An Optimal Insurance Design Problem under Knightian Uncertainty

Carole Bernard *
University of Waterloo

Shaolin Ji†
Shandong University

Weidong Tian‡
University of North Carolina at Charlotte.

*Department of Statistics and Actuarial Science, University of Waterloo, Canada.

†Department of Mathematics, Shandong University, China.

‡Corresponding author. Belk College of Business, University of North Carolina at Charlotte, NC 28223, USA. Phone: 704-687-7702. Fax:704-687-6987. Email: wtian1@uncc.edu. We are very grateful to the editor and an anonymous referee for several constructive and insightful comments on how to improve the paper.

An Optimal Insurance Design Problem under Knightian Uncertainty

Abstract

This paper solves an optimal insurance design problem in which both the insurer and the insured are subject to Knightian uncertainty about the loss distribution. The Knightian uncertainty is modeled in a multi-prior g-expectation framework. We obtain an endogenous characterization of the optimal indemnity that extends classical theorems of Arrow (1971) and Raviv (1979) in the classical situation. In the presence of Knightian uncertainty, it is shown that the optimal insurance contract is not only contingent on the realized loss but also on another source of uncertainty coming from the ambiguity.

Keywords: Knightian uncertainty, Insurance, Contingency
JEL Classification Codes: C61, D81.

1 Introduction

In the classical optimal insurance indemnity framework, it is assumed that both the insurer and the insured have full knowledge about the loss distribution. See, for example, Arrow (1971), Borch (1962), Raviv (1979), Gollier (1987), Gollier and Schlesinger (1996), Golubin (2006) and Bernard and Tian (2009a, 2009b). Some authors have been incorporated the presence of asymmetric information into the insurance market, such as in Breuer (2005), Crocker and Snow (1985), Cummins and Mahul (2003), Landsberger and Meilijson (1999), Rothschild and Stiglitz (1976)), but the assumption on the loss distribution is still retained.

Accurate estimation of the loss distribution is crucial for insurance market, since market participants purchase or sell the insurance contract and compute the insurance premium relying on the distribution of the underlying loss variable. While accurate estimation is possible in some insurance markets with a large amount of data and a stable history (for instance, automobile insurance and life insurance market), it becomes problematic in some other insurance markets, such as insurance against natural diasters or terrorism. In those insurance markets, the specific events in the insurance indemnity are so rare that the estimation methodology is subjective and limited. The realized loss, however, could be substantial. [1] Therefore, market participants tend to have serious concerns on the loss distribution.

This paper solves Arrow-Raviv's optimal insurance indemnity problem in the presence of agents' ambiguities on the loss distribution in a multi-prior framework of Knightian uncertainty. Knightian uncertainty has been first proposed in Knight (1921) and evidence is given by Ellsberg (1961). Bewley (2002) proposes a Knightian uncertainty using an incompleteness preference. We follow an axiomatic setting, initiated by Gilboa and Schmeidler (1989) and Schmeidler (1989). Intuitively, since the agent (insured or insurer) has no precise knowledge on the loss distribution, she believes that the true probability belongs to a set of probability distributions and her welfare is described by the worst case element in the set of probability distributions of the loss variable. In a precise manner, $\mathcal{P}$ is a set of probability measures, and the expected wealth of the agent is represented by a nonlinear expectation (represented by $\min_{P\in\mathcal{P}} \mathbb{E}_P[.]$), namely in a non-expected Choquet utility framework.

There is a vast literature extending the classical insurance design problem to non-expected utility framework, for instance, Doherty and Eeckhoudt (1995), Carlier and Dana (2008), Carlier, Dana and Shahidi (2003), Chateauneuf, Dana and Tallon (2000), Johnson et al. (1993). Among others, Chateauneuf, Dana and Tallon (2000) examine the optimal risk-sharing rule among agents in a Choquet capacity framework, and show that the classical characterization of Pareto efficient allocation holds when the Choquet capacities are identical for all agents. They further show that Pareto efficient is comonotone under some conditions on the capacities. Moreover, the characterization of Pareto efficient is solved as full insurance if there is no aggregate uncertainty. See also Billot, Chateauneuf, Gilboa and Tallon (2000), and Rigotti, Shannon and Strzalecki (2008). The characterization of Pareto efficient allocation on the insurance market, however, are not available as the insured risk ensures a *random* aggregate risk.

This paper is devoted to characterizing precisely the optimal insurance policy for a general class of non-linear expectation, a g-expectation, that extend the classical theorems of Arrow

[1] We refer to Cummins et al. (2004) and Froot (2001) for the catastrophe market and the insurance risk.

(1971) and Raviv (1979). A g-expectation is developed by Duffie and Epstein (1992) and by Chen and Epstein (2002) to deal with Knightian uncertainty (or ambiguity) in a continuous time setting. We argue that such a g-expectation is plausible in the insurance market for several reasons. First, this framework has received wide success in other economic contexts. For instance, Nishimura and Ozaki (2004, 2007), Obstfeld (1994) in economics, Uppal and Wang (2003) in finance. Second, even though the insurance indemnity's payout considered in this paper is realized only at one future time, the uncertainty in the insurance market might be driven by other factors in addition to the loss risk variable, and these factors could be path-dependent and state-dependent. Third, as g-expectation satisfies the time consistency property, we can extend the approach to discuss the optimal multi-period insurance contract, such as in Cooper and Hayes (1987), Vaquez and Watt (1999) in a Knightian uncertainty setting. On the other hand, Choquet-expected-utility is not time consistent unless it is generated by a recursive multiple-priors (Epstein and Schneider (2003)).

There are three main contributions in this paper. (1). We derive the optimal risk sharing rule when the insured has ambiguity on the loss distribution, while the insurer has perfect information and is risk neutral. We study the case when an insurance company has more information and is able to fully pool risk, and the insured has less information and is risk-averse. We derive an endogenous relationship to characterize the optimal insurance policy in this situation. It is shown that the optimal insurance indemnity is a deductible contract but its deductible depends on state event. We refer to is as a *contingency* contract. This reflects the ambiguity level of the insured: it reduces to Arrow's deductible when the ambiguity disappears. (2). We propose a "*quantity*" to measure the effect of ambiguity by using the variance of the (random) deductible. As it turns out, both the ambiguity and the risk aversion jointly affect the randomness of the deductible (as was also mentioned in Alary et al. (2010)). On the one hand, the variance tends to increase when the ambiguity level increases, while the risk aversion is fixed. On the other hand, the variance tends to increase when the agent becomes more risk-averse in a reasonable range of the risk aversion parameter. Hence, the effect of ambiguity becomes stronger for more risk-averse insured. (3). We extend Raviv (1979)'s optimal insurance policy in the g-expectation framework. This situation might occur in a reinsurance market in which both parties of the contract have ambiguities on the loss distribution. We characterize the optimal policy for the insurer by a nonlinear equation that extends Raviv (1979).

Our results are distinct from previous literature in that the optimal insurance indemnity is allowed to be state-dependent. The intuition is as follows. A nonissuable risk is generated in the presence of ambiguity, hence a contingency insurance contract emerges in the optimal insurance sharing agreement. In other words, given the ambiguity on the loss distribution, an insurance contract that depends only on the underlying loss variable is not rich enough to insure all the risk from the insured's perspective. Therefore, an optimal deductible, at least in theory, should write on the loss event itself instead of the loss amount. [2] To some extent our analysis shares some insights with analysis in incomplete market in Doherty and Schlesinger

[2] Actually, many insurance contracts in the real world are contingent. We use a simple example to illustrate a contingent contract. First consider an individual exposed to a loss risk of \$1 million or \$2 million with equal probability. If the insured is one hundred percent confident with her estimation of the risk, a deductible contract, say \$1 million deductible, works perfectly for her by paying \$0.6 million dollars (we choose the load factor as 20 percent). We now assume that the insured has ambiguity on the loss distribution estimation,

(1983a). Doherty and Schlesinger (1983a) show that the classical deductible contract is not optimal anymore in the presence of nonissuable risk (see also Gollier and Schlesinger (1995)). In this paper, we examine the optimality and the contingency of the insurance contract in the ambiguity framework, so we are able to study the effect of ambiguity on the insurance contract and how to risk sharing efficiently in insurance market in which the ambiguity level is strong.

Some previous papers have discussed the insurance contract with ambiguity from different perspectives. Hogarth and Kunreuther (1985, 1989) analyze the effect of ambiguity on the insurance market. Kunreuther, Hogarth and Meszaros (1993) investigate the market failure when the insurer has ambiguity on the market. Mukerji and Tallon (2001) demonstrate that the market has no trading at all when the difference of ambiguities is large enough among the agents in the market. See also Alary et al. (2010), Cabantous (2007), Ho, Keller and Kunreuther (1989) for related analysis in ambiguity frameworks. This paper has a different focus point. Instead of examining specific insurance contracts from the marketplace (such as a deductible or a coinsurance), we study the optimal insurance indemnity in a Knightian framework and characterize the contingency form of the optimal indemnity.

The layout of this paper is as follows. Section 2 introduces an insurance market with Knightian uncertainty, and presents the optimal insurance design problem in which both the insured and the insurer have ambiguity on loss distribution. In Section 3 we first characterize the optimal insurance contract for the insured when the insurer has perfect knowledge on the loss variable and risk-neutral. The solution in Section 3 is extended in Section 4 in which both agents have Knightian uncertainties. We discuss some special cases in details to illustrate our characterizations of the optimal indemnity. Conclusions are formulated in Section 5. All proofs are presented in the Appendix.

2 Model

In this section, we first introduce the notation related to the modeling of ambiguity, then present the optimal insurance design problem that we solve in subsequent sections.

