capacity can be written as

$$(1-\varepsilon)E^P u + \varepsilon \min_{s \in S} u(s),$$

where $u : S \to R$ is a von Neumann-Morgenstern utility function, and $E^P u$ is the expected utility with respect to probability distribution P.

The above model of ε-contamination of a prior P can be intuitively interpreted as follows. A decision-maker holds a subjective prior belief P about the states of the world. However, she is not confident about her belief, and assumes that with a probability ε, her belief may be incorrect (the correct belief can be any probability). A decision maker assumes that the worst outcome will occur when her belief is incorrect.

To describe the behavior of an ambiguity seeking decision maker by ε-contamination of a given prior[2], we need to use the capacity defined by

$$\nu(A) = (1-\varepsilon)P(A) + \varepsilon \text{ for any measurable set } A \neq \emptyset, \text{ and } \nu(\emptyset) = 0.$$

The Choquet expected utility with respect to this capacity becomes

$$(1-\varepsilon)E^P u + \varepsilon \max_{s \in S} u(s).$$

A decision-maker with the above capacity assumes that with probability ε, her prior belief P may be incorrect. The decision maker assumes that the best outcome will occur when her belief is incorrect.

For a decision maker having ε-contamination of a given prior, researchers agree that the degree of ambiguity can be represented by value ε, and an ambiguity averse (ambiguity seeking) decision-maker assumes that the worst (best) outcome will occur in the presence of ambiguity. (When $\varepsilon = 0$, the decision-maker is ambiguity neutral). However, for a general Choquet expected utility (not necessarily in the form of integral with respect to ε-contaminated prior), there is no results about the measurement of the degree of ambiguity or ambiguity aversion, and there are hot debates about the separation of ambiguity and attitude towards ambiguity; see, for example, Ghirardato et al. (2004).

There exist many different ways to incorporate ambiguity in the study of the effects of ambiguity in the insurance market, as we discussed in the introduction of the paper. In our paper we assume there are only two types of customers, and the state space has only two states with $S = \{H, L\}$. The following Lemma can justify the use of ε-contaminated prior for a Choquet integral in our case.

Lemma 1. *If there are only two states, a Choquet expected utility can always be represented by an integral with respect to ε contamination of a prior P. Specifically, if ν is convex with $\nu(H) + \nu(L) < 1$, then there exist a unique pair (P, ε) such that $\int u d\nu = (1-\varepsilon)E^P u + \varepsilon \min_{s \in S} u(s)$. If ν is concave with $\nu(H) + \nu(L) > 1$, then there exist a unique pair (P, ε) such that $\int u d\nu = (1-\varepsilon)E^P u + \varepsilon \max_{s \in S} u(s)$.*

[2] In the literature, most of the work only discussed the ambiguity aversion in the ε contamination model.

Therefore, when there are only two states, a Choquet expected utility can always be written as integral with respect to ε-contamination of a prior, and the convexity (concavity) of the capacity implies ambiguity aversion (ambiguity seeking). However, if there are more than two states, a Choquet expected utility can no longer always be written as an integral with respect to ε contamination of a prior.

Empirical or experimental evidences suggest that most decision-makers facing ambiguity are ambiguity averse, while some of them are ambiguity seeking or ambiguity neutral. We consider a monopolist insurer who has ambiguity about the composition of the customers, and we focus on the case with an ambiguity averse insurer, while the case for an ambiguity-seeking insurer will be briefly discussed in a similar way. First, we look at some results when there is no ambiguity.

3. THE OPTIMAL INSURANCE CONTRACT WHEN THERE IS NO AMBIGUITY

We assume that there is only one insurer in the market and we use the notation as in Chade and Schlee (2012), Szalay (2008). A risk-neutral monopolistic insurer faces two types of consumers (the insurees). All the consumers have initial wealth $w > 0$, and they face a potential loss $l \in (0, w)$. The consumers know their own types. The high-type consumers face a loss with a high probability $\bar{\theta}$, and the low-type consumers face a loss with a low probability $\underline{\theta}$, where $0 < \underline{\theta} < \bar{\theta} < 1$. A consumer's preference over risks is represented by a strictly increasing, strictly concave von Neumann-Morgenstern utility function u. A type-θ consumer's expected utility without insurance is $\theta u(w - l) + (1 - \theta)u(w)$.

The insurer cannot observe a consumer's type. As in the standard adverse selection model, a monopolistic insurer offers a menu of contracts to the consumers, and the consumers can select from the offered menu of contracts. An insurance contract has the form (x, t), where t is the premium (insurance price), and $x - t$ is a consumer's net reimbursement in case of loss. If a consumer of type-θ accepts a contract (x, t), the profit of the insurer is $\pi(x, t, \theta) = t - \theta x$.

By accepting a contract (x, t), the expected utility of a type-θ consumer is $U(x, t, \theta) = \theta u(w - l + x - t) + (1 - \theta)u(w - t)$. As in most of the literature, we assume that function U satisfies the strict single crossing property: if $(x', t') \geq (x, t)$ and $\theta' > \theta$, then $U(x', t', \theta) \geq U(x, t, \theta)$ implies that $U(x', t', \theta') > U(x, t, \theta')$. When function u is differentiable and strictly increasing, the strict single crossing property implies that a higher type consumer is willing to pay more for a marginal increase in insurance.

Using the revelation principle, we only need to consider the menus of contracts in which consumers reveal truthfully their types. For a market with only two types of consumers, a menu of contracts becomes $(\underline{x}, \underline{t}, \bar{x}, \bar{t})$, where $(\underline{x}, \underline{t})$ is the contract offered to the low-type consumers and $(\bar{x}, \bar{t})$ is the contract offered to the high-type consumers.

A menu of contracts should satisfy both incentive and participation constraints. A low-type consumer prefers contract $(\underline{x}, \underline{t})$ over contract $(\bar{x}, \bar{t})$ or no contract at all: $U(\underline{x}, \underline{t}, \underline{\theta}) \geq U(\bar{x}, \bar{t}, \underline{\theta})$, $U(\underline{x}, \underline{t}, \underline{\theta}) \geq U(0, 0, \underline{\theta})$. Likewise, a high-type consumer prefers contract $(\bar{x}, \bar{t})$ over contract $(\underline{x}, \underline{t})$ or no contract at all: $U(\bar{x}, \bar{t}, \bar{\theta}) \geq U(\underline{x}, \underline{t}, \bar{\theta})$, $U(\bar{x}, \bar{t}, \bar{\theta}) \geq U(0, 0, \bar{\theta})$.

For a menu of contracts $(\underline{x}, \underline{t}, \bar{x}, \bar{t})$, the insurer's profit earned from a low-type consumer is $\pi_L = \underline{t} - \underline{\theta x}$ and her profit earned from a high-type consumer is $\pi_H = \bar{t} - \bar{\theta}\bar{x}$.

To simplify our analysis, we define the following notations (where subindex "a" stands for accident and "na" stands for no-accident):

$$u_a = u(w-l), u_{na} = u(w), \underline{u}_a = u(w - l + \underline{x} - \underline{t}), \underline{u}_{na} = u(w - \underline{t}),$$

$$\bar{u}_a = u(w - l + \bar{x} - \bar{t}), \text{ and } \bar{u}_{na} = u(w - \bar{t}).$$

The utility of a non-insured consumer is u_a when loss occurs and u_{na} when no loss occurs. Under menu of contracts $(\underline{x}, \underline{t}, \bar{x}, \bar{t})$, the utility of a low-type consumer is $\underline{u}_a$ when loss occurs and $\underline{u}_{na}$ when no loss occurs; the utilities of a high-type consumer are $\bar{u}_a$ and $\bar{u}_{na}$, respectively. Because function u is strictly increasing, we can describe menu of contracts $(\underline{x}, \underline{t}, \bar{x}, \bar{t})$ by $(\underline{u}_a, \underline{u}_{na}, \bar{u}_a, \bar{u}_{na})$. Such a transformation can simplify our analysis. The incentive and participation constraints can be written using the terms $(\underline{u}_a, \underline{u}_{na}, \bar{u}_a, \bar{u}_{na})$.

Let v be the inverse of function u. Function v is strictly convex and strictly increasing. With $\underline{u}_a$, $\underline{u}_{na}$, $\bar{u}_a$, $\bar{u}_{na}$, the insurer's profit $\pi_L = \underline{t} - \underline{\theta x}$ from a low-type consumer and her profit $\pi_H = \bar{t} - \bar{\theta}\bar{x}$ from a high-type consumer can be written as:

$$\begin{aligned} \pi_L &= -\underline{\theta} v(\underline{u}_a) - (1 - \underline{\theta}) v(\underline{u}_{na}) + w - \underline{\theta} l, && (1) \\ \pi_H &= = \bar{\theta} v(\bar{u}_a) - (1 - \bar{\theta}) v(\bar{u}_{na}) + w - \bar{\theta} l. && (2) \end{aligned}$$

We assume now that the insurer has no ambiguity about the composition of the customers. The insurer has a subjective prior belief $P = (p, 1-p)$ about the distribution of consumer types: a consumer is of low-type with probability p and a consumer is of high-type with probability $1-p$ (or equivalently, a proportion p of the consumers are of the low-type and a proportion $1-p$ of the consumers are of the high-type). This belief can be the insurer's best guess based on her experience or market research.

For a menu of contracts $(\underline{u}_a, \underline{u}_{na}, \bar{u}_a, \bar{u}_{na})$, when there is no ambiguity, the expected profit of the monopolistic insurer is

$$\pi = p\pi_L + (1-p)\pi_H,$$

where π_L and π_H are defined in Eq. (1) and Eq. (2). The optimal menu of contracts is the one that maximizes the insurer's expected profit π. We denote the optimal menu of contracts as $(\underline{u}_a^*, \underline{u}_{na}^*, \bar{u}_a^*, \bar{u}_{na}^*)$ in the following. (For convenience, we sometimes express $(\underline{u}_a^*, \underline{u}_{na}^*, \bar{u}_a^*, \bar{u}_{na}^*)$ simply as $(\underline{u}_a, \underline{u}_{na}, \bar{u}_a, \bar{u}_{na})$ when no confusion arises).

When the insurer has no ambiguity about the composition of the customers, the optimal menu of contracts that maximizes expected profit π satisfies the following properties:

Lemma 2.

1.In the optimal menu of contracts in the form of $(\underline{x}^*, \underline{t}^*, \bar{x}^*, \bar{t}^*)$, we have

(a)$\bar{x}^* \geq \underline{x}^* \geq 0; \bar{t}^* \geq \underline{t}^* \geq 0; \bar{x}^* - \bar{t}^* \geq \underline{x}^* - \underline{t}^* \geq 0$,

(b)the high-type consumers are fully insured with $\bar{x} = l$,

(c)the incentive constraint is binding for a high-type consumer,

(d)the participation constraint is binding for a low-type consumer.

2.For the optimal menu of contracts $(\underline{u}_a^*, \underline{u}_{na}^*, \bar{u}_a^*, \bar{u}_{na}^*)$, we have

$$u_a \leq \underline{u}_a^* \leq \bar{u}_a^* = \bar{u}_{na}^* \leq \underline{u}_{na}^* \leq u_{na}, \tag{3}$$

$$\underline{u}_a^* = u_a + \frac{1-\underline{\theta}}{\underline{\theta}} u_{na} - \frac{1-\underline{\theta}}{\underline{\theta}} \underline{u}_{na}^* := s(\underline{u}_{na}^*), \tag{4}$$

$$\bar{u}_a^* = \bar{u}_{na}^* = \bar{\theta}\left(u_a + \frac{1-\underline{\theta}}{\underline{\theta}} u_{na}\right) - \frac{\bar{\theta}-\underline{\theta}}{\underline{\theta}} \underline{u}_{na}^* := z(\underline{u}_{na}^*). \tag{5}$$

Stiglitz (1977) derived (b)(d) for the two-type case and for the continuum case with a smooth positive density when no ambiguity exists. Chade and Schlee (2012) proved (a), (b), and (d) for a distribution with arbitrary types of consumers under no ambiguity. The properties (a), (b), and (d) of Lemma 2 in our paper can be considered direct outcomes of Lemma 2 and Theorem 1 in Chade and Schlee (2012). Since the proof is standard, we omit the proof here. From Lemma 2, we have Proposition 1 (see the proof in the appendix).

PROPOSITION 1. *The optimal menu of contracts* $(\underline{u}_a^*, \underline{u}_{na}^*, \bar{u}_a^*, \bar{u}_{na}^*)$ *takes the form* $(s(\underline{u}_{na}), \underline{u}_{na}, z(\underline{u}_{na}), z(\underline{u}_{na}))$ *with* $\underline{u}_{na} \in [\underline{\theta} u_a + (1-\underline{\theta}) u_{na}, u_{na}]$.

From Proposition 1, the choice variables in the monopolist's optimal menu of contracts are reduced from four variables $(\underline{u}_a, \underline{u}_{na}, \bar{u}_a, \bar{u}_{na})$ to only one, $\underline{u}_{na}$. Once $\underline{u}_{na}^*$ in the optimal menu of contracts is determined, the other three variables in the optimal menu of contracts — $\underline{u}_a^*$, $\bar{u}_a^*$, and $\bar{u}_{na}^*$ — will be determined according to Eq. (4) and Eq. (5). For simplicity of notation, we define a menu of contracts in Proposition 1 to be a feasible menu of contracts:

DEFINITION 3.1. A feasible menu of contracts is a menu in the form of $(s(\underline{u}_{na}), \underline{u}_{na}, z(\underline{u}_{na}), z(\underline{u}_{na}))$ with $\underline{u}_{na} \in [\underline{\theta} u_a + (1-\underline{\theta}) u_{na}, u_{na}]$.

From $\underline{u}_{na} = u(w - \underline{t})$, we obtain $\underline{t} = w - v(\underline{u}_{na})$, and the choice of $\underline{u}_{na}$ is the choice of the premium $\underline{t}$ for a low-type consumer. From Eq. (4) and Eq. (5), $\underline{u}_a$ and $\bar{u}_a$ (thus, also $\bar{u}_{na}$) are decreasing function of $\underline{u}_{na}$ in the feasible menus of contracts. From $\underline{u}_a = u(w - l + \underline{x} - \underline{t})$, we get $\underline{x} - \underline{t} = v(\underline{u}_a) - w + l$. When the premium $\underline{t}$ for a low-type consumer increases (i.e., when $\underline{u}_{na}$ decreases), the net reimbursement to a low-type consumer $\underline{x} - \underline{t}$ will increase. Therefore, we can refer to the expression "the decrease in $\underline{u}_{na}$" as a more intuitive "the increase in the coverage of a low-type consumer". From $\bar{u}_{na} = u(w - \bar{t})$, when the premium $\underline{t}$ for a low-type consumer increases, the premium $\bar{t}$ for a high-type consumer decreases. Thus, when the coverage of a low-type consumer increases in a feasible menu of contracts, a high-type consumer will pay a lower premium (but she remains fully insured).

We now discuss the choice of the optimal value $\underline{u}_{na}^*$.

The feasible menus of contracts when $\underline{u}_{na}$ takes the end values of the interval $[\underline{\theta} u_a + (1-\underline{\theta}) u_{na}, u_{na}]$ correspond to two special cases. If $\underline{u}_{na} = \underline{\theta} u_a + (1-\underline{\theta}) u_{na}$, then $\underline{u}_a = \underline{u}_{na} = \bar{u}_a = \bar{u}_{na} = \underline{\theta} u_a + (1-\underline{\theta}) u_{na}$. The contract offered to a low-type consumer is the first-best contract when there is no private information, and the utilities of all the consumers are the same. If $\underline{u}_{na} = u_{na}$ (i.e., $\underline{t} = 0$), then $\underline{u}_a = u_a$ and $\bar{u}_a = \bar{u}_{na} = \bar{\theta} u_a + (1-\bar{\theta}) u_n a$. The low-type consumers are excluded from the market and the contract offered to a high-type consumer is the first-best contract when there is no private information. If $\underline{\theta} u_a + (1-\underline{\theta}) u_{na} \le \underline{u}_{na} < u_{na}$, then both types of consumers are served under the corresponding feasible menu of contracts.

For a feasible menu of contracts characterized by $\underline{u}_{na}$, the insurer's profit from a low-type consumer becomes $\pi_L(\underline{u}_{na})$, and the profit from a high-type consumer becomes $\pi_H(\underline{u}_{na})$. We can show that for a feasible menu of contracts characterized by $\underline{u}_{na}$, the profit $\pi_L(\underline{u}_{na})$ of the insurer from a low-type consumer is a decreasing function of $\underline{u}_{na}$, and the profit $\pi_H(\underline{u}_{na})$ from a high-type consumer is an increasing function of $\underline{u}_{na}$. We obtain the following Lemma (see proof in the appendix).

LEMMA 3. *A unique value $\hat{u}$ with $\underline{\theta}u_a + (1-\underline{\theta})u_{na} < \hat{u} < u_{na}$ exists such that*

$$\begin{cases} \pi_H(\underline{u}_{na}) < \pi_L(\underline{u}_{na}) & \text{if } \underline{u}_{na} < \hat{u}, \\ \pi_H(\underline{u}_{na}) = \pi_L(\underline{u}_{na}) & \text{if } \underline{u}_{na} = \hat{u}, \\ \pi_L(\underline{u}_{na}) < \pi_H(\underline{u}_{na}) & \text{if } \underline{u}_{na} > \hat{u}. \end{cases}$$

It is worth noting that the value $\hat{u}$ does not depend on the value of p in prior P. Lemma 3 points out that for the feasible menu of contracts with $\underline{u}_{na} = \hat{u}$, the profits of the insurer from the two types of consumers are equal. For feasible menus of contracts with $\underline{u}_{na} < \hat{u}$, the profit of the insurer from a high-type consumer is lower, and for those with $\underline{u}_{na} > \hat{u}$, the profit from a low-type consumer is lower. Intuitively, for feasible menus of contracts, if $\underline{u}_{na}$ decreases (so that the coverage for a low-type consumer increases), the premium $\underline{t}$ paid by a low-type consumer will increase, whereas that paid by a high-type consumer will decrease; the net reimbursement $x - t$ in case of loss for both types of consumers will increase. Overall, the insurers profit $t - \theta x$ from a high-type consumer will be lower than that from a low-type consumer if $\underline{u}_{na} < \hat{u}$.

The expected profit of the insurer with prior $P = (p, 1-p)$ from a feasible menu of contracts characterized by $\underline{u}_{na}$ is described by a function of two variables $\underline{u}_{na}$ and p, and we denote it as $S(\underline{u}_{na}, p)$ with

$$S(\underline{u}_{na}, p) := p\pi_L(\underline{u}_{na}) + (1-p)\pi_H(\underline{u}_{na}) \tag{6}$$

where $\underline{u}_{na} \in [\underline{\theta}u_a + (1-\underline{\theta})u_{na}, u_{na}]$.

We denote the value $\underline{u}_{na}$ that maximizes function $S(\underline{u}_{na}, p)$ for a given value p as $\tilde{\underline{u}}_{na}(p)$, and denote the expected profit from the optimal menu of contracts as $\tilde{\pi}(p) = S(\tilde{\underline{u}}_{na}(p), p)$. It can be shown that $S(\underline{u}_{na}, p)$ is a strictly concave function of $\underline{u}_{na}$ and $\tilde{\underline{u}}_{na}(p)$ is well defined over the interval $[0, 1]$. We introduce a value $\hat{p}$ such that

$$\tilde{\underline{u}}_{na}(\hat{p}) = \hat{u}. \tag{7}$$

We can show that such a value $\hat{p}$ exists and is uniquely determined. If the insurer has prior $(\hat{p}, 1-\hat{p})$, then $\underline{u}^*_{na} = \hat{u}$ in the optimal menu of contracts when no ambiguity exists. We have the following lemma (see proof in the appendix).