2.1 Ambiguity Setting

The insurance contract is a one-period contract with maturity T and written on a loss variable X. For simplicity and to restrict ourselves on the ambiguity only, we assume that the state is perfectly observable at the maturity T of the contract. Because of the ambiguity

say, 50% certain on the loss \$1 million but has a serious concern that the loss can be significantly large, say \$10 million, in extremely bad situation. However the insured who is not sure how small the probability of this rare event is. In this case, a loss variable X is represented by $X(\omega_1) = 1, X(\omega_2) = 2, X(\omega_3) = 10$ on a sample space $\{\omega_1, \omega_2, \omega_3\}$. Clearly X captures the risk of loss but can't capture the uncertainty on the loss risk. Therefore, a better insurance contract for the insured might be written upon the occurrence of a specified event, instead of the loss variable X since the expected loss amount is not certain under ambiguity. Without the ambiguity, the distribution of the loss variable X is unique so a contract written on the loss variable is enough to insured the risk. However, when the distribution of the loss variable is not unique as it is the case in the ambiguity setting, a contract contingent on the state variable appears.

on the loss risk, however, the market participants are uncertain about the loss distribution and henceforth of the optimal insurance contract.

We follow the economic setting stated in Chen and Epstein (2002) (See page 1409-1415 and Appendix A-C therein) to introduce the ambiguity, which is written as follows. Let $B(\cdot) = \{B(t)\}_{t\geqslant 0}$ be a standard one-dimension Brownian motion that is defined on a complete probability space $(\Omega, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, P)$, X is defined over $\mathcal{F}_T$, and P is a benchmark probability measure. The market participant's ambiguity about the loss distribution is represented by a so called *aggregation* $g(y, z, t) : \mathbb{R} \times \mathbb{R} \times [0, T] \to \mathbb{R}$. Given any random variable observed at the maturity time T, under mild conditions which will be specified in the next section, it is known that there exists a *unique* process (Y_t) such that[3]

$$\begin{cases} Y_t &= Y + \int_t^T g(Y_s, z_s, s)ds - \int_t^T z_s dB(s), \quad t \in [0, T], \\ Y_T &= Y. \end{cases} \tag{1}$$

where $\{Y_t, z_t\}$ are determined endogenously. In this insurance market, we are interested by Y_0. This initial value Y_0 is called the *g-expectation* of Y_T, and can be written as

$$\mathcal{E}_g[Y_T] := Y_0. \tag{2}$$

The g-expectation Y_0 extends the expected value of the terminal reward Y_T, which represents the minimal expected value of Y_T among all possible ambiguity distributions.

This setting to model Knightian uncertainty using g-expectation is quite general. To illustrate this setting we next give three special cases that include the standard expectation under a subjective probability, the case of heterogeneous beliefs, and the case of stochastic recursive utility.

A. Classical Case.

When the aggregation $g(y, z, t) = 0$, the g-expectation is reduced to the classical expectation $\mathbb{E}_P[\cdot]$ under the subjective probability measure P. That is, $Y_0 = \mathbb{E}_P[Y]$. Therefore, when both the insured and insurer have a zero aggregation, and Y is the utility of the final wealth, it reduced to the classical cases studied in Arrow (1971) or Raviv (1979). This classical situation will serve as a benchmark case for our subsequent discussion.

B. Heterogeneous Beliefs

This heterogeneous probability situation arises when the aggregation is chosen as a linear functional form, $g(y, z, t) = bz$. In this case, it can be shown that the g-expectation is

$$\mathcal{E}_g[Y] = \mathbb{E}_P\left[\exp\left(-\frac{1}{2}b^2T + bB(T)\right)Y\right]. \tag{3}$$

[3]Technically speaking, a g-expectation utility is a solution of a backward stochastic differential equation (BSDE) in the terminology of Peng (1997). See El Karoui, Peng and Quenez (1997), and Peng (1997) for an exhaustive review on the theory of g-expectations. The idea to represent Knightian uncertainty in terms of g-expectation has been developed in Duffie and Epstein (1992), Chen and Epstein (2002).

The parameter b represents various beliefs of the market participants. To see the importance of b, consider one agent who has no ambiguity on the market and uses the expectation under a subjective probability P. Consider another agent who has ambiguity modeled by the aggregation $g(y, z, t) = bz$. We see that $\mathcal{E}_g[Y] = \mathbb{E}_Q[Y]$, where

$$\frac{dQ}{dP} = \exp\left\{-\frac{1}{2}b^2T + bB(T)\right\} \tag{4}$$

or equivalently, for all measurable subset A,

$$Q(A) = \mathbb{E}_P\left[\exp\left\{-\frac{1}{2}b^2T + bB(T)\right\}1_A\right]. \tag{5}$$

Therefore, the second agent has a different belief about the loss distribution. The relative entropy can help us to compare the ambiguity level. After some computations (using Girsanov theorem), it can be shown that the entropy is equal to

$$\mathbb{E}_P\left[\frac{dQ}{dP}\log\left(\frac{dQ}{dP}\right)\right] = \frac{1}{2}b^2T. \tag{6}$$

which can be intuitively interpreted as follows. The ambiguity level increases when b increases, and the ambiguity disappears when $b = 0$. The positive parameter b directly measures the ambiguity level or the divergence among the heterogeneous beliefs.

This example is related to two particular cases of asymmetric information situations. Consider one case in which the insured has a perfect observation of her risk and thus no uncertainty. But the insurer might have another probability distribution about the loss risk. Another sensible case corresponds to the opposite situation. The insurer has access to historical data and has therefore a better knowledge of the loss while the insured has a different probability distribution. Overall, the probability distributions of the same loss variable for both the insured and the insurer can be different and thus the ambiguity occurs. Therefore, this example is related to Rothschild and Stiglitz (1976)'s analysis on the insurance market under asymmetric information.

We now extend the case in which both agents believe that the loss distribution belongs to a class of probability distributions, instead of a single distribution. This kind of ambiguity can be modeled by the (stochastic) recursive utility, which is also a special case of g-expectation.

C. Recursive Stochastic Utility

The stochastic recursive utility is developed in Duffie and Epstein (1992) and extended in Chen and Epstein (2002). Consider a *subliner* function g such that there exists a convex subset $\Lambda \subseteq \mathbb{R}$ that verifies for all z, t,

$$g(z, t) = \sup_{\lambda \in \Lambda} < \lambda, z > .$$

It has been proved in Chen and Epstein (2002), Theorem 2.2, that the g-expectation is obtained as the worse expectation (with respect to a set $\mathcal{P}$ of probability measures)

$$\mathcal{E}_g[Y] = \min_{P_\theta \in \mathcal{P}} \mathbb{E}_{P_\theta}[Y], \quad \text{with} \quad \frac{dP_\theta}{dP} = \exp\left\{\int_0^T \theta_t dB(t) - \frac{1}{2}\int_0^T |\theta_t|^2 dt\right\} \tag{7}$$

where θ takes value in Λ and satisfies the Novikov condition [4].

The set of probability measures $\mathcal{P}$ represents the Knightian uncertainty and can be described as follows

$$\mathcal{P} := \left\{ P_\theta : B(t) - \int_0^t \theta_s ds \ \text{ is a } P_\theta\text{-Brownian motion.} \right\}$$

In this setting, the g-expectation corresponds to the multi-prior $\min_{P_\theta \in \mathcal{P}} \mathbb{E}_{P_\theta}[.]$ introduced by Gilboa and Schmeidler (1989) and Schmeidler (1989). The set Λ represents the ambiguity level. The larger Λ the larger of the ambiguity on the loss distribution. For example, let $\Lambda = \{x : |x_i| \leqslant \kappa, \forall i\}$, then $g = \kappa|z|$ and this example has been used in several economical situations for ambiguity [5]. The larger the positive number κ is, the higher the ambiguity on the loss distribution.

2.2 Pareto-efficient Insurance Design

This section presents the optimal insurance design problem in the presence of Knightian uncertainty. Assume the insured faces a risk of loss denoted by the positive random variable X which is observed at its maturity T. The insured pays an upfront premium P_0 to an insurer at time zero in return for a promise to receive a contractually agreed indemnity payment $I(X)$ upon the occurrence of a specified event. Assume the initial wealth of the insured and insurer is W_0 and W_0^n, respectively. Then the terminal wealth of the insured is

$$W = W_0 - P_0 - X + I(X) \tag{8}$$

and the terminal wealth of the insurer is

$$W^n = W_0^n + P_0 - c(I(X)) - I(X) \tag{9}$$

where $c(.)$ represents the cost. To highlight the effect of ambiguity we choose a linear cost structure, $c(I(X)) = \eta I(X)$ for a constant $\eta > 0$. We also assume that the premium P_0 is a function of the actuarial value $\mathbb{E}_P[I(X)]$. These two assumptions are standard in the optimal insurance design literature and we will explain more the second assumption below.

The insured and insurer's risk aversion are interpreted by the strictly concave functions $U(\cdot)$ and $V(\cdot)$, respectively. As demonstrated in the last section, both the insured and insurer's ambiguity on the loss variable are introduced by a g-aggregation $g(\cdot)$ and another g-aggregation $f(\cdot)$, respectively. We assume that the functions $g(y, z, t)$ and $f(x, \pi, t)$: $\mathbb{R} \times \mathbb{R} \times [0, T] \to \mathbb{R}$ satisfy

(H1) $g(\cdot)$ and $f(\cdot)$ are continuous in $\mathbb{R} \times \mathbb{R} \times [0, T]$ for $a.a.\omega$ and have continuous bounded derivatives respectively in (y, z) and (x, π) for each $t \in [0, T]$;

(H2) both $g(0, 0, \cdot)$ and $f(0, 0, \cdot)$ are square-integrable.

[4] That is, $\mathbb{E}_P \left[exp \left(\frac{1}{2} \int_0^T |\theta_t|^2 dt \right) \right] < \infty$.

[5] See Chen and Epstein (2002) and Nishimura and Ozaki (2007) for its application of this particular g-expectation.