LEMMA 4. *For an insurer with prior belief P, the optimal menu of contracts exists and $\underline{u}^*_{na} = \tilde{\underline{u}}_{na}(p)$. There is a value p_0 with $0 \le p_0 \le 1$ such* *that the low-type consumers are excluded (with $\tilde{\underline{u}}_{na}(p) = u_{na}$) if $p \leq p_0$ and $\tilde{\underline{u}}_{na}(p)$ is a decreasing function of p for $p \in (p_0, 1]$. Furthermore, the expected profit $\tilde{\pi}(p)$ in the optimal menu of contracts under prior P is the lowest when $p = \hat{p}$.*

Therefore, when there is no ambiguity about the composition of the customers, if the proportion of low-type consumers is low enough such that $p \leq p_0$, all the low-type consumers will be excluded from the market, and the high-type consumers are offered the first-best contract. For $p > p_0$, both types of consumers are served. As the proportion p of the low-type consumers increases, the insurer will set a lower $\underline{u}_{na}$ and increase the coverage of the low-type consumers; the high-type consumers are fully covered, but they pay a lower premium. For different priors of the insurer, the expected profit of the insurer at the optimal is the lowest when the prior is $P = (\hat{p}, 1 - \hat{p})$.

4. THE OPTIMAL MENU OF CONTRACTS WHEN AMBIGUITY EXISTS

We now assume that the insurer has ambiguity about the proportion of two types of customers. From section 2, this ambiguity can be represented by the Choquet expected utility with respect to ε-contamination of a prior P. The insurer assumes that with a probability ε, her subjective belief P is not correct, and the true distribution can be any possible distribution. When ε increases, the degree of the insurer's ambiguity increases. The insurer may be ambiguity averse or ambiguity seeking — she assumes that the worst or best outcome will occur when her belief is not correct.

With ambiguity, the revelation principle still holds (for example, see Vierø, 2012). We only need to consider the menus of contracts in which consumers reveal truthfully their types, as in the previous section. For a menu of contracts $(\underline{u}_a, \underline{u}_{na}, \bar{u}_a, \bar{u}_{na})$, when ambiguity exists, the expected profit of an ambiguity averse monopolistic insurer is

$$\pi = (1 - \varepsilon)[p\pi_L + (1 - p)\pi_H] + \varepsilon \min[\pi_L, \pi_H], \tag{8}$$

and the expected profit of an ambiguity seeking insurer is

$$\pi = (1 - \varepsilon)[p\pi_L + (1 - p)\pi_H] + \varepsilon \max[\pi_L, \pi_H]. \tag{9}$$

The optimal menu of contracts is the one that maximizes the insurer's expected profit. When ambiguity about the composition of the customer exists, we can prove the following:

LEMMA 5. *When there is ambiguity about the proportion of the types of consumers, the results in Lemma 2, Lemma 3 and Proposition 1 are still valid when the insurer is ambiguity averse or ambiguity seeking.*

With information ambiguity, although the expressions of the expected profit in Eq. (8) (for an ambiguity averse insurer) or in Eq. (9) (for an ambiguity seeking insurer) are more complicated, the expected profit will increase if the insurer's profit from one type of consumers increases (with the profit from another type of consumers unchanged). To prove the results in Lemma 2 when there is ambiguity, we only need to show that if a menu of contracts satisfying the incentive and participation constraints does not satisfy properties of (a) to (d), then we can identify another menu satisfying the constraints that strictly increases the profit from one type of consumers (with the profit from another type of consumers unchanged). This is exactly the method used in proving Lemma 2 and Theorem 1 in Chade and Schlee (2012). Once the results in Lemma 2 of our paper hold, the results in Lemma 3 and Proposition 1 follow directly when ambiguity exists, and the value $\hat{u}$ does not depend on the value of ε. As the proof is standard, we dont provide it here.

When the insurer has ambiguity about the composition of the customers, we still refer to the menus in the form of $(s(\underline{u}_{na}), \underline{u}_{na}, z(\underline{u}_{na}), z(\underline{u}_{na}))$ with $\underline{u}_{na} \in [\underline{\theta} u_a + (1-\underline{\theta}) u_{na}, u_{na}]$ as the feasible menus of contracts.

4.1. Optimal menu of contracts for an ambiguity averse insurer

We now assume that the insurer is ambiguity averse. Given that the insurer assumes the worst outcome in the case of ambiguity, we need to determine which outcome is the worst for the insurer.

When the ambiguity averse insurer has ε-contaminated prior, she assumes that an extra ε proportion of the consumers will provide her with the lowest profit. From Lemma 5 (and Lemma 3), the type of customer from which the insurer gets the lowest profit depends on the menu of contracts. For a feasible menu of contracts with $\underline{u}_{na} < \hat{u}$, the profit from a high-type consumer is lower; thus, ε is added to the proportion of high-type consumers $(1-\varepsilon)(1-p)$. Therefore, the insurer assumes that the proportion of high-type consumers is $(1-\varepsilon)(1-p)+\varepsilon = 1-(1-\varepsilon)p$ (and thus the proportion of low-type consumers is $(1-\varepsilon)p$). Similarly, for a feasible menu of contracts with $\underline{u}_{na} > \hat{u}$, the insurer assumes that the proportion of low-type consumers is $(1-\varepsilon)p+\varepsilon$.

We introduce the notations

$$p_1 = (1-\varepsilon)p, p_2 = (1-\varepsilon)p + \varepsilon. \tag{10}$$

Compared with prior $P = (p, 1-p)$, prior belief $P_1 = (p_1, 1-p_1)$ features a lower proportion of low-type consumers, and prior belief $P_2 = (p_2, 1-p_2)$ features a higher proportion of low-type consumers. Thus, we get

Lemma 6. *The expected profit of an ambiguity averse insurer with ε-contamination of prior $P = (p, 1-p)$ can be written as:*

$$\pi(\underline{u}_{na}) = \begin{cases} S(\underline{u}_{na}, p_1), & \text{if } \underline{u}_{na} \leq \hat{u} \\ S(\underline{u}_{na}, p_2), & \text{if } \underline{u}_{na} \geq \hat{u} \end{cases} \tag{11}$$

Without ambiguity about the composition of the customers, the insurer applies prior belief P for any feasible menu of contracts. With information ambiguity, for all feasible menus of contracts with $\underline{u}_{na} \leq \hat{u}$, the ambiguity averse insurer applies prior belief $P_1 = (p_1, 1-p_1)$ to calculate her expected profit; she applies prior belief $P_2 = (p_2, 1-p_2)$ for all feasible menus of contracts with $\underline{u}_{na} \geq \hat{u}$. There is a switch of belief at the menu of contracts with $\underline{u}_{na} = \hat{u}$.

Because of the switch of belief at $\underline{u}_{na} = \hat{u}$, the insurer's expected profit function $\pi(\underline{u}_{na})$ is now a conjunction of two strictly concave functions $S(\underline{u}_{na}, p_1)$ and $S(\underline{u}_{na}, p_2)$, which intersect at $\underline{u}_{na} = \hat{u}$. Function $\pi(\underline{u}_{na})$ is continuous with respect to $\underline{u}_{na}$, but it has a kink at $\underline{u}_{na} = \hat{u}$. By Lemma 4, functions $S(\underline{u}_{na}, p_1)$ and $S(\underline{u}_{na}, p_2)$ achieve their maximum at $\tilde{\underline{u}}_{na}(p_1)$ and $\tilde{\underline{u}}_{na}(p_2)$, respectively, with $\tilde{\underline{u}}_{na}(p_1) \geq \tilde{\underline{u}}_{na}(p_2)$. We can now identify the maximum of $\pi(\underline{u}_{na})$ from the properties of $S(\underline{u}_{na}, p_1)$ and $S(\underline{u}_{na}, p_2)$.

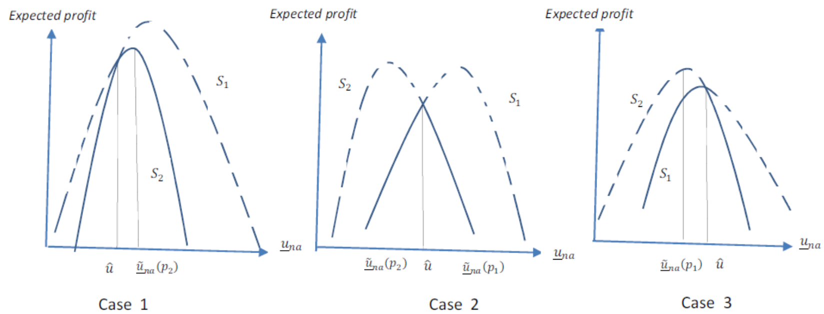
Lemma 7. *For an ambiguity averse insurer with ε-contamination of prior belief $P = (p, 1-p)$ about the composition of the customers, the optimum menu of contract is characterized by $\underline{u}^*_{na} = \tilde{\underline{u}}_{na}(p_2)$ if $p_2 < \hat{p}$; $\underline{u}^*_{na} = \tilde{\underline{u}}_{na}(p_1)$ if $p_1 > \hat{p}$; and $\underline{u}^*_{na} = \hat{u}$ if $p_1 \leq \hat{p} \leq p_2$.*

Note that $\hat{p}$ is defined in Eq. (7) as the value such that $\tilde{\underline{u}}_{na}(\hat{p}) = \hat{u}$, and p_1 and p_2 are defined in Eq. (10). Since p_1 and p_2 are functions of ε, with $p_2 - p_1 = \varepsilon$, a higher ε increases the likelihood of the case $p_1 \leq \hat{p} \leq p_2$. Using Lemma 7, we can characterize the optimal menu of contracts by characterizing $\underline{u}^*_{na}$ (see the proof in the appendix).

Proposition 2. *For an ambiguity averse monopolistic insurer with ε-contamination of prior belief $P = (p, 1-p)$ about the composition of the customers, we have*

if $p < \hat{p}$, then $\underline{u}^*_{na} = \max(\tilde{\underline{u}}_{na}(p_2), \hat{u})$;
if $p = \hat{p}$, then for any $\varepsilon \geq 0, \underline{u}^*_{na} = \hat{u}$;
if $p > \hat{p}$, then $\underline{u}^*_{na} = \min(\tilde{\underline{u}}_{na}(p_1), \hat{u})$.

FIG. 1. Expected profit when ambiguity exists



We can provide an intuitive illustration of Proposition 2 using Figure 1. We have three cases in Figure 1. In each case, the x-axis is the value of $\underline{u}_{na}$, and the y-axis is the value of the insurer's expected profit $\pi(\underline{u}_{na})$. Curve S_1 corresponds to function $S(\underline{u}_{na}, p_1)$ (as a function of $\underline{u}_{na}$), and curve S_2 corresponds to function $S(\underline{u}_{na}, p_2)$ (as a function of $\underline{u}_{na}$). When no ambiguity exists, $p_1 = p_2$ and the two curves S_1 and S_2 are the same. When ambiguity exists, curve S_2 shift to the left and S_1 shift to the right, and the two curves intersect at $\underline{u}_{na} = \hat{u}$. The expected profit $\pi(\underline{u}_{na})$ of the ambiguity averse insurer is the solid line of the two curves.

If $p < \hat{p}$ and the proportion of low consumer is low, then $\tilde{\underline{u}}_{na}(p) > \hat{u}$ when there is no ambiguity and the coverage of a low-type consumer is low. When ambiguity exists but ambiguity is low such that $p_2 < \hat{p}$, we get $\tilde{\underline{u}}_{na}(p_1) > \tilde{\underline{u}}_{na}(p_2) > \hat{u}$ (case 1 in Figure 1). For the feasible menus of contracts with $\underline{u}_{na} \leq \hat{u}$, the insurer applies prior P_1 and her expected profit $S(\underline{u}_{na}, p_1)$ achieves its maximum at $\underline{u}_{na} = \hat{u}$ for all $\underline{u}_{na} \leq \hat{u}$. For the feasible menus of contracts with $\underline{u}_{na} \geq \hat{u}$, the insurer applies prior P_2 and her expected profit $S(\underline{u}_{na}, p_2)$ reaches its maximum at $\tilde{\underline{u}}_{na}(p_2)$. The overall maximum of $\pi(\underline{u}_{na})$ is achieved at $\tilde{\underline{u}}_{na}(p_2)$, and the insurer assumes p_2 proportion of low valuation consumers.

However, for $p < \hat{p}$, when ambiguity is high enough (the difference between p_1 and p_2 is big since $p_2 - p_1 = \varepsilon$), we may get $p_1 \leq \hat{p} \leq p_2$ and thus $\tilde{\underline{u}}_{na}(p_2) \leq \hat{u} \leq \tilde{\underline{u}}_{na}(p_1)$ (case 2 in Figure 1). For the feasible menus of contracts with $\underline{u}_{na} \leq \hat{u}$, the insurer assumes p_1 proportion of low-type consumers; her expected profit achieves its maximum at $\underline{u}_{na} = \hat{u}$ for all $\underline{u}_{na} \leq \hat{u}$. For the feasible menus of contracts with $\underline{u}_{na} \geq \hat{u}$, the insurer assumes p_2 proportion of low-type consumers. Her expected profit $S(\underline{u}_{na}, p_2)$ also achieves its maximum at $\underline{u}_{na} = \hat{u}$ for all $\underline{u}_{na} \geq \hat{u}$. In the overall optimum, $\underline{u}_{na} = \hat{u}$. Therefore, if $p < \hat{p}$, the monopolist sets $\underline{u}^*_{na} = \tilde{\underline{u}}_{na}(p_2)$ (with $\tilde{\underline{u}}_{na}(p_2) > \hat{u}$) and uses prior P_1 when the degree of ambiguity is low, and she sets $\underline{u}^*_{na} = \hat{u}$ (with $\tilde{\underline{u}})_{na}(p_2) < \hat{u}$) when the degree of ambiguity is high. We have $\underline{u}^*_{na} = \max(\tilde{\underline{u}}_{na}(p_2), \hat{u})$ for $p < \hat{p}$.

Similarly, if $p > \hat{p}$, case 3 in Figure 1 will happen when the degree of ambiguity is low, and case 2 will happen when the degree of ambiguity is high, and $\underline{u}^*_{na} = \min(\tilde{\underline{u}}_{na}(p_1), \hat{u})$. If $p = \hat{p}$, we always have case 2 and $\underline{u}^*_{na} = \hat{u}$.

When ambiguity exists, we obtain another outcome which does not occur without ambiguity:

Proposition 3. *For an ambiguity averse insurer and for a given $\varepsilon > 0$, the optimal menus of contracts for ε contamination of all the beliefs $P = (p, 1-p)$ with $p \in \left(\frac{p-\varepsilon}{1-\varepsilon}, \frac{\hat{p}}{1-\varepsilon}\right)$ are the same menu with $\underline{u}^*_{na} = \hat{u}$.*[3]

When the prior belief of the ambiguity averse insurer changes, the optimal menu of contracts may be the same for a range of priors even when both types of consumers are served.

For a given prior belief $P = (p, 1-p)$, we denote $\underline{u}^*_{na}$ as $\underline{u}^*_{na}(\varepsilon)$ when the degree of ambiguity ε varies. We can study the comparative statics of the optimal menu of contracts when the degree of ambiguity changes. Actually, Lemma 7 includes the case wherein the low-type consumers are excluded. For $p < p_0$, if ε is small such that $p_2 \leq p_0$ (and thus $p_2 < \hat{p}$, as $\hat{p} > p_0$), we obtain $\underline{u}^*_{na} = \tilde{\underline{u}}_{na}(p_2) = u_{na}$, and the low-type consumers are excluded. In such a case, the optimal menus of contracts with or without ambiguity are the same. If the degree of ambiguity increases (i.e, as ε increases) such that $p_2 \in (p_0, \hat{p})$, then $\underline{u}^*_{na} = \tilde{\underline{u}}_{na}(p_2) > u_n a$ and the low-type consumers will get positive coverage. When ε continues to rise such that $p_2 \geq \hat{p}$, we have $\underline{u}^*_{na} = \hat{u}$. The succeeding result describes the comparative statics when ambiguity increases.

Proposition 4. *As ambiguity increases, $\underline{u}^*_{na}$ for an ambiguity averse insurer moves toward $\hat{u}$. Once $\underline{u}^*_{na}$ reaches $\hat{u}$, it remains in this position as ε increases.*

Therefore, an "attraction" menu of contracts characterized by $\underline{u}^*_{na} = \hat{u}$ appears when the degree of ambiguity increases. If $p > \hat{p}$, then we obtain $\underline{u}^*_{na} < \hat{u}$ when there is no ambiguity. When ambiguity increases, $\underline{u}^*_{na}$ increases and the coverage of a low-type consumer decreases until $\underline{u}^*_{na}$ reaches the level $\hat{u}$. If $p < \hat{p}$, then $\underline{u}^*_{na} > \hat{u}$ when there is no ambiguity. When ambiguity increases, $\underline{u}^*_{na}$ decreases and the ambiguity averse insurer will increase the coverage of a low-type consumer until the coverage reaches the level determined by $\underline{u}^*_{na} = \hat{u}$.

[3] If $\hat{p} - \varepsilon < 0$, then the interval becomes $\left(0, \frac{\hat{p}}{1-\varepsilon}\right)$.

In the attraction contract, the profits of the ambiguity averse insurer from both types of consumers are equal. When the degree of ambiguity is high, an ambiguity averse insurer becomes very conservative: the insurer uses a menu of contracts that equalizes the profits earned from the two types of consumers and disregards the information included in her prior about the distribution of the consumers.

4.2. Optimal menu of contracts for an ambiguity seeking insurer

As there is empirical evidence that some decision-makers are ambiguity seeking, it is interesting to explore the optimal menu of contracts for an ambiguity seeking insurer. We can use the same method used in the discussion for an ambiguity averse insurer. The main difference is that the insurer assumes that the best outcome will occur when ambiguity exists, and thus $\pi = (1-\varepsilon)[p\pi_L + (1-p)\pi_H] + \varepsilon \max[\pi_L, \pi_H]$.

We know that the results in Lemma 5 are valid for both an ambiguity averse and an ambiguity seeking insurer. For a feasible menu of contracts with $\underline{u}_{na} < \hat{u}$, the profit from a low-type consumer is higher; thus, ε is added to the proportion of low-type consumers. Therefore, the insurer assumes that the proportion of low-type consumers is $p_2 = (1-\varepsilon)p + \varepsilon$. Similarly, for a feasible menu of contracts with $\underline{u}_{na} > \hat{u}$, the insurer assumes that the proportion of low-type consumers is $p_1 = (1-\varepsilon)p$. The expected profit of the ambiguity seeking insurer becomes:

$$\pi(\underline{u}_{na}) = \begin{cases} S(\underline{u}_{na}, p_2), & \text{if } \underline{u}_{na} \leq \hat{u} \\ S(\underline{u}_{na}, p_1), & \text{if } \underline{u}_{na} \geq \hat{u} \end{cases}$$

The optimal menu of contracts can be obtained in a similar way as for an ambiguity averse insurer. The easiest way is to look at the three cases in Figure 1. In each case, the expected profit of an ambiguity seeking insurer is now the dotted line of the two curves S_1 and S_2. In case 1 where the proportion p of low-type consumers is low and $p_2 < \hat{p}$ (the degree of ambiguity is low), the insurer sets $\underline{u}^*_{na} = \tilde{\underline{u}}_{na}(p_1)$, and assumes an even lower proportion p_1 of low-type consumers in the optimum menu of contracts. In case 3 where the proportion p of low-type consumers is high and $p_1 > \hat{p}$, the insurer sets $\underline{u}^*_{na} = \tilde{\underline{u}}_{na}(p_2)$, and assumes an even higher proportion p_2 of low-type consumers in the optimal menu of contracts. In case 2 where $p_1 le\hat{p} \leq p_2$, the insurer will set either $\underline{u}^*_{na} = \tilde{\underline{u}}_{na}(p_1)$ or $\underline{u}^*_{na} = \tilde{\underline{u}}_{na}(p_2)$, depending on which of $S(\tilde{\underline{u}}_{na}(p_2), p_2)$ or $S(\tilde{\underline{u}}_{na}(p_1), p_1)$ is higher[4].