(H3) $g(\cdot)$, $f(\cdot)$ are strictly concave functions with respect to both y and z.

The problem under consideration is the following optimal insurance design problem:

$$\max_{I:0\leqslant I(X)\leqslant X, x(0)\geqslant k} y(0)$$

where $x(0)$ is the g-expectation of $V(W^n)$ under the aggregation $f(\cdot)$, and $y(0)$ is the g-expectation of $U(W)$ under the aggregation $g(\cdot)$. The constant k is a given acceptable utility floor of the insurer, $x(0) \geqslant k$ is often called the *participation constraint* in principal-agency literature (see Raviv (1979)). Given the interpretation of $x(0)$ in terms of the minimal expected value of $V(W^n)$ under a set of probability measures, the participation constraint states that under any available probability measure under ambiguity, the corresponding expected value of $V(W^n)$ is always greater or equal to k. $0 \leqslant I(X) \leqslant X$ is also standard in literature.

This problem can be reduced to a sequence of the following optimal insurance design problem (for each $0 \leqslant \Delta \leqslant \mathbb{E}_P[X]$):

$$(A) \qquad \max_I y(0) \quad s.t. \quad \begin{cases} 0 \leqslant I(X) \leqslant X \\ \mathbb{E}_P[I(X)] = \Delta \\ x(0) \geqslant k. \end{cases}$$

under a further constraint $\mathbb{E}_P[I(X)] = \Delta$ for any $\Delta \in [0, \mathbb{E}_P[X]]$. When the insurer is risk-neutral this condition can be viewed as a premium condition (see Problem (B) below). [6] In the general case, this condition states that under the benchmark measure P the expected value of the indemnity $I(X)$ is fixed as Δ. In what follows we focus on Problem (A) for a fixed Δ.

Definition 2.1 *The insurance indemnity $I(X)$ is called* admissible *for a given loss X and an acceptable utility floor k, if $I(X)$ satisfies the constraints in Problem (A). We shall denote by $\mathcal{A}(X,k)$, the set of all admissible insurance indemnities $I(X)$.*

An admissible insurance indemnity $I^*(X)$ is called *optimal* if it attains the maximum of $y(0)$ over $\mathcal{A}(X,k)$. There is a remarkable feature in this setting. In the presence of ambiguity, as will be shown in detail, the ambiguity aggregation $g(\cdot)$ and the loss variable X *jointly* affect the insured and the insurer's decision. In the classical situation, since the aggregation is null, only the loss variable plays a role in the optimal insurance contract.

3 Optimal Insurance Design from the Insured's Perspective

In this section we consider a special case of Problem (A) in which the insurer is risk-neutral and has no ambiguity about the loss distribution, that is $f = 0$. Note that $x(0) =$

[6] To impose such a condition is quite standard to derive the optimal insurance contract as in Arrow (1971). It is also available to extend it to be a g-expectation such as in Ji and Zhou (2010).

$\mathbb{E}_P\left[W_0^n + P_0 - (1+\eta)I(X)\right]$. Problem (A) is reduced to be an extended Arrow problem as follows

$$(B) \qquad \max_I y(0) \quad s.t. \quad \begin{cases} 0 \leqslant I(X) \leqslant X \\ \mathbb{E}_P[I(X)] = \Delta. \end{cases}$$

Problem (B) can be viewed as a Pareto-efficient risk sharing between a risk neutral insurer and an insured is subject to Knightian uncertainty.

Let $I^*(x)$ be optimal and $(y^*(\cdot), z^*(\cdot))$ be the solution of

$$\begin{cases} -dy(t) = g(y(t), z(t), t)dt - z(t)dB(t), \\ y(T) = U(W^*). \end{cases} \tag{10}$$

Set

$$g_y^*(t) = g_y(y^*(t), z^*(t), t) \text{ and } g_z^*(t) = g_z(y^*(t), z^*(t), t). \tag{11}$$

where g_y, g_z represent the first-order derivatives of the function $g(y, z, t)$ with respect to y and z, respectively.

In order to derive the maximum principle, it is useful to introduce the adjoint process associated with $I^*(X)$, $n(\cdot)$, which is defined by the following stochastic differential equation

$$\begin{cases} dn(t) = g_y^*(t)n(t)dt + g_z^*(t)n(t)dB(t), \\ n(0) = 1. \end{cases} \tag{12}$$

The next theorem presents a complete characterization of the optimal insurance indemnity when the insured has ambiguity while the insurer is risk-neutral and has no ambiguity about the loss variable.

Theorem 3.1 *Assume (H1), (H2) and (H3). Then there exists a unique optimal insurance indemnity to Problem (B). The optimal insurance indemnity $I^*(X)$ is characterized by*

$$I^*(X) = \min\left\{(X-d)^+, X\right\} \tag{13}$$

where

$$d = W_0 - P_0 - (U')^{-1}\left(\frac{\mu^*}{n(T)}\right), \tag{14}$$

μ^* is a constant and $(U')^{-1}$ is the inverse function of $U'(\cdot)$, $n(T)$ is the solution of the adjoint equation (12) at time T and μ^* is solved by the equation $\mathbb{E}_P[I^*(X)] = \Delta$.

Proof. See Appendix. □

The equation (13) is an *endogenous* characterization of the optimal insurance contract $I^*(X)$. According to Theorem 3.1, the optimal insurance contract can be interpreted as a deductible contract with a random deductible d (since $n(T)$ is random). The deductible d is deterministic if and only if there is no ambiguity. In the right side of equation (13), the adjoint process $n(T)$ relies on $I^*(X)$ in the forward stochastic differential equation (12).

In fact, $\frac{1}{n(T)}$ is the shadow price of the state as a Lagrange multiplier in the optimization problem (B). The hard part in this characterization lies in the determination of $\{y^*(\cdot), z^*(\cdot)\}$. In the g-expectation literature (see El Karoui, Peng and Quenez (1997, 2001)), $y(\cdot)$ is often calculated as the worse case of the expectations among a group of probability measures and $z(\cdot)$ is the Malleation derivative of $y(\cdot)$.

We illustrate this characterization in the three special cases, which were introduced before in Section 2.

Case A. Classical Case.

When $g(\cdot) = 0$, it is easy to see that the adjoint process $n(T)$ is independent of the insurance contract $I^*(X)$, that is, $n(T) = n(0) = 1$. As a consequence, the deductible level is constant, and equal to

$$d = W_0 - P_0 - (U')^{-1}(\mu^*),$$

and the optimal contract $I^*(X)$ is reduced to the Arrow (1971)'s optimal deductible.

Case B. Heterogenous Beliefs.

Let $g(y, z, t) = bz$. As explained above, Problem (B) becomes

$$(C) \quad \max_I \quad \mathbb{E}_Q[U(W)] \quad \text{subject to} \begin{cases} 0 \leqslant I(X) \leqslant X, \\ \mathbb{E}_P[I(X)] = \Delta. \end{cases}$$

where Q is the probability measure by the insured. This problem simply states that the insured is a subjective expected utility maximizer with subjective probability Q and the insurer is risk-neutral and uses the objective probability P to calculate the premium. The optimal contract is written as

$$I^*(\omega) = \max\{X(\omega) - d(\omega), 0\}$$

where $d(\omega)$ is a state-dependent deductible. In this case, as $g(\cdot)$ is a linear function, the adjoint process $n(T)$ is also independent of the optimal insurance contract $I^*(\cdot)$ and depends only on the ambiguity. In fact, $n(T) = \exp\left\{-\frac{1}{2}b^2T + bB(T)\right\}$, then the deductible level is $d = W_0 - P_0 - (U')^{-1}\left(\frac{\mu^*}{n(T)}\right)$ where the constant μ^* solves

$$\mathbb{E}_P\left[\min\left\{\left(X - W_0 + P_0 + (U')^{-1}\left(\frac{\mu^*}{n(T)}\right)\right)^+, X\right\}\right] = \Delta.$$

There are several important implications about the solution $I^*(\cdot)$ as follows. *(1)*, $I^*(\cdot)$ depends on the loss X and on the state event ω. In other words, the optimal indemnity cannot be written on the loss variable X itself, and the optimal insurance indemnity I^* must depend on the state event ω when the insured has ambiguity. *(2)*, the contingency feature of the optimal indemnity $I^*(X)$ is related to the incompleteness of insurance market: as the insured has not a perfect knowledge on some states (say, for example, catastrophe events), the insurance contract that depends on the loss only can't fully insure the risk of X in some scenarios. As there is uncertainty on the probability distribution, there is an uninsurable risk in the insurance market and $I^*(X)$ has to depend on this additional source of uncertainty.

To some extent our solution is similar to the analysis in Doherty and Schlesinger (1983b), Gollier and Schlesinger (1995). In Doherty and Schlesinger (1983b), Arrow's deductible is shown not optimal anymore in the presence of uninsurable risk. Gollier and Schlesinger (1995) examine the optimal form of insurance for multiple risks. We demonstrate the same result in the presence of ambiguity, and present the explicit solution of the optimal insurance contract. *(3),* to determine the parameter μ^* in the last formula, the joint distribution of X and the Radon-Nikodym derivative $\frac{dQ}{dP}$ is required. *(4),* the insured's ambiguity is related to the randomness of the deductible level.

The last point (4) captures the effect of ambiguity on the optimal indemnity. To understand this point, let us ignore the computation of the parameter μ^* and use the variance of $U'^{-1}\left(\frac{1}{n(T)}\right)$ as a key *quantity* to measure the effect of the ambiguity of the insured. In what follows we perform some computations for a class of utility functions $U(\cdot)$ and verify that the variance of the deductible is a helpful proxy to measure the ambiguity effect .