[4] In this case, when ambiguity increases, it is theoretically possible that $\underline{u}^*_{na}$ in the optimal menu of contracts jumps from $S(\tilde{\underline{u}}_{na}(p_2), p_2)$ to $S(\tilde{\underline{u}}_{na}(p_1), p_1)$, or vice-versa.

For a given prior belief $P = (p, 1-p)$, let $\underline{u}_{na}^*(\varepsilon)$ be the value of $\underline{u}_{na}^*$ for an ambiguity seeking insurer when the degree of ambiguity is ε. For $p = \hat{p}$, we get $\underline{u}_{na}^* = \hat{u}$ when there is no ambiguity. As ambiguity increases, $\underline{u}_{na}^*(\varepsilon)$ can be either $\underline{\tilde{u}}_{na}(p_2)$ or $\underline{\tilde{u}}_{na}(p_1)$. For $p < \hat{p}$, we get $\underline{u}_{na}^* > \hat{u}$ when there is no ambiguity. As ambiguity increases (but $p_2 < \hat{p}$), we get $\underline{u}_{na}^*(\varepsilon) = \underline{\tilde{u}}_{na}(p_1)$, which further increases with ε.[5] The coverage of the low-type consumers will further decrease as ambiguity increase. When ε continue to increase such that $p_2 \geq \hat{p}$ (corresponding to the case $p_1 \leq \hat{p} \leq p_2$), the insurer will set either $\underline{u}_{na}^* = \underline{\tilde{u}}_{na}(p_1)$ or $\underline{u}_{na}^* = \underline{\tilde{u}}_{na}(p_2)$.

Similarly, for $p > \hat{p}$, starting from the case without ambiguity in which $\underline{u}_{na}^*$ is low (or the coverage of the low-type consumer is high), as ambiguity increases (but $p_1 > \hat{p}$), then $\underline{u}_{na}^*(\varepsilon) = \underline{\tilde{u}}_{na}(p_1)$, which decreases with ε, and the coverage of the low-type consumers will further increase. When ambiguity is very high such that $p_1 \leq \hat{p} \leq p_2$, the insurer will set either $\underline{u}_{na}^* = \underline{\tilde{u}}_{na}(p_1)$ or $\underline{u}_{na}^* = \underline{\tilde{u}}_{na}(p_2)$.

We can check that for a given $\varepsilon > 0$, we have either $\underline{u}_{na}^* \geq \underline{u}_{na}^*(\hat{p}(1-\varepsilon))$ or $\underline{u}_{na}^* \leq \underline{u}_{na}^*(\hat{p}(1-\varepsilon)+\varepsilon)$ for any prior P.[6] As ambiguity increases, $\underline{u}_{na}^*$ for an ambiguity seeking insurer moves away from $\hat{u}$. Therefore, the menu of contract with $\underline{u}_{na}^* = \hat{u}$ looks like a repulsion center.

For an ambiguity averse insurer, the menu of contracts characterized by $\underline{u}_{na}^* = \hat{u}$ is the attraction center, but it becomes an expulsion center for an ambiguity seeking insurer. The reason for this to happen is that the insurer adjusts her belief about the composition of customers in the opposite ways. An ambiguous averse insurer tends to adjust her belief in a pessimist way and an ambiguity seeking insurer tends to adjust her belief in an optimist way. As we know from Lemma 4 that the expected profit $\tilde{\pi}(p)$ achieves its minimum at $p = \hat{p}$, an ambiguity averse insurer tends to adjust her prior in favor of $p = \hat{p}$ (corresponding to $\underline{u}_{na}^* = \hat{u}$), but an ambiguity seeking insurer tends to adjust her belief against the prior with $p = \hat{p}$ (the insurer uses the extreme prior p_1 or p_2, and $\underline{u}_{na}^*$ takes the value of $\underline{\tilde{u}}_{na}(p_2)$ or $\underline{\tilde{u}}_{na}(p_1)$).

4.3. An example

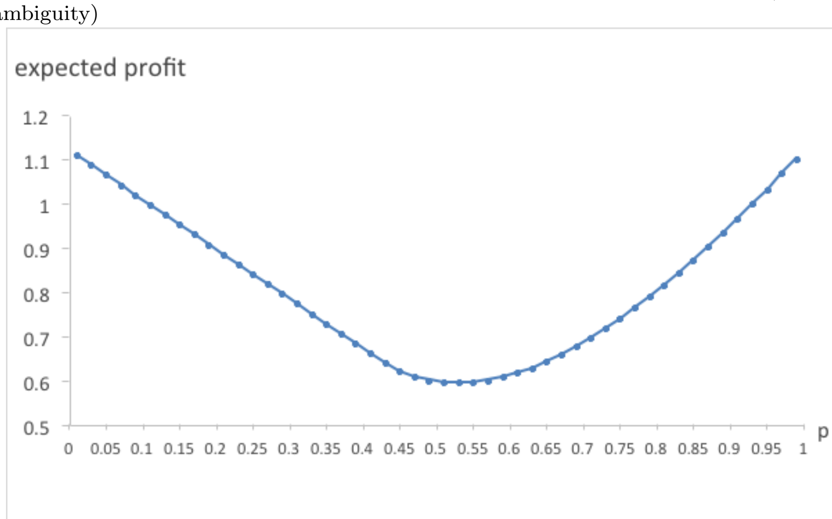
We provide a numerical example for the optimal menu of contracts. Suppose $u(x) = \sqrt{x}$, $w = 10$, $l = 9$, $\underline{\theta} = 0.4$, $\bar{\theta} = 0.6$. Then we get $u_a = 1$, $u_{na} = 3.16$, $p_0 = 0.418$, $\hat{p} = 0.530$ with $\hat{u} = 2.890$. When $p < 0.418$, low-type consumers will be excluded in the market. When $p = \hat{p} = 0.530$ and there is no ambiguity, the insurer's profit in the optimum from the two types of consumers are equal, and $\underline{u}_{na}^*$ is at the level $\hat{u} = 2.890$, which

[5] If the low-type consumers are excluded without ambiguity, they will also be excluded when ambiguity exists.

[6] For $p = \hat{p}$, we get $\underline{u}_{na}^* = \hat{u}$ when there is no ambiguity; for $\varepsilon > 0$, $\underline{u}_{na}^*(\varepsilon)$ is either $\underline{u}_{na}^*(\hat{p}(1-\varepsilon))$ or $\underline{u}_{na}^*(\hat{p}(1-\varepsilon)+\varepsilon)$. For $p < \hat{p}$ and $p > \hat{p}$, we can verify that either $\underline{u}_{na}^* \geq \underline{u}_{na}^*(\hat{p}(1-\varepsilon))$ or $\underline{u}_{na}^* \leq \underline{u}_{na}^*(\hat{p}(1-\varepsilon)+\varepsilon)$.

corresponds to $\underline{t} = 2.132$, $\underline{x} = 3.488$, $\bar{t} = 5.826$ (with $\bar{x} = 9$). In this case, the insurer's expected profit is $\pi = 0.596$. Figure 2 shows the expected profit of the insurer (as a function of the proportion p of the low valuation consumers) under the optimal menu of contracts when there is no ambiguity.

FIG. 2. Expected profit of the insurer in the optimum as a function of p (without ambiguity)



For an ambiguity averse insurer, if $p = 0.530$, the insurer will use the menu of contract with $\underline{u}^*_{na} = \hat{u} = 2.890$ for any $\varepsilon > 0$. If $p > 0.530$, for example, $p = 0.6$, $\underline{u}^*_{na} = 2.757$ for $\varepsilon = 0$, $\underline{u}^*_{na} = \underline{\tilde{u}}_{na}(p_1) = 2.811$ for $\varepsilon = 0.05$, and $\underline{u}^*_{na} = \hat{u} = 2.890$ for $\varepsilon = 0.15$. The value of $\underline{u}^*_{na}$ increases and moves towards $\hat{u}$, and the coverage the low-type consumer decreases when ambiguity increases.

For an ambiguity seeking insurer with prior P and $\varepsilon > 0$, the insurer compares the value of $\tilde{\pi}(p_1)$ and $\tilde{\pi}(p_2)$, where $\tilde{\pi}(p) = S(\underline{\tilde{u}}_{na}(p), p)$ is the expected profit from the optimal menu of contract with prior $P = (p, 1-p)$ and $\varepsilon = 0$. The expected profit $\tilde{\pi}(p)$ of the insurer is shown in Figure 2. For $p = \hat{p} = 0.530$, we already know that $\underline{u}^*_{na} = \hat{u} = 2.890$ without ambiguity. For $p = \hat{p} = 0.530$ and with ambiguity, we can check that for all $0 < \varepsilon < 1$, we always have $\tilde{\pi}(p_1) > \tilde{\pi}(p_2)$ in this example, and thus $\underline{u}^*_{na} = \underline{\tilde{u}}_{na}(p_1)$. When ambiguity increases, the value of $\underline{u}^*_{na}$ increases and moves away from $\hat{u}$ until $\underline{u}^*_{na} = u_{na}$ where low valuation buyers are excluded from the market.

If $p > \hat{p}$ (so that the proportion of low-type consumer is high), for example, $p = 0.6$, we can check that $\tilde{\pi}(p_2) > \tilde{\pi}(p_1)$ for $\varepsilon = 0.05, 0.1, 0.5, 0.7$, the insurer applies prior P_2 with a higher proportion of low-type consumers. When ambiguity increases, the value of $\underline{u}^*_{na}$ decreases and moves away from

$\hat{u}$, and the coverage of the low-type consumers increases. If $p < \hat{p}$, for example, $p = 0.48$, we can check that $\tilde{\pi}(p_1) > \tilde{\pi}(p_2)$ for $\varepsilon = 0.05, 0.1, 0.5, 0.7$, the insurer assumes prior P_1 with an even lower proportion of low-type consumers. When ambiguity increases, the value of $\underline{u}^*_{na}$ increases (until $\underline{u}^*_{na} = u_{na}$ where low-type consumers are excluded from the market) and moves away from $\hat{u}$, and the coverage of the low-type consumers decreases.

5. CONCLUSION

We have explored the effect of an insurers ambiguity about the distribution of consumers on the optimal menu of contracts. We find that in the optimum, the high-type consumers get full coverage and the low-type consumers get less than full coverage, and the optimal contract is a separating contract. When ambiguity exists, an insurer adjusts her prior belief for different menus of contracts. The existence of ambiguity can increase or decrease the coverage of low-type consumers. As ambiguity increases, the optimal menu of contracts for an ambiguity averse insurer moves towards the menu that equalize the profits from the two types of consumers, while the optimal menu of contracts for an ambiguity seeking insurer moves away from the menu that equalize the profits from the two types of consumers.

For an insurer entering a new market or facing varying population so that there exists ambiguity about the composition of the consumers, if there is no opportunity of renegotiation in the short run (for long term contract, for example), the insurer has no opportunity to learn from the consumers' choice when designing the contract. Our results suggests that the choice of the insurer depends on the insurer's prior belief, the degree of ambiguity, and the insurer's attitude towards ambiguity. It may turn out that the composition of the consumers after observing the choices of the consumers is very different from the insurers prior (or the ambiguity-adjusted prior). However, the insurer cannot do better ex-ante.

Once the insurer has opportunity to learn, for example, the population is fixed and the contract is short-term or renegotiable, our result of separating contract suggests that the insurer can learn perfectly the composition of consumers by observing the consumers' choices, even when ambiguity exists. Ambiguity can be solved and incorrect prior can be corrected, and the insurer can use the correct information about the composition of consumers in the renegotiation of contracts.

The above results are restricted for two types of consumers, which may not be true for more than two types of consumers. When there are more than two types of consumers, we can use similar ideas as in the case of two types of consumers. However, Choquet integral for more than two states cannot always be written in a form using ε contaminated prior. For more than two states, there is a hot theoretical debate about the description of ambiguity and ambiguity aversion (or ambiguity seeking) in the Choquet integral representation. There is even no consensus about whether we can separate ambiguity from ambiguity attitude. For more than two types of consumers, even with ε-contaminated prior, the rank of profits from consumers for a menu of contract can be more complicated, so are the incentive constraints and participation constraints.

In actual insurance market, we need to consider more than two types of consumers. It will be interesting to explore the choice of the insurer when the insurer has no opportunity to learn the composition of the consumers, and to which degree the insurer can learn the composition of the consumers if the insurer has opportunity to do so. Our paper with two types of consumers provides a possible starting point for further studies.

APPENDIX: PROOFS

Proof of Lemma 1. A capacity ν over $S = H, L$ is determined by values of $\nu(H)$ and $\nu(L)$, with $0 \leq \nu(H), \nu(L) \leq 1$. (For a capacity, we have $\nu(\emptyset) = 0$ and $\nu(S) = 1$). It is known (see Gilboa and Schmeidler, 1994) that if ν is convex with $\nu(H) + \nu(L) < 1$, Choquet integral $\int u d\nu$ can be written as

$$\int u d\nu = u(H)v(H) + u(L)v(L) + (1 - \nu(H) - \nu(L)) \min(U(H), u(L)).$$

Thus, the Choquet integral can be writeen as $(1-\varepsilon)E^P u + \varepsilon \min_{s \in S} u(s)$, with $P(H) = \frac{\nu(H)}{\nu(H)+\nu(L)}, P(L) = \frac{\nu(L)}{\nu(H)+\nu(L)}$ and $\varepsilon = 1 - \nu(H) - \nu(L)$. If ν is concave with $\nu(H) + \nu(L) > 1$, then

$$\int u d\nu = u(H)v(H) + u(L)v(L) + (1 - \nu(H) - \nu(L)) \min(U(H), u(L)),$$

and we have $(1 - \nu(H) - \nu(L)) < 0$. Using $\min(U(H), u(L)) = U(H) + u(L) - \max(U(H), u(L))$, we can write

$$\int u d\nu = (1-v(L))u(H) + (1-v(H))u(L) + (\nu(H)+\nu(L)-1) \max(U(H), u(L)).$$

Thus, the Choquet integral can be written as $(1-\varepsilon)E^P u + \varepsilon \max_{s \in S} u(s)$, with $P(H) = \frac{1-\nu(L)}{1-\nu(H)+1-\nu(L)}, P(L) = \frac{1-\nu(H)}{1-\nu(H)+1-\nu(L)}$ and $\varepsilon = \nu(H)+\nu(L)-1$. ∎

Proof of Proposition 1. From Lemma 2, the optimal menu of contracts $(\underline{u}_a^*, \underline{u}_{na}^*, \bar{u}_a^*, \bar{u}_{na}^*)$ has the form $(s(\underline{u}_{na}), \underline{u}_{na}, z(\underline{u}_{na}), z(\underline{u}_{na}))$. We can determine the range of possible values that $\underline{u}_{na}^*$ may take. From $\underline{u}_{na}^* \geq \underline{u}_a^*$

and Eq. (4), we obtain $\underline{u}^*_{na} \geq u_a + \frac{1-\theta}{\theta}u_{na} - \frac{1-\theta}{\theta}\underline{u}^*_{na}$, which implies $\underline{u}^*_{na} \geq \underline{\theta}u_a + (1-\underline{\theta})u_{na}$. Therefore, $\underline{u}^*_{na} \in [\underline{\theta}u_a + (1-\underline{\theta})u_{na}, u_{na}]$. ▮

Proof of Lemma 3. Since $s'(\underline{u}_{na}) = -\frac{1-\theta}{\underline{\theta}}$, and $z'(\underline{u}_{na}) = -\frac{\bar{\theta}-\underline{\theta}}{\underline{\theta}}$, it is easy to see that $\pi'_L(\underline{u}_{na}) = -\underline{\theta}v'(s(\underline{u}_{na}))\left(-\frac{1-\theta}{\underline{\theta}}\right) - (1-\underline{\theta})v'(\underline{u}_{na}) = (1-\underline{\theta})[v'(s(\underline{u}_{na})) - v'(\underline{u}_{na})]$ and $\pi'_H(\underline{u}_{na}) = v'(z(\underline{u}_{na}))\frac{\bar{\theta}-\underline{\theta}}{\underline{\theta}}$. As v is strictly increasing and strictly convex, and $s(\underline{u}_{na}) \leq \underline{u}_{na}$, we obtain $v'(z(\underline{u}_{na})) > 0$, $v'(s(\underline{u}_{na})) - v'(\underline{u}_{na}) \leq 0$. Thus, $\pi_L(\underline{u}_{na})$ is a decreasing function and $\pi_H(\underline{u}_{na})$ is a strictly increasing function.

We now compare the value of $\pi_L(\underline{u}_{na})$ and $\pi_H(\underline{u}_{na})$ at the end values of interval $[\underline{\theta}u_a + (1-\underline{\theta})u_{na}, u_{na}]$. We will prove that $\pi_H(u_a) < \pi_L(u_a)$ at $\underline{u}_{na} = \underline{\theta}u_a + (1-\underline{\theta})u_{na}$ and $\pi_L(\underline{u}_{na}) < \pi_H(\underline{u}_{na})$ at $\underline{u}_{na} = u_{na}$.

At $\underline{u}_{na} = u_{na}$ (i.e., $\underline{t} = 0$), we know that $\underline{u}_a = u_a$ and $\bar{u}_a = \bar{u}_{na} = \bar{\theta}u_a + (1-\bar{\theta})u_{na}$, and the low-type consumers are excluded from the market. Thus, $\pi_L(u_{na}) = -\underline{\theta}v(u_a) - (1-\underline{\theta})v(u_{na}) + w - \underline{\theta}l = 0$, and $\pi_H(u_{na}) = -v(\bar{\theta}u_a + (1-\bar{\theta})u_{na}) + w - \bar{\theta}l$. Since u is strictly concave, we have $u(\bar{\theta}(w-l) + (1-\bar{\theta})w) > \bar{\theta}u(w-l) + (1-\bar{\theta})u(w) = \bar{\theta}u_a + (1-\bar{\theta})u_{na}$, and thus $v(\bar{\theta}u_a + (1-\bar{\theta})u_{na}) < \bar{\theta}(w-l) + (1-\bar{\theta})w = w - \bar{\theta}l$, which implies $\pi_H(u_{na}) = -v(\bar{\theta}u_a + (1-\bar{\theta})u_{na}) + w - \bar{\theta}l > 0$. Therefore, $\pi_L(\underline{u}_{na}) < \pi_H(\underline{u}_{na})$ at $\underline{u}_{na} = u_{na}$.