Case of Log utility

Assume first that $U(x) = \log(x)$. After straightforward calculations, the variance of the deductible $d(\omega)$ is proportional to $e^{b^2T} - 1$ which we write as

$$Variance(d) = \alpha\left(e^{b^2T} - 1\right) \tag{15}$$

where $\alpha > 0$ is a constant. By virtue of equation (6), we have an expression of the variance of the deductible as a function of the entropy

$$Variance(d) = \alpha\left(e^{2\mathbb{E}_P\left[\frac{dQ}{dP}\log\left(\frac{dQ}{dP}\right)\right]} - 1\right). \tag{16}$$

Equation (16) is useful for our analysis: it links the variance of the deductible and the relative entropy explicitly. The left side of (16) is the variance of the deductible level. The right side is expressed explicitly by the relative entropy of Q with respect to the subjective probability measure P. The higher the ambiguity, the higher the relative entropy $\mathbb{E}_P\left[\frac{dQ}{dP}\log\left(\frac{dQ}{dP}\right)\right]$, therefore, the larger the variance of the deductible. In the extreme case, when $P = Q$, that is, no ambiguity at all, the variance of $d(\omega)$ is equal to 0, so the deductible level becomes deterministic.

To further understand the joint effect of the ambiguity and risk aversion on the variance of the deductible, we next compute the explicit solution for a CRRA utility function.

Case of CRRA utility

Consider now $U(x) = \frac{x^{1-A}}{1-A}, A > 0, A \neq 1$. Straightforward calculation yields

$$Variance(d) = \beta\left(e^{\frac{b^2T(2-A)}{A^2}} - e^{\frac{b^2T(1-A)}{A^2}}\right), \tag{17}$$

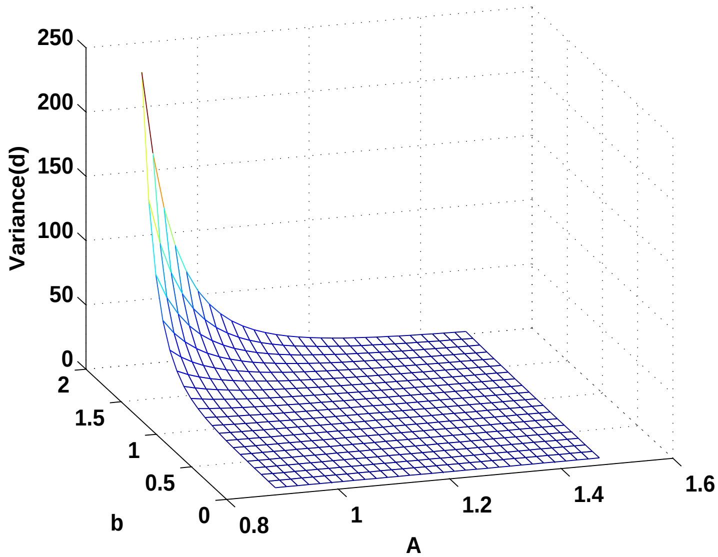
where $\beta > 0$ is constant and one can verify that [7]

$$\frac{\partial Variance(d)}{\partial b} \geqslant 0, \quad \text{for } A \leqslant 2 \tag{18}$$

[7] Note that $\frac{\partial Variance(d)}{\partial b} = \beta\frac{2bT}{A^2}e^{b^2T(1-A)/A^2}\{e^{b^2T/A^2}(2-A) - (1-A)\}$ and $\frac{\partial Variance(d)}{\partial A} = -\beta\frac{e^{b^2T(1-A)/A^2}}{A^3}\{e^{b^2T/A^2}(4-A) - (2-A)\}$.

Figure 1: Effect of Ambiguity and Risk Aversion

This graph displays the variance of the deductible as a function of the relative aversion parameter A and the ambiguity index b. A moves between 0.9 and 1.5, b moves between 0.1 and 2.



and

$$\frac{\partial Variance(d)}{\partial A} \leqslant 0, \quad \text{for } A \leqslant 4 \tag{19}$$

Equation (18) shows that the variance of the deductible level increases when the ambiguity level b increases, while keeping the risk aversion parameter A constant and bounded by 2. This result is intuitive as the insured's ambiguity increases the randomness of the deductible level. Equation (19) states that the effect of ambiguity becomes stronger when the insured is more risk-averse for a reasonable range of the risk aversion parameter A. As insured tends to be very risk averse and ambiguity simultaneously on some loss distribution in insurance market, our analysis show that the classical deductible insurance is far away from the optimal one from the insured's perspective.

Figure 1 displays the joint effect of the ambiguity index and the risk aversion. As shown, the variance increases when the ambiguity index b increases, or when the risk aversion parameter A decreases. Moreover, the effect of ambiguity is more significant than the effect of the risk aversion. We also observe that the variance behaviors relatively "flat" when the risk aversion parameter changes. It implies that the ambiguity contributes a first-order effect while risk averse effect is of second order.

Remark. Problem (C) can be solved in a slightly more general framework, alternatively, when the probabilities of both the insured and the insurer are equivalent. [8] Similar to Bernard and Tian (2009a, 2009b), it can be shown that the optimal contract $I^*(\cdot)$ has the

[8] It means that both the insured and the insurer agree on all zero probability events and one hundred percent sure events, but might have different probabilities otherwise. If so, $\frac{dQ}{dP}$ exists.

same form as stated in (13) and the deductible level is equal to

$$d = W_0 - P - U'^{-1}\left(\mu^* \frac{dP}{dQ}\right)$$

which is identical to the previous solution of $I^*(X)$ since $\frac{dQ}{dP} = n(T)$. Therefore, the variance of the deductible level is the same as (up to a constant) of the variance of $U'^{-1}\left(\mu^* \frac{dP}{dQ}\right)$ in general.

Case C. Stochastic Recursive Utility.

Let $g(y, z, t) = k|z|$ in Chen and Epstein (2002). In this case Assumption (H1) does not hold since the absolute function is not differentiable. It can still be shown that $I^*(X) = \max\{(X - d)^+, X\}$ is a sufficient condition for the optimal contract.[9] The numerical method for backward stochastic differential equation can be used to find $n(T)$ and μ^* numerically. See Ma, Protter, Martin and Torres (2002).

4 Pareto-optimal insurance contract

In this section we present the characterization of the optimal indemnity solution of Problem (A). Let $I^*(X)$ be optimal and $(y^*(\cdot), z^*(\cdot))$ (resp. $(x^*(\cdot), \pi^*(\cdot))$ be the corresponding utility processes of (10) for aggregation $g(\cdot)$ and aggregation $f(\cdot)$, respectively.

Set

$$g_y^*(t) = g_y(y^*(t), z^*(t), t) \text{ and } g_z^*(t) = g_z(y^*(t), z^*(t), t);$$
$$f_x^*(t) = f_x(x^*(t), \pi^*(t), t) \text{ and } f_\pi^*(t) = f_\pi(x^*(t), \pi^*(t), t)$$

Similar to the extended Arrow problem of the last section, the adjoint processes $n(\cdot)$ and $m(\cdot)$ associated with I^* are defined as follows.

$$\begin{cases} dn(t) = g_y^*(t)n(t)dt + g_z^*(t)n(t)dB(t), \\ n(0) = 1, \\ dm(t) = f_x^*(t)m(t)dt + f_\pi^*(t)m(t)dB(t), \\ m(0) = 1. \end{cases} \tag{20}$$

The main result of this paper is given as follows.

Theorem 4.1 *Assume (H1), (H2) and (H3). The insurance indemnity $I^*(X)$ is optimal to Problem (A) if and only if there exists constants $\lambda > 0$ and constant μ such that*

$$I^*(X) = \min\left\{(X - H(\lambda, \mu; X))^+, X\right\} \tag{21}$$

where $H(\lambda, \mu; X)$ is a random variable such that $H(\lambda, \mu; X)(\omega)$, for any $\omega \in \Omega$, is the unique solution y of the following equation

$$U'(W_0 - P_0 - y)\, n(T) - \lambda V'(W_0^n + P_0 - (1 + \eta)(X(\omega) + y))\, m(T) = \mu, \tag{22}$$

[9] Precisely, the derivative involved in $n(t)$ is replaced by the undifferentiated. See Ji and Zhou (2010) for details.

where $m(T)$ and $n(T)$ are the solutions of the adjoint equations (20) at time T. The parameters λ and μ in the solution are solved by both the premium constraint and the participation constraint simultaneously.

Proof. See Appendix. □

When $f(\cdot) = g(\cdot) = 0$, then $n(T) = m(T) = 1$, and equation (22) is reduced to be

$$U'(W_0 - P_0 - y) = \mu + \lambda V'\left(W_0^n + P_0 - (1+\eta)(X(\omega) + y)\right). \tag{23}$$

It is easy to see that y is a (deterministic) function of the loss variable and the last equation becomes the fundamental functional equation (10) in Raviv (1979). When $f(\cdot) = 0$, and the insurer is risk-neutral, then $m(T) = 1$ and $V'(\cdot)$ is a constant, and Theorem 4.1 is reduced to Theorem 3.1. In general, $H(\lambda, \mu; X)$ depends also on the state variable so the optimal insurance characterized by equation (21) is contingency.