Since $\underline{u}_{na} = \underline{\theta}u_a + (1-\underline{\theta})u_{na}$, we know that $\underline{u}_a = \bar{u}_a = \bar{u}_{na} = \underline{u}_{na} = \underline{\theta}u_a + (1-\underline{\theta})u_{na}$. We have $\pi_L(\underline{u}_{na}) = -\underline{\theta}v(\underline{u}_{na}) - (1-\underline{\theta})v(\underline{u}_{na}) + w - \underline{\theta}l = -v(\underline{u}_{na}) + w - \underline{\theta}l$, and $\pi_H(u_a) = -v(\underline{u}_{na}) + w - \bar{\theta}l$. Now it is easy to see that $\pi_H(u_a) < \pi_L(u_a)$ at $\underline{u}_{na} = \underline{\theta}u_a + (1-\underline{\theta})u_{na}$. Since $\pi_L(\underline{u}_{na})$ is decreasing and $\pi_H(\underline{u}_{na})$ is increasing, and $\pi_H(u_a) < \pi_L(u_a)$ at $\underline{u}_{na} = \underline{\theta}u_a + (1-\underline{\theta})u_{na}$ and $\pi_L(\underline{u}_{na}) < \pi_H(\underline{u}_{na})$ at $\underline{u}_{na} = u_{na}$, a unique value $\hat{u}$ exists with $\underline{\theta}u_a + (1-\underline{\theta})u_{na} < \hat{u} < u_{na}$, such that $\pi_L(\hat{u}) = \pi_H(\bar{u})$, and $\pi_H(\underline{u}_{na}) < \pi_L(\underline{u}_{na})$ if $\underline{u}_{na} < \hat{u}$, $\pi_L(\underline{u}_{na}) < \pi_H(\underline{u}_{na})$ if $\underline{u}_{na} > \hat{u}$. ▮

Proof of Lemma 4. The proof is standard, and we only provide a sketch of the proof. We prove the lemma by using the concavity of $S(\underline{u}_{na}, p)$ and the monotonicity of $S(\underline{u}_{na}, p)$ (both as a function of $\underline{u}_{na}$) at the two ends of interval $[\underline{\theta}u_a + (1-\underline{\theta})u_{na}, u_{na}]$.

Take partial derivative with respect to $\underline{u}_{na}$ (see also the proof of Lemma 3), we get

$$\frac{\partial S(\underline{u}_{na}, p)}{\partial \underline{u}_{na}} = p(1-\underline{\theta})[v'(s(\underline{u}_{na})) - v'(\underline{u}_{na})] + (1-p)v'(z(\underline{u}_{na}))\frac{\bar{\theta}-\underline{\theta}}{\underline{\theta}},$$

and $\frac{\partial^2 S(\underline{u}_{na},p)}{\partial \underline{u}^2_{na}} < 0$ from $s'(\underline{u}_{na}) < 0, z'(\underline{u}_{na}) < 0$, and $v'' > 0$. Thus, $S(\underline{u}_{na}, p)$ is a strictly concave function of $\underline{u}_{na}$.

At the left end of the interval $\underline{u}_{na} = \underline{\theta}u_a + (1-\underline{\theta})u_{na}$, as $s(\underline{u}_{na}) = \underline{u}_{na}$, we get $\frac{\partial S(\underline{u}_{na},p)}{\partial \underline{u}_{na}} > 0$ for $p < 1$ and $S(\underline{u}_{na}, p)$ is increasing. (For $p = 1$,

$\frac{\partial S(\underline{u}_{na},p)}{\partial \underline{u}_{na}} = 0$ at $\underline{u}_{na} = \underline{\theta} u_a + (1-\underline{\theta}) u_{na})$. At the right end of the interval $\underline{u}_{na} = u_{na}$, we get $\frac{\partial S(\underline{u}_{na},p)}{\partial \underline{u}_{na}} = A - p[A + (1-\underline{\theta})[v'(u_{na}) - v'(u_a)]]$, where $A = v'(\bar{\theta} u_a + (1-\bar{\theta}) u_{na}) \frac{\bar{\theta}-\underline{\theta}}{\underline{\theta}} > 0$.

We denote $p_0 = \frac{A}{(1-\underline{\theta})[v'(u_{na}) - v'(u_a)] + A}$. Since v is strictly convex, we have $v'(u_{na}) - v'(u_a) > 0$, and it is easy to see that $0 < p_0 < 1$. Therefore, if $p \le p_0$, then $\frac{\partial S(\underline{u}_{na},p)}{\partial \underline{u}_{na}} \ge 0$ at $\underline{u}_{na} = u_{na}$. If $p > p_0$, then $\frac{\partial S(\underline{u}_{na},p)}{\partial \underline{u}_{na}} < 0$ at $\underline{u}_{na} = u_{na}$.

$S(\underline{u}_{na}, p)$ is a strictly concave function of $\underline{u}_{na}$ over the interval $[\underline{\theta} u_a + (1-\underline{\theta}) u_{na}, u_{na}]$ and it is non-decreasing at the left end of the interval. For $p \le p_0$, it is non-decreasing at the right end of the interval; therefore, $S(\underline{u}_{na}, p)$ is an increasing function of $\underline{u}_{na}$ over the interval and the maximum of $S(\underline{u}_{na}, p)$ will be reached at $\underline{u}_{na} = u_{na}$. For $p > p_0$, $S(\underline{u}_{na}, p)$ is decreasing at the right end of the interval, the maximum of $S(\underline{u}_{na}, p)$ will be reached at the interior point $\underline{u}_{na}$ uniquely determined by the first order condition $\frac{\partial S(\underline{u}_{na},p)}{\partial \underline{u}_{na}} = 0$.

If we denote $\underline{u}_{na}$ that maximizes $S(\underline{u}_{na}, p)$ for a value of $p \in [0,1]$ as $\tilde{\underline{u}}_{na}(p)$, then $\tilde{\underline{u}}_{na}(p)$ is well defined: it is equal to u_{na} for $p \le p_0$, and it is equal to the unique interior value of $\underline{u}_{na}$ determined by the condition $\frac{\partial S(\underline{u}_{na},p)}{\partial \underline{u}_{na}} = 0$ for $p > p_0$.

For $p > p_0$, function $\tilde{\underline{u}}_{na}(p)$ satisfies the first order condition:

$$p(1-\underline{\theta})[v'(s(\tilde{\underline{u}}_{na}(p))) - v'(\tilde{\underline{u}}_{na}(p))] + (1-p) v'(z(\tilde{\underline{u}}_{na}(p))) \frac{\bar{\theta}-\underline{\theta}}{\underline{\theta}} = 0.$$

Taking derivative with respect to p, and use the conditions $v'' > 0, s'(u) < 0$ and $z'(u) < 0$, we get $\frac{d\tilde{\underline{u}}_{na}(p)}{dp} < 0$ for $p > p_0$.

Since $\hat{p}$ satisfies $\tilde{\underline{u}}_{na}(\hat{p}) = \hat{u}$, given that $u_a < \hat{u} < u_{na}$ and $\frac{d\tilde{\underline{u}}_{na}(p)}{dp} < 0$ for $p > p_0$, such a value $\hat{p}$ exists and is uniquely determined, and $p \in (p_0, 1)$.

To prove the property of $\tilde{\pi}(p)$, we use the envelop theorem, which lead to $\tilde{\pi}'(p) = \pi_L(\tilde{\underline{u}}_{na}(p)) - \pi_H(\tilde{\underline{u}}_{na}(p))$. Since $\pi_L(\tilde{\underline{u}}_{na}(p)) - \pi_H(\tilde{\underline{u}}_{na}(p)) \le 0$ for $p < \hat{p}$ from Lemma 3 (as $\tilde{\underline{u}}_{na}(p) \ge \hat{u}$), we get that $\tilde{\pi}(p)$ is decreasing for $p < \hat{p}$. Similarly, $\tilde{\pi}(p)$ is increasing for $p > \hat{p}$, and the minimum of $\tilde{\pi}(p)$ is achieved at $p = \hat{p}$. ▮

Proof of Lemma 6: When $\underline{u}_{na} \le \hat{u}$, we obtain $\pi_H \le \pi_L$ by Lemma 3, thus $\pi = (1-\varepsilon)[p\pi_L + (1-p)\pi_H] + \varepsilon \min[\pi_L, \pi_H] = (1-\varepsilon)[p\pi_L + (1-p)\pi_H] + \varepsilon\pi_H = (1-\varepsilon)p\pi_L + (1-(1-\varepsilon)p)\pi_H = S(\underline{u}_{na}, p_1)$. Similarly, when $\underline{u}_{na} \ge \hat{u}$, we get $\pi_L \le \pi_H$ by Lemma 3, thus $\pi = (1-\varepsilon)[p\pi_L + (1-p)\pi_H] + \varepsilon \min[\pi_L, \pi_H] = (1-\varepsilon)[p\pi_L + (1-p)\pi_H] + \varepsilon\pi_L = S(\underline{u}_{na}, p_2)$. ▮

Proof of Lemma 7: Since $p_1 < p_2$, we have three possibilities: (1) $p_1 < p_2 < \hat{p}$, (2) $p_1 \le \hat{p} \le p_2$, (3) $\hat{p} < p_1 < p_2$. These three possibilities can be written as (1) $p_2 < \hat{p}$, (2) $p_1 \le \hat{p} \le p_2$, (3) $\hat{p} < p_1$. By Lemma 4, these three cases correspond to the following three cases. Case 1: $\hat{u} < \underline{\tilde{u}}_{na}(p_2) < \underline{\tilde{u}}_{na}(p_1)$; case 2: $\underline{\tilde{u}}_{na}(p_2) \le \hat{u} \le \underline{\tilde{u}}_{na}(p_1)$; case 3: $\underline{\tilde{u}}_{na}(p_2) < \underline{\tilde{u}}_{na}(p_1) < \hat{u}$. (See also Figure 1).

In case 1 where $\hat{u} < \underline{\tilde{u}}_{na}(p_2) < \underline{\tilde{u}}_{na}(p_1)$, for feasible menus of contracts with $\underline{u}_{na} \le \hat{u}$, the insurer's expected profit takes the form of $S(\underline{u}_{na}, p_1)$. As $S(\underline{u}_{na}, p_1)$ is a concave function of $\underline{u}_{na}$, it is an increasing function to the left of $\underline{\tilde{u}}_{na}(p_1)$. Since $\hat{u} < \underline{\tilde{u}}_{na}(p_1)$, for feasible menus of contracts with $\underline{u}_{na} \le \hat{u}$, the insurer's expected profit increases with $\underline{u}_{na}$. For feasible menus of contracts with $\underline{u}_{na} \ge \hat{u}$, the insurer's expected profit takes the form of $S(\underline{u}_{na}, p_2)$, which reaches its maximum at $\underline{\tilde{u}}_{na}(p_2)$. Since $\hat{u} < \underline{\tilde{u}}_{na}(p_2)$, the optimum of $\pi(\underline{u}_{na})$ is reached at $\underline{\tilde{u}}_{na}(p_2)$. Therefore, we obtain $\underline{u}^*_{na} = \underline{\tilde{u}}_{na}(p_2)$ in the optimal contract if $p_2 < \hat{p}$.