In the heterogeneous belief environment, $H(\lambda, \mu; X)$ can be given explicitly. Let $g(y, z, t) = b_1 z$, and $f(y, z, t) = b_2 z$. The parameters b_1 and b_2 represent the ambiguity level on the loss variable for the insured and the insurer, respectively. Assume $U(x) = V(x) = \log(x)$. Then $H(\lambda, \mu; X)$ is the solution y of the following equation

$$\frac{n(T)}{W_0 - P_0 - y} - \lambda \frac{m(T)}{W_0^n + P_0 - (1+\eta)(X + y)} = \mu \tag{24}$$

where

$$\begin{cases} n(T) \equiv \frac{dQ^{ed}}{dP} = \exp\left\{-\frac{1}{2} b_1^2 T + b_1 B(T)\right\}, \\ m(T) \equiv \frac{dQ^{er}}{dP} = \exp\left\{-\frac{1}{2} b_2^2 T + b_2 B(T)\right\} \end{cases} \tag{25}$$

in which Q^{ed} and Q^{er} represent the respective estimated probability measures of the insured and the insurer. The solution y in equation (24) is obtained from

$$a_1 + a_2 y + (1+\eta)\mu y^2 = 0$$

and written as follows.

$$H(\lambda, \mu; X) = \frac{-a_2 - \sqrt{a_2^2 - 4\mu(1+\eta)a_1}}{2(1+\eta)\mu} \tag{26}$$

where

$$a_2 = -\mu\left(W_0^n + P_0 + (1+\eta)(W_0 + P_0 - X)\right) - \lambda \frac{dQ^{er}}{dP} + (1+\eta)\frac{dQ^{ed}}{dP}$$

and

$$\begin{aligned} a_1 &= \mu(W_0 - P_0)\left(W_0^n + P_0 - (1+\eta)X\right) + \lambda(W_0 - P_0)\frac{dQ^{er}}{dP} \\ &\quad - \left(W_0^n + P_0 - (1+\eta)X\right)\frac{dQ^{ed}}{dP}. \end{aligned}$$

We now demonstrate the optimal insurance indemnity when both the insured and the insurer have ambiguity on the loss variable. By Theorem 4.1, the optimal insurance indemnity depends on the loss variable X and the ambiguities, represented by $\{\frac{dQ^{ed}}{dP}, \frac{dQ^{er}}{dP}\}$. By equation (25), $\frac{dQ^{ed}}{dP}$ and $\frac{dQ^{er}}{dP}$ satisfy

$$\log\left(\frac{dQ^{er}}{dP}\right) = \frac{b_2}{b_1}\log\left(\frac{dQ^{ed}}{dP}\right) + \frac{1}{2}b_2(b_2 - b_1)T. \tag{27}$$

Therefore, we are able to plot the optimal insurance indemnity I^* that depend on the loss variable X and the Radon-Nikodym $\frac{dQ^{ed}}{dP}$ of the insured's probability measure. Figure 2 displays the optimal indemnity in three different cases: $b_1 < b_2, b_1 = b_2$, and $b_1 > b_2$, respectively. As shown, the optimal indemnity is, in its shape, similar to Raviv's optimal indemnity when $b_1 = b_2 = 0$, but the effects on the "*coinsurance*", the marginal coverage $\frac{\partial I^*}{\partial X}$, are different. In Panel A, $b_1 < b_2$, so the insurer is more ambiguity on the loss distribution than the insured. We see that the coinsurance of insurance indemnity decreases when the insured has more insurance. The intuition is as follows. As the insurer has a higher ambiguity, the Pareto-efficient contract requires a higher coinsurance for the insurer. Hence, the coinsurance becomes relatively smaller for the insured even when the insured's ambiguity increase. On the other hand, when $b_1 > b_2$ in Panel C, the insured is more ambiguity on the loss distribution than the insurer. By the same intuition, the coinsurance of the insurance indemnity increases when $\frac{dQ^{ed}}{dP}$ increases. When the ambiguity of the insured dominates, the coinsurance increases with the increase of the ambiguity of the insured, in a Pareto-efficient agreed insurance contract. In Panel B, when the insured and the insurer have the same ambiguity, the coinsurance behaves stable. In all Panels, the coinsurance and the deductible are all state-dependent.

The three following tables give a_1, a_2 and H for the range of parameters considered in Panel A, B and C of Figure 2.

Table 1: Values of a_1, a_2 and H used for Figure 2, case $b_1 < b_2$

X	0			3			6			9		
$\frac{dQ^{ed}}{dP}$	a_1	a_2	H	a_1	a_2	H	a_1	a_2	H	a_1	a_2	H
0.5	89.3	-27.4	3.85	67.8	-24.1	3.31	46.4	-20.8	2.58	24.9	-17.5	1.58
0.75	92	-28	3.87	71.3	-24.7	3.4	50.7	-21.4	2.76	30.1	-18.1	1.88
1	97	-28.9	3.95	77.2	-25.6	3.56	57.4	-22.3	3.02	37.6	-19	2.28
1.25	104	-30.2	4.07	85.5	-26.9	3.76	66.5	-23.6	3.34	47.6	-20.3	2.76
1.5	114	-31.8	4.21	96.2	-28.5	3.99	78	-25.2	3.7	59.9	-21.9	3.28

5 Conclusion

This paper characterizes in a theoretical manner the optimal insurance contract when both the insurer and the insured have uncertainties on the loss distribution. It turns out that the

Table 2: Values of a_1, a_2 and H used for Figure 2, case $b_1 = b_2$

X	0			3			6			9		
$\frac{dQ^{ed}}{dP}$	a_1	a_2	H	a_1	a_2	H	a_1	a_2	H	a_1	a_2	H
0.5	88	-27.3	3.82	66.5	-23.9	3.27	45.1	-20.7	2.52	23.6	-17.3	1.51
0.75	86.5	-27.2	3.74	65.9	-23.9	3.23	45.3	-20.6	2.54	24.6	-17.3	1.58
1	85	-27.2	3.67	65.2	-23.9	3.2	45.4	-20.6	2.55	25.6	-17.3	1.65
1.25	83.5	-27.2	3.6	64.5	-23.9	3.16	45.5	-20.6	2.57	26.6	-17.3	1.73
1.5	82	-27.2	3.5	63.8	-23.9	3.13	45.7	-20.6	2.58	27.5	-17.3	1.8

Table 3: Values of a_1, a_2 and H used for Figure 2, case $b_1 > b_2$

X	0			3			6			9		
$\frac{dQ^{ed}}{dP}$	a_1	a_2	H	a_1	a_2	H	a_1	a_2	H	a_1	a_2	H
0.5	88.9	-27.4	3.84	67.4	-24.1	3.3	46	-20.8	2.56	24.5	-17.5	1.56
0.75	86.6	-27.2	3.75	66	-23.9	3.24	45.3	-20.6	2.54	24.7	-17.3	1.59
1	84.2	-27.1	3.65	64.4	-23.8	3.17	44.6	-20.5	2.52	24.8	-17.2	1.61
1.25	81.7	-26.9	3.55	62.7	-23.6	3.1	43.7	-20.3	2.49	24.7	-17	1.62
1.5	79.1	-26.7	3.45	60.9	-23.4	3.03	42.8	-20.1	2.45	24.6	-16.8	1.64

optimal insurance contract has a similar shape with the insurance contract in the classical framework, but the insurance contract is contingent on another source of uncertainty, in addition to the loss variable. This additional source of uncertainty is driven by the ambiguity on the probability distribution and the optimal contract itself depends on the uncertainty and how different the insurer and the insured view about the distribution of the loss risk are.

Some loss risk in the insurance market such as catastrophe risk is rare but has a significant impact on market participants when it occurs. As a consequence, agents could have very different estimation on the loss distribution for the catastrophe risk. The paper demonstrates that the higher the ambiguity level of the market participants, the bigger the difference between the "contingent" optimal contract with the one without ambiguity. This paper shed some light on the optimal risk sharing in those insurance markets where the ambiguity could be strong such as earthquake insurance, tornadoes risk, volcanoes and insurance against natural disasters in general.

Appendix

Before proving Theorem 4.1, we first prove two lemmas. We use the notations defined in the main text of the paper and will refer to the assumptions (H1), (H2) and (H3) as introduced in Section 2.

Note that the following $\mathbb{R}-$valued functionals

$$\begin{aligned} I(X) &\mapsto y(0), \\ I(X) &\mapsto x(0), \end{aligned}$$

are concave under Assumption (H3). Applying classical results of convex analysis (Luenberger (1969)), it is straightforward to prove the following lemma.

Lemma A.1 *We suppose (H1), (H2) and (H3). There exist real numbers $\lambda > 0$ and μ such that the maximum objective function is given by*

$$\max_{I(X)} \left\{ y(0) + \lambda(x(0) - k) - \mu(\mathbb{E}_P[I(X)] - \Delta) \right\}. \tag{A1}$$

Furthermore, if the maximum is attained in Problem (A) by $I^*(X)$, then it is attained in (A1) by $I^*(X)$. Conversely, suppose that there exist $\lambda^o > 0$, μ^o and $0 \leqslant I^o(X) \leqslant X$ such that the maximum of

$$\max_I \{ y(0) + \lambda^o(x(0) - k) - \mu^o(\mathbb{E}_P[I] - \Delta) \}$$

is achieved with $I^o(X)$, then the maximum is achieved in Problem (A) by $I^o(X)$.

The second result is related to the *first-order necessary condition* of the optimal indemnity.

Let $I^*(X)$ be optimal and $(y^*(\cdot), z^*(\cdot))$ (resp. $(x^*(\cdot), \pi^*(\cdot))$) be the corresponding state processes of (10) for aggregation $g(\cdot)$ and aggregation $f(\cdot)$, respectively. Take an *arbitrary* $I(X)$ such that $0 \leqslant I(X) \leqslant X$, and $0 \leqslant \rho \leqslant 1$ such that $I^\rho(X) = I^*(X) + \rho(I(X) - I^*(X))$ falls in $[0, X]$. Define $(y^\rho(\cdot), z^\rho(\cdot))$ (resp. $(x^\rho(\cdot), \pi^\rho(\cdot))$) similarly when $I(X)$ is replaced by $I^\rho(X)$. To derive the first-order necessary condition in terms of small ρ, we let $(\hat{y}(\cdot), \hat{z}(\cdot))$ and $(\hat{x}(\cdot), \hat{\pi}(\cdot))$ be the solutions of the following BSDEs (*variational equations*):

$$\begin{cases} -d\hat{y}(t) = \left[g_y^*(t)\hat{y}(t) + g_z^*(t)\hat{z}(t) \right] dt - \hat{z}(t)dB(t), \\ \quad \hat{y}(T) = U'(W_0 - P_0 - X + I^*(X))(I(X) - I^*(X)), \end{cases} \tag{A2}$$

$$\begin{cases} -d\hat{x}(t) = \left[f_x^*(t)\hat{x}(t) + \ f_\pi^*(t)\hat{\pi}(t) \right] dt - \hat{\pi}(t)dB(t), \\ \quad \hat{x}(T) = -V'(W_0^n + P_0 - (1+\eta)I^*(X))(1+\eta)(I(X) - I^*(X)), \end{cases} \tag{A3}$$

Set

$$\begin{aligned} \widetilde{y}_\rho(t) &= \rho^{-1}[y_\rho(t) - y^*(t)] - \hat{y}(t), \\ \widetilde{z}_\rho(t) &= \rho^{-1}[z_\rho(t) - z^*(t)] - \hat{z}(t). \\ \widetilde{x}_\rho(t) &= \rho^{-1}[x_\rho(t) - x^*(t)] - \hat{x}(t), \\ \widetilde{\pi}_\rho(t) &= \rho^{-1}[\pi_\rho(t) - \pi^*(t)] - \hat{\pi}(t). \end{aligned}$$

The following lemma is essential in our proof of the main result.

Lemma A.2 *Assuming (H1), (H2) and (H3), we have*

$$\lim_{\rho\to 0} \sup_{0\leqslant t\leqslant T} \mathbb{E}_P\left[\widetilde{y}_\rho(t)^2\right] = 0,$$
$$\lim_{\rho\to 0} \mathbb{E}_P\left[\int_0^T \mid \widetilde{z}_\rho(t) \mid^2 dt\right] = 0.$$
$$\lim_{\rho\to 0} \sup_{0\leqslant t\leqslant T} \mathbb{E}_P\left[\widetilde{x}_\rho(t)^2\right] = 0,$$
$$\lim_{\rho\to 0} \mathbb{E}_P\left[\int_0^T \mid \widetilde{\pi}_\rho(t) \mid^2 dt\right] = 0.$$

Proof. We only prove the first two equalities, the other two being similar. Set

$$W^\rho = W_0 - P_0 - X + I^*(X) + \rho(I(X) - I^*(X)),$$
$$W^* = W_0 - P_0 - X + I^*(X).$$

From (10) and (A2), we have

$$\begin{cases} -d\widetilde{y}_\rho(t) & = \rho^{-1}[g(y_\rho(t), z_\rho(t), t) - g(y^*(t), z^*(t), t) - \rho g_y^*(t)\hat{y}(t) \\ & \quad -\rho g_z^*(t)\hat{z}(t)]dt - \widetilde{z}_\rho(t)dB(t), \\ \widetilde{y}_\rho(T) & = \rho^{-1}[V(w_2^\rho(T)) - V(w_2^*(T))] - \hat{y}(T). \end{cases}$$

Let

$$A^\rho(t) = \int_0^1 g_y(y^*(t) + \lambda\rho(\hat{y}(t) + \widetilde{y}_\rho(t)), z^*(t) + \lambda\rho(\hat{z}(t) + \widetilde{z}_\rho(t)), t)d\lambda,$$
$$B^\rho(t) = \int_0^1 g_z(y^*(t) + \lambda\rho(\hat{y}(t) + \widetilde{y}_\rho(t)), z^*(t) + \lambda\rho(\hat{z}(t) + \widetilde{z}_\rho(t)), t)d\lambda,$$
$$C^\rho(t) = [A^\rho(t) - g_y^*(t)]\hat{y}(t) + [B^\rho(t) - g_z^*(t)]\hat{z}(t).$$

Thus

$$\begin{cases} -d\widetilde{y}_\rho(t) & = (A^\rho(t)\widetilde{y}_\rho(t) + B^\rho(t)\widetilde{z}_\rho(t) + C^\rho(t))dt - \widetilde{z}_\rho(t)dB(t), \\ \widetilde{y}_\rho(T) & = \rho^{-1}[V(W^\rho) - V(W^*)] - \hat{y}(T). \end{cases}$$

Using Itô's formula to $\mid \widetilde{y}_\rho(t) \mid^2$ we get

$$\begin{aligned} &\mathbb{E}_P \mid \widetilde{y}_\rho(t) \mid^2 + \mathbb{E}_P \int_t^T \mid \widetilde{z}_\rho(s) \mid^2 ds \\ &= 2\mathbb{E}_P \int_t^T \langle \widetilde{y}_\rho(s), A^\rho(s)\widetilde{y}_\rho(s) + B^\rho(s)\widetilde{z}_\rho(s) + C^\rho(s)\rangle ds + \mathbb{E}_P \mid \widetilde{y}_\rho(T) \mid^2 \\ &\leqslant K\mathbb{E}_P \int_t^T \mid \widetilde{y}_\rho(s) \mid^2 ds + \frac{1}{2}\mathbb{E}_P \int_t^T \mid \widetilde{z}_\rho(s) \mid^2 ds + \mathbb{E}_P \int_t^T \mid C^\rho(s) \mid^2 ds + \mathbb{E}_P \mid \widetilde{y}_\rho(T) \mid^2, \end{aligned}$$

where K is a constant. So

$$\begin{aligned} &\mathbb{E}_P \mid \widetilde{y}_\rho(t) \mid^2 + \frac{1}{2}\mathbb{E}_P \int_t^T \mid \widetilde{z}_\rho(s) \mid^2 ds \\ &\quad \leqslant K\mathbb{E}_P \int_t^T \mid \widetilde{y}_\rho(s) \mid^2 ds + \mathbb{E}_P \int_t^T \mid C^\rho(s) \mid^2 ds + \mathbb{E}_P \mid \widetilde{y}_\rho(T) \mid^2 . \end{aligned}$$

However, the Lebesgue dominated convergence theorem implies

$$\lim_{\rho\to 0}\mathbb{E}_P\int_0^T \mid C^\rho(t)\mid^2 dt = 0,$$
$$\lim_{\rho\to 0}\mathbb{E}_P \mid \widetilde{y}_\rho(T)\mid^2 = 0.$$

Hence, applying Gronwall's inequality, we obtain the result. □

Proof of Theorem 4.1:
The proof is divided into several steps. Let $\mathcal{U} = \{\zeta \in L^2(\Omega, \mathcal{F}_T, P), 0 \leqslant \zeta \leqslant X\}$.

Step 1. (Necessary Condition). We prove that the insurance indemnity $I^*(X)$ is optimal to Problem (A), then there exists a constant $\lambda > 0$ such that

$$\begin{array}{ll} U'(W^*)n(T) - \lambda V'(W^{n,*}(T))(1+\eta)m(T) - \mu \leqslant 0, \text{ a.s.} & \text{on } \mathcal{M}, \\ U'(W^*)n(T) - \lambda V'(W^{n,*})(1+\eta)m(T) - \mu = 0, \text{ a.s.} & \text{on } (\mathcal{M}\cup\mathcal{N})^c, \\ U'(W^*)n(T) - \lambda V'(W^{n,*})(1+\eta)m(T) - \mu \geqslant 0, \text{ a.s.} & \text{on } \mathcal{N} \end{array}$$

where $W^* = W_0 - P_0 - X + I^*(X)$ and $W^{n*} = W_0^n + P_0 - c(I^*(X)) - I^*(X)$, with $x^{I^*(X)}(0) = k$, where $m(T)$ and $n(T)$ are the solutions of the adjoint equations (20) at time T, and where $\mathcal{N}$ and $\mathcal{M}$ are defined respectively as the states of no-insurance and full-insurance,

$$\mathcal{M} \triangleq \{\omega \in \Omega \mid I^*(X) = 0\},$$
$$\mathcal{N} \triangleq \{\omega \in \Omega \mid I^*(X) = X\}$$

For $I^\rho(X) = I^*(X) + \rho(I(X) - I^*(X))$, by Lemma A.1, there exists a constant $\lambda > 0$ such that

$$y_0^{I^\rho(X)} + \lambda(x_0^{I^\rho(X)} - k) - \mu(\mathbb{E}_P[I^\rho(X)] - \Delta) \leqslant y_0^{I^*(X)} + \lambda(x_0^{I^*(X)} - k) - \mu(\mathbb{E}_P[I^*(X)] - \Delta).$$

Dividing the inequality by ρ, sending ρ to 0, and applying Lemma A.2, we obtain

$$\hat{y}_0 + \lambda\hat{x}_0 - \mu\mathbb{E}_P[I(X) - I^*(X)] \leqslant 0 \tag{A4}$$

where $\hat{y}_0$ and $\hat{x}_0$ are the corresponding solutions of (A2) and (A3) at time 0.

Applying Itô's lemma to $n(t)\hat{y}(t) + \lambda m(t)\hat{x}(t)$ yields

$$\begin{aligned} & d[n(t)\hat{y}(t) + \lambda m(t)\hat{x}(t)] \\ = & [\hat{y}(t)g_y^*(t)n(t) - n(t)g_y^*(t)\hat{y}(t) - \langle n(t), g_z^*(t)\hat{z}(t)\rangle + \langle \hat{z}(t), g_z^*(t)n(t)\rangle]dt \\ & +\lambda[\hat{x}(t)f_x^*(t)m(t) - m(t)f_x^*(t)\hat{x}(t) + \langle m(t)f_\pi^*(t), \hat{\pi}(t)\rangle + \langle \hat{\pi}(t), f_\pi^*(t)m(t)\rangle]dt + \{...\}dB(t) \\ = & \{...\}dB(t). \end{aligned}$$

Integrating from 0 to T and taking the expectation, we obtain

$$\begin{aligned} \mathbb{E}_P\left[n(T)U'(W^*)(I(X) - I^*(X)) - \lambda m(T)V'(W^{n,*})(1+\eta)(I(X) - I^*(X))\right] - \mu\mathbb{E}_P[I(X) - I^*(X)] \\ = \hat{y}_0 + \lambda\hat{x}_0 - \mu\mathbb{E}_P[I(X) - I^*(X)] \leqslant 0, \end{aligned} \tag{A5}$$

where (A4) is used in the second line.