Similarly, in case 2 where $\underline{\tilde{u}}_{na}(p_2) \le \hat{u} \le \underline{\tilde{u}}_{na}(p_1)$. The insurer's expected profit $\pi(\underline{u}_{na})$ increases with $\underline{u}_{na}$ to the left of $\hat{u}$, and decreases with $\underline{u}_{na}$ to the right of $\hat{u}$. The maximum of $\pi(\underline{u}_{na})$ is reached at $\underline{u}_{na} = \hat{u}$. Therefore, we get $\underline{u}^*_{na} = \hat{u}$ in the optimum contract if $p_1 \le \hat{p} \le p_2$.

In case 3 where $\underline{\tilde{u}}_{na}(p_2) < \underline{\tilde{u}}_{na}(p_1) < \hat{u}$, the insurer's profit $\pi(\underline{u}_{na})$ decreases with $\underline{u}_{na}$ to the right of $\hat{u}$. The maximum of $\pi(\underline{u}_{na})$ is reached at $\underline{u}_{na} = \underline{\tilde{u}}_{na}(p_1)$. We get $\underline{u}^*_{na} = \underline{\tilde{u}}_{na}(p_1)$ in the optimum menu of contracts if $p_1 > \hat{p}$. ▮

Proof of Proposition 2.

If $p = \hat{p}$, we get $p_1 < \hat{p} < p_2$ for any $\varepsilon \ge 0$, thus, $\underline{u}_{na} = \hat{u}$ from Lemma 7.

If $p > \hat{p}$, for $\varepsilon \le 1 - \frac{\hat{p}}{p}$, we get $p_1 > \hat{p}$, and thus $\underline{u}^*_{na} = \underline{\tilde{u}}_{na}(p_1)$ by Lemma 7; for $\varepsilon \ge 1 - \frac{\hat{p}}{p}$, we obtain $p_1 \le \hat{p} < p_2$ and $\underline{u}^*_{na} = \hat{u}$ by Lemma 7. Thus $\underline{u}^*_{na} = \min(\underline{\tilde{u}}_{na}(p_1), \hat{u})$.

If $p < \hat{p}$, for $\varepsilon \le \frac{\hat{p}-p}{1-p}$, then we have $p_2 < \hat{p}$ and thus $\underline{u}^*_{na} = \underline{\tilde{u}}_{na}(p_2)$. For $p > \hat{p}$ and $\varepsilon \ge \frac{\hat{p}-p}{1-p}$, we obtain $p_1 < \hat{p} \le p_2$ and $\underline{u}^*_{na} = \hat{u}$. Thus $\underline{u}^*_{na} = \max(\underline{\tilde{u}}_{na}(p_2), \hat{u})$. ▮

Proof of Proposition 3. For a given $\varepsilon > 0$, if $p \in \left(\frac{\hat{p}-\varepsilon}{1-\varepsilon}, \frac{\hat{p}}{1-\varepsilon}\right)$, then $p_1 < \hat{p} < p_2$ and thus $\underline{u}^*_{na} = \hat{u}$ by Lemma 7. ▮

Proof of Proposition 4. We now only need to consider the case when both types of consumers are served when there is no ambiguity. If $p > \hat{p}$, for $\varepsilon < 1 - \frac{\hat{p}}{p}$, we have $\underline{u}^*_{na} = \underline{\tilde{u}}_{na}(p_1) < \hat{u}$, thus $\frac{\partial \underline{u}^*_{na}(\varepsilon)}{\partial \varepsilon} = \frac{\partial \underline{u}^*_{na}(p_1)}{\partial p_1}\frac{\partial p_1}{\partial \varepsilon} > 0$; for $\varepsilon > 1 - \frac{\hat{p}}{p}$, we have $\underline{u}^*_{na} = \hat{u}$ and thus $\frac{\partial \underline{u}^*_{na}(\varepsilon)}{\partial \varepsilon} = 0$. If p< $\hat{p}$, for $\varepsilon < \frac{\hat{p}-p}{1-p}$ we have $\underline{u}^*_{na} = \underline{\tilde{u}}_{na}(p_2) > \hat{u}$, and $\frac{\partial \underline{u}^*_{na}(\varepsilon)}{\partial \varepsilon} = \frac{\partial \underline{u}^*_{na}(p_2)}{\partial p_2}\frac{\partial p_2}{\partial \varepsilon} < 0$; if $\varepsilon > \frac{\hat{p}-p}{1-p}$, then $\underline{u}^*_{na} = \hat{u}$ and thus $\frac{\partial \underline{u}^*_{na}(\varepsilon)}{\partial \varepsilon} = 0$. Therefore, as information becomes more ambiguous, $\underline{u}^*_{na}$ moves toward $\hat{u}$. Once $\underline{u}^*_{na}$ reaches $\hat{u}$, it remains in this position as ε increases. ▮

REFERENCES

Amarante, Massimiliano, Mario Ghossoub, and Edmund Phelps, 2015. Ambiguity on the insurer's side: The demand for insurance. *Journal of Mathematical Economics* **58**, 61-78.

Anwar, Sajid, and Mingli Zheng, 2012. Competitive insurance market in the presence of ambiguity. *Insurance: mathematics and economics* **50(1)**, 79-84.

Berg, Kimmo, and Harri Ehtamo, 2009. Learning in nonlinear pricing with unknown utility functions. *Annals of Operations Research* **172(1)**, 375-392.

Cao, H. Henry, and Bao-Hong Sun, 2007. Value of Learning and Acting on Customer Information. *Tepper School of Business*, 484.

Chade, Hector, and Edward Schlee, 2012. Optimal insurance with adverse selection. *Theoretical Economics* **7(3)**, 571-607.

Chen, Zengjing, and Larry Epstein, 2002. Ambiguity, risk, and asset returns in continuous time. *Econometrica*, 1403-1443.

Dimmock, Stephen G., Roy Kouwenberg, Olivia S. Mitchell, and Kim Peijnenburg, 2013. Ambiguity attitudes and economic behavior. National Bureau of Economic Research.

Eichberger, Jürgen, and David Kelsey, 1999. E-capacities and the Ellsberg paradox. *Theory and decision* **46(2)**, 107-138.

Ellsberg, Daniel, 1961. Risk, ambiguity, and the Savage axioms. *The quarterly journal of economics*, 643-669.

Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci, 2004. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory* **118(2)**, 133-173.

Gilboa, Itzhak, 1987. Expected utility with purely subjective non-additive probabilities. *Journal of mathematical Economics* **16(1)**, 65-88.

Gilboa, Itzhak, and David Schmeidler, 1989. Maxmin expected utility with non-unique prior. *Journal of mathematical economics* **18(2)**, 141-153.

Gilboa, Itzhak, and David Schmeidler, 1994. Additive representations of non-additive measures and the Choquet integral. *Annals of Operations Research* **52(1)**, 43-65.

Gollier, Christian, 2014. Optimal insurance design of ambiguous risks. *Economic Theory* **57(3)**, 555-576.

Hansen, Lars Peter, and Thomas J. Sargent, 2001. Robust control and model uncertainty. *American Economic Review* **91(2)**, 60-66.

Hogarth, Robin M., and Howard Kunreuther, 1989. Risk, ambiguity, and insurance. *Journal of Risk and Uncertainty* **2(1)**, 5-35.

Huang, Rachel J., Arthur Snow, and Larry Y. Tzeng, 2015. Competitive Insurance Contracting with Ambiguity and Asymmetric Information.

Jeleva, Meglena, and Bertrand Villeneuve, 2004. Insurance contracts with imprecise probabilities and adverse selection. *Economic Theory* **23(4)**, 777-794.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, 2005. A Smooth Model of Decision Making under Ambiguity. *Econometrica* 1849-1892.

Koufopoulos, Kostas, and Roman Kozhan, 2015. Optimal insurance under adverse selection and ambiguity aversion. *Economic Theory*, 1-29.

Koufopoulos, Kostas, and Roman Kozhan, 2014. Welfare-improving ambiguity in insurance markets with asymmetric information. *Journal of Economic Theory* **151**, 551-560.

Knight, Frank H., 1921. Risk, uncertainty and prot. New York: Hart, Schaffner and Marx.

Kunreuther, Howard, Robin Hogarth, and Jacqueline Meszaros, 1993. Insurer ambiguity and market failure. *Journal of Risk and Uncertainty* **7(1)**, 71-87.

Maskin, Eric, and John Riley, 1984. Monopoly with incomplete information. *The RAND Journal of Economics* **15(2)**, 171-196.

Quiggin, John, 1982. A theory of anticipated utility. *Journal of Economic Behavior & Organization* **3(4)**, 323-343.

Rothschild, Michael, and Joseph Stiglitz, 1976. Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics* **90(4)**, 629-649.

Salo, Ahtia, and Martin Weber, 1995. Ambiguity aversion in first-price sealed-bid auctions. *Journal of Risk and Uncertainty* **11(2)**, 123-137.

Stiglitz, Joseph E., 1977. Monopoly, non-linear pricing and imperfect information: the insurance market. *The Review of Economic Studies*, 407-430.

Szalay, Dezs, 2008. Monopoly, non-linear pricing, and imperfect information: A reconsideration of the insurance market. Working Paper. Coventry: University of Warwick, Department of Economics. (Warwick economic research papers).

Vergote, Wouter, 2010. Insurance contracts with one-sided ambiguity. Available at SSRN 1609583.

Vier, Marie-Louise, 2012. Contracting in vague environments. *American Economic Journal: Microeconomics* **4(2)**, 104-130.

Zheng, Mingli, Chong Wang, and Chaozheng Li, 2015. Optimal nonlinear pricing by a monopolist with information ambiguity. *International Journal of Industrial Organization* **40**, 60-66.