We first consider the case that $\mathcal{M}$ is a nonempty set. Since $0 \leqslant I(X) \leqslant X$ is arbitrary, we choose $I(X) = I^*(X)$ on the complementary set of $\mathcal{M}$ and arbitrary over $\mathcal{M}$. Then by (A5) and $I^*(X) = 0$ over $\mathcal{M}$, we obtain:

$$\mathbb{E}_P\left[(n(T)U'(W^*) - \lambda m(T)V'(W^{n,*})(1+\eta) - \mu)(I(X) - I^*(X))\right] \leqslant 0 \tag{A6}$$

As $I(X)$ can be arbitrage non-negative over $\mathcal{M}$, it shows that for each $\varepsilon > 0$

$$P\left\{\omega \mid \omega \in \mathcal{M}, \quad U'(W^*)n(T) - \lambda V'(W^{n,*})(1+\eta)m(T) - \mu > \varepsilon\right\} = 0. \tag{A7}$$

By the continuous property of probability measure, we have

$$U'(W^*(T))n(T) - \lambda V'(W^{n,*}(T))(1+\eta)m(T) - \mu \leqslant 0, \forall \omega \in \mathcal{M}, a.s.. \tag{A8}$$

We now consider the characterization over $\mathcal{N}$. Similarly, we choose $I(X) = I^*(X)$ over the complementary set of $\mathcal{N}$, and arbitrary over $\mathcal{N}$. Then we have

$$\mathbb{E}_P\left[(n(T)U'(W^*) - \lambda m(T)V'(W^{n,*})(1+\eta) - \mu)(I(X) - I^*(X))\right] \leqslant 0. \tag{A9}$$

Since $I(X) - X$ can be arbitrary negative random variable over $\mathcal{N}$, by the same reason as above for $\mathcal{M}$, we show that

$$U'(W^*)n(T) - \lambda V'(W^{n,*})(1+\eta)m(T) - \mu \geqslant 0, \forall \omega \in \mathcal{N}, a.s.. \tag{A10}$$

At last, choose $I(X) = I^*(X)$ over $(\mathcal{M} \cup \mathcal{N})$ and arbitrage over its complementary set. Then

$$\mathbb{E}_P\left[(n(T)U'(W^*) - \lambda m(T)V'(W^{n,*})(1+\eta) - \mu)(I(X) - I^*(X))\right] \leqslant 0. \tag{A11}$$

Since $I(X) - I^*(X)$ can be any non-negative or negative over $(\mathcal{M} \cup \mathcal{N})^c$, we see that

$$U'(W^*)n(T) - \lambda V'(W^{n,*})(1+\eta)m(T) - \mu = 0, \forall \omega \in (\mathcal{M} \cup \mathcal{N})^c, a.s..$$

Therefore, the necessary conditions for $I^*(X)$ have been proved.

Step 2. (Sufficient condition). We prove in this step the conditions in Step 1 are sufficient condition for the optimal insurance contract $I^*(X)$.

Let $I(X) \in \mathcal{U}$ with $(y(\cdot), z(\cdot))$ and $(x(\cdot), \pi(\cdot))$ be the corresponding trajectory. For simplicity, we denote $y_t^{I^*(X)} = y^*(t)$, $x_t^{I^*(X)} = x^*(t)$ and $z_t^{I^*(X)} = z^*(t)$ etc. By Lemma A.1, it suffices to show that for any $I(X) \in \mathcal{U}$,

$$y_0 + \lambda(x_0 - k) - \mu(\mathbb{E}_P[I(X)] - \Delta) \leqslant y_0^* + \lambda(x_0^* - k) - \mu(\mathbb{E}_P[I^*(X)] - \Delta), \tag{A12}$$

or equivalently, to prove

$$y_0 - y_0^* + \lambda(x_0 - x_0^*) - \mu\mathbb{E}_P[I(X) - I^*(X)] \leqslant 0. \tag{A13}$$

Set

$$\begin{aligned}\hat{I}(X) &= I(X) - I^*(X),\\ g_1(y,z,t) &= g(y^*(t)+y, z^*(t)+z, t) - g(y^*(t), z^*(t), t),\\ g_2(y,z,t) &= g_x(y^*(t), z^*(t), t)y + g_z(y^*(t), z^*(t), t)z.\end{aligned}$$

Consider the following equation

$$\begin{cases} -d(y(t)-y^*(t)) &= [g(y(t),z(t),t) - g(y^*(t),z^*(t),t)]dt - (z(t)-z^*(t))'dB(t),\\ &= [g_1(y(t)-y^*(t), z(t)-z^*(t), t)]dt - (z(t)-z^*(t))'dB(t),\\ y(T)-y^*(T) &= U(W) - U(W^*).\end{cases}$$

By Assumption (H3),

$$\begin{aligned} &g_1(y,z,t) \leqslant g_2(y,z,t) \qquad \forall y,z, \quad dP\otimes dt - a.s.\\ &U(W) - U(W^*) \leqslant U'(W^*)\hat{I}(X) \quad a.s.\end{aligned}$$

We now appeal to the comparison theorem for BSDEs (see El Karoui, Peng and Quenez (1997)), obtain, for all t

$$y(t) - y^*(t) \geqslant \hat{y}(t), P - a.s., \tag{A14}$$

where $\hat{y}(\cdot)$ is the solution of (A2). Similar analysis shows that $x(t) - x^*(t) \geqslant \hat{x}(t)$, $\forall t\ P - a.s.$, where $\hat{x}(\cdot)$ is the solution of (A3).

Thus, we have

$$\begin{aligned} & y_0 - y_0^* + \lambda(x_0 - x_0^*) - \mu\mathbb{E}_P[I(X) - I^*(X)]\\ \leqslant\ & \hat{y}_0 + \lambda\hat{x}_0 - \mu\mathbb{E}_P[\hat{I}(X)]\\ \leqslant\ & \mathbb{E}_P\left[(n(T)U'(W^*) - \lambda m(T)V'(W^{n,*})(1+\eta) - \mu)\hat{I}(X)\right]\end{aligned}$$

By similar proofs in the necessary part, those necessary conditions imply that

$$\mathbb{E}_P\left[(n(T)U'(W^*) - \lambda m(T)V'(W^{n,*})(1+\eta) - \mu)\hat{I}(X)\right] \leqslant 0.$$

The sufficient conditions for $I^*(X)$ have been proved.

Step 3.

Define the function

$$h(y) \triangleq U'(W_0 - P_0 + y)n(T) - \lambda V'(W_0^n + P_0 - (1+\eta)(X(\omega)+y))m(T).$$

It is easy to check that $h(y)$ is a monotonic function. Then $H(\lambda,\mu;X)$ is defined uniquely. By step 1 and step 2, the optimal insurance contract is characterized as follows

$$I^*(X) = \min\left\{(X - H(\lambda,\mu;X))^+, X\right\}.$$

Hence the proof is completed. □

Proof of Theorem 3.1: It is a special case of Theorem 4.1, where $\lambda = 0$. Thus, $h(y)$ is degenerated to be

$$h(y) \triangleq U'(W_0 - P_0 + y)n(T)$$

and

$$H(\lambda,\mu;X) = W_0 - P_0 - (U')^{-1}\left(\frac{\mu}{n(T)}\right).$$

It is easy to see that this theorem holds. □

References

Alary, D., C. Gollier, and N. Treich (2010): "Risk, Ambiguity and Insurance," *Working paper presented at the World Risk and Insurance Congress in Singapore.*

Arrow, K. (1971): *Essays in the Theory of Risk Bearing*. Chicago: Markham.

Bernard, C., and W. Tian (2009a): "Optimal Insurance Policies When Insurers Implement Risk Management Metrics," *The Geneva Risk and Insurance Review*, 34, 74–107.

——— (2009b): "Optimal Reinsurance Arrangements Under Tail Risk Measures," *Journal of Risk and Insurance*, 76(3), 709–725.

Bewley, T. F. (2002): "Knightian Decision Theory, Part I," *Decisions in Economics and Finance*, 25, 79–110.

Billot, A., A. Chateauneuf, I. Gilboa, and J.-M. Tallon (2000): "Sharing Beliefs: Between Agreeing and Disagreeing," *Econometrica*, 68, 685–694.

Borch, K. (1962): "Equilibrium in a Reinsurance Market," *Econometrica*, 30(3), 424–44.

Breuer, M. (2005): "Multiple Losses, Ex Ante Moral Hazard and the Implications for Umbrella Policies," *Journal of Risk and Insurance*, 72(4), 525–538.

Cabantous, L. (2007): "Ambiguity Aversion in the Field of Insurance: Insurer's Attitude to Imprecise and Conflicting Probability Estimates," *Theory and Decision*, 62, 219–240.

Carlier, G., and R.-A. Dana (2008): "Two-Persons Efficient Risk-Sharing and Equilibria for Concave Law-Invariant Utilities," *Economic Theory*, 36(2), 189–223.

Carlier, G., R.-A. Dana, and N. Shahidi (2003): "Efficient Insurance Contracts under Epsilon-Contaminated Utilities," *Geneva Papers on Risk and Insurance Theory*, 28, 59–71.

Chateauneuf, A., R.-A. Dana, and J.-M. Tallon (2000): "Optimal Risk Sharing Rules and Equilibria with Choquet expected utilities," *Journal of Mathematical Economics*, 34(2), 191–214.

Chen, Z., and L. Epstein (2002): "Ambiguity, Risk and Asset Returns in Continuous Time," *Econometrica*, 70(4), 1403–1443.

Cooper, R., and B. Hayes (1987): "Multi-Period Insurance Contracts," *International Journal of Industrial Organization*, 5, 211–231.

Crocker, K., and A. Snow (1985): "The Efficiency of Competitive Equilibria in Insurance Markets with Asymmetric Information," *Journal of Public Economics*, 26, 207–219.

Cummins, J., and O. Mahul (2003): "Optimal Insurance with Divergent Beliefs about Insurer Total Default Risk," *The Journal of Risk and Uncertainty*, 27, 121–138.

Cummins, J. D., D. Lalonde, and R. Phillips (2004): "The Basis Risk of Index Linked Catastrophe Loss Securities," *Journal of Financial Economics*, pp. 77–111.

Doherty, N., and H. Schlesinger (1983a): "Optimal Insurance in Incomplete Markets," *Journal of Political Economy*, 91(6), 1045–1054.

——— (1983b): "The Optimal Deductible for an Insurance Policy When Insurance Wealth is Random," *The Journal of Business*, 56, 555–565.

Doherty, N. A., and L. Eeckhoudt (1995): "Optimal Insurance Without Expected Utility: The Dual Theory and the Linearity of Insurance Contracts," *Journal of Risk and Uncertainty*, 10, 157–179.

Duffie, D., and L. Epstein (1992): "Stochastic Differential Utility," *Econometrica*, 60, 353–394.

El Karoui, N., S. Peng, and M. Quenez (1997): "Backward Stochastic Differential Equations in Finance," *Mathematical Finance*, 7, 1–71.

——— (2001): "A Dynamic Maximum Principle for the Optimization of Recursive Utilities under Constraints," *Annals of Applied Probability*, 11(3), 664–694.

Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75(4), 643–669.

Epstein, L., and M. Schneider (2003): "Recursive Multiple-Priors," *Journal of Economic Theory*, 113, 1–33.

Froot, K. (2001): "The Market for Catastrophe Exchange: a Clinic Exam," *Journal of Financial Economics*, pp. 529–571.

Gilboa, I., and D. Schmeidler (1989): "Maximum Expected Utility with a Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

Gollier, C. (1987): "Pareto-Optimal Risk Sharing with Fixed Costs per Claim," *Scandinavian Actuarial Journal*, 13, 62–73.

Gollier, C., and H. Schlesinger (1995): "Second-Best Insurance Contract Design in an Incomplete Market," *Scandinavian Journal of Economics*, 97(1), 123–135.

Gollier, C., and H. Schlesinger (1996): "Arrow's Theorem on the Optimality of Deductibles: A Stochastic Dominance Approach," *Economic Theory*, 7, 359–363.

Golubin, A. (2006): "Pareto-Optimal Insurance Policies in the Models with a Premium Based on the Actuarial Value," *Journal of Risk and Insurance*, 73, 469–487.

Ho, J., L. Keller, and H. Kunreuther (1989): "Risk, Ambiguity and Insurance," *Journal of Risk and Uncertainty*, 2, 5–35.

Hogarth, R., and H. Kunreuther (1985): "Ambiguity and Insurance Decisions," *American Economic Review*, 75, 386–390.

——— (1989): "Risk, Ambiguity and Insurance," *Journal of Risk and Uncertainty*, 2, 5–35.

Ji, S., and X. Zhou (2010): "A Generalized Neyman-Person Lemma for g-Probabilities," *Probability Theory and Related Fields*, 148, 645–669.

Johnson, E., J. Hershey, J. Meszaros, and H. Kunreuther (1993): "Framing, Probability Distortions and Insurance Decisions," *The Journal of Risk and Uncertainty*, 7, 35–51.

Knight, F. (1921): *Risk, Uncertainty and Profit*. Houghton Mifflin Company.

Kunreuther, H., R. Hogarth, and J. Meszaros (1993): "Insurer Ambiguity and Market Failure," *Journal of Risk and Uncertainty*, 7, 71–87.

Landsberger, M., and I. Meilijson (1999): "General Model of Insurance under Adverse Selection," *Economic Theory*, 14, 331–352.

Luenberger, D. (1969): *Optimization by Vector Space Methods*. Wiley, New York.

Ma, J., P. Protter, J. S. Martin, and S. Torres (2002): "Numerical Methods for Backward Stochastic Differential Equations," *Annals of Applied Probability*, 12, 302–316.

Mukerji, S., and J.-M. Tallon (2001): "Ambiguity Aversion and Incompleteness of Financial Markets," *Review of Economic Studies*, 68, 883–904.

Nishimura, K., and H. Ozaki (2004): "Search and Kinghtian Uncertainty," *Journal of Economic Theory*, 119, 299–333.

——— (2007): "Irereversible investment and Kinghtian Uncertainty," *Journal of Economic Theory*, 136, 668–694.

Obstfeld, M. (1994): "Risk-Taking, Global Diversification, and Growth," *American Economic Review*, 84, 1310–1329.

Peng, S. (1997): "BSDE and Related g-Expectation," *edited by N. El Karoui and L. Mazliak Esses, Pitman Research Notes in Mathematics 364, Addison Wesley Longman*, pp. 141–159.

Raviv, A. (1979): "The Design of an Optimal Insurance Policy," *American Economic Review*, 69(1), 84–96.

Rigotti, L., C. Shannon, and T. Strzalecki (2008): "Subjective Beliefs and Ex Ante Trade," *Econometrica*, 76, 1167–1190.

Rothschild, M., and J.E.Stiglitz (1976): "Equilibrium in Competitive Insurance Markerts: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 90(4), 629–649.

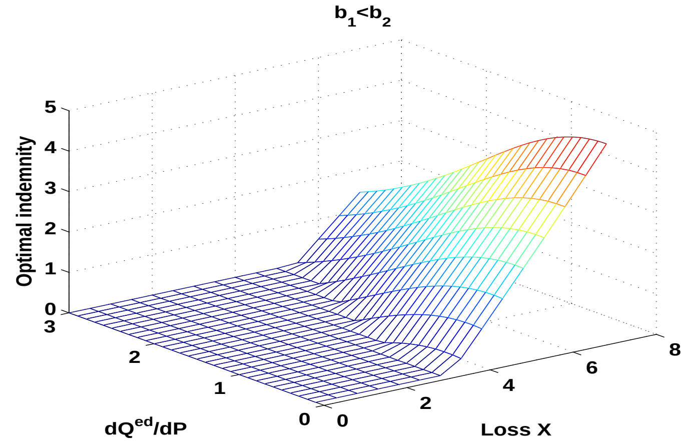
Schmeidler, D. (1989): "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57, 571–587.

Uppal, R., and T. Wang (2003): "Model Misspecification and Under-Diversification," *Journal of Finance*, 58, 2465–2486.

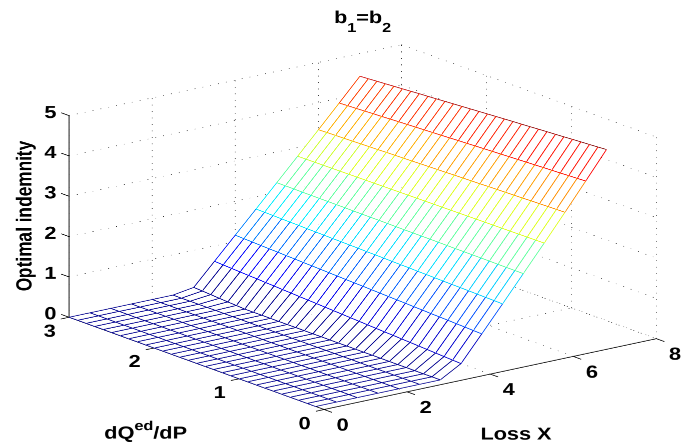
Vazquez, F., and R. Watt (1999): "A Theorem on Optimal Multi-Period Insurance Contracts with Commitment," *Insurance: Mathematics and Economics*, 24, 273–280.

Figure 2: Optimal Insurance Indemnity under Ambiguity

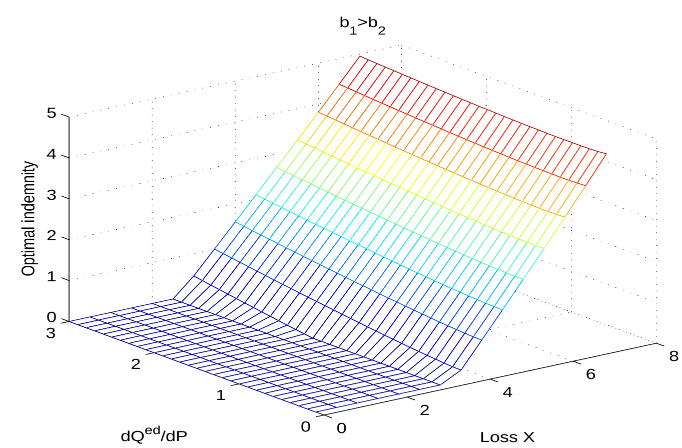
This graph displays the optimal insurance indemnity when both the insured and the insurer have ambiguities. Parameters are $W_0 = 10$, $W_0^n = 10$, $P_0 = 3$, $\eta = 0.1$, $\lambda = 1$, $\mu = 1$, $T = 1$. Panel A plots the indemnity when the insured is less ambiguity than the insurer, $b_1 = 1, b_2 = 2$; Panel B plots the indemnity when both the insured and insurer have the same ambiguity, $b_1 = b_2 = 1$; Panel C plots the indemnity when the insured is more ambiguity than the insured, $b_1 = 1$, $b_2 = 0.5$.



Panel A: $b_1 < b_2$



Panel B: $b_1 = b_2$



Panel C: $b_1 > b_2$