Insurance Contracts with deductibles and upper limits

G. Carlier*, R.A. Dana**

March 2002

Abstract: Insurance policies seldom provide policy holders with complete coverage against losses. Most policies limit coverage with deductibles or include an upper limit. A new explanation for existence of deductibles and upper limits on coverage is given based on multiplicity of agents' priors and cautious behavior. Two models are studied. In the first, both agents have capacities which are epsilon-contaminations of a given prior with different epsilons. It is shown that if the insurer (respectively the insured) takes more into account maximal losses than the insured (respectively the insurer), then the optimal contract has an upper limit (respectively includes a deductible). In the second model, agents are RDEU maximizers. It is proven that if the insurer is very sensitive to the certainty effect (and therefore overweights his maximal loss) while the insured is mildly sensitive, then the optimal contract has an upper limit. If furthermore, the insured disregards events of low probability and neglects his minimal loss, the optimal contract includes a deductible. We lastly combine and compare the two models.

* Grape, Université Bordeaux 4 Montesquieu, Avenue Léon Duguit, 33608 Pessac, France

** Ceremade, Université Paris-Dauphine, Place du Maréchal De Lattre de Tassigny, 75775 Paris cedex 16, France

1 Introduction

In his 1963 seminal insurance paper, Arrow showed that risk-averse agents with von-Neumann Morgenstern utility demand full insurance coverage if insurance is available at actuarially fair price. Yet insurance policies seldom provide policy holders with a complete coverage against losses. Most insurance policies limit coverage of losses with deductibles and include an upper limit on coverage. Many explanations have already been given: transaction costs, economies of scale in administrative expenses, limited liability of policy holders, informationnal asymetries between the insurer and the insured...(see for example Huberman et al (1983) and Winter (1992) and the references included in those papers).

In this paper, we provide a new explanation for deductibles and upper limits on coverage. We assume that the insurer and the insured are not able to derive the probability of the loss and have multiple priors. Being uncertainty averse, agents maximize the minimal expected utility of their wealths over their sets of priors. We show by mean of two examples that ambiguity and uncertainty aversion may explain deductibles and upper limits on coverage.

Since early work on equilibrium asset pricing by Dow and Werlang (1992) and Epstein-Wang (1994), the effect of ambiguity and uncertainty aversion has been discussed in various settings, for example, game theory, asymetric information, equilibrium asset pricing. Risk or uncertainty sharing has also been the subject of a subtantive body of literature on which this paper builds upon. Our main contribution is to show by mean of examples that optimal contracts may have qualitative properties that cannot be obtained if agents are expected utility maximizers with differentiable utilities. Following a stream of research that started with Schmeidler (1989), we assume here that agents are Choquet-expected-utility maximizers with respect to convex capacities. We provide two examples where agents have different sets of priors and optimal contracts include a deductible or an upper limit or both.

Let us first briefly recall the model and the properties of efficient contracts when agents have v.N.M. utilities with differentiable, strictly concave increasing index. A risk averse agent faces a risk W to his initial wealth w_0. The market provides insurance contracts for this risk. A contract is characterized by a premium Π and an indemnity schedule $\bar{I}$ from the state space into $\mathbb{R}_+$. Contracts are such that $0 \leq \bar{I} \leq W$. When the insured buys the contract, he is endowed with the random wealth $w_0 - \Pi - W + \bar{I}$, while by selling the contract the insurer gets Π and promises to pay $\bar{I}(s)$ if state s occurs. Assuming that his initial wealth is zero, his wealth equals $\Pi - \bar{I}$. When agents have v.N.M. utilities with differentiable, strictly concave increasing index, then agents' wealths are comonotone. Hence $\bar{I}$ and $\bar{I} - W$ are non decreasing functions of the loss W that we denote by $I(W)$ and $W - I(W)$. If furthermore, utility index are C^2, then I is a differentiable function whenever it is positive and an efficient contract is such that there exists $D \geq 0$ such that $I(t) = 0$ for $t \leq D$ and $0 < I(t) < t$ and $0 < I'(t) < 1$ for $t \geq D$.

In a finite number of states setting, Machina (1995) expresses the view that these results should be robust to Non expected utility if agents'utilities are smooth functions of the distribution. Safra and Zilcha (1987) express a similar view for utilities compatible with second order stochastic dominance while providing an example of an insurance market with a weakly averse utility and an efficient contract with an upper limit. We partly share Safra and Zilcha's view: when agents have RDEU utilities with strictly convex distortions and strictly concave index, then utilities are compatible with second order stochastic dominance, hence agents' wealths are comonotone and therefore $\bar{I}$ and $\bar{I} - W$ are non decreasing functions of the loss. But we shall show, in an infinite state space setting, that the behavior of the distortion at the boundaries, plays an important role in explaining deductibles and upper limits.

Let $((\Omega, \mathcal{B}))$ be an infinite state space endowed with a σ-field. Assume that agents' set of countably additive priors is of the form $\varepsilon P + (1-\varepsilon)Q$ where Q is "any" probability measure on $(\Omega, \mathcal{B})$ and that agents' utilities are $\varepsilon-$ contamination of an expected utility: $u(X) = \varepsilon E_P(U(X)) + (1-\varepsilon)\inf U(X)$. If one assumes "ex-ante" the existence of "continuous efficient contracts", it can be proven that, necessarely, they are non decreasing function of the loss. In a pure exchange economy setting, this hypothesis is equivalent to the comonotony of Pareto-optimal allocations. Since we believe that existence and continuity of efficient contracts should be a result of the model and not an assumption, we assume in this paper that agents' set of countably additive priors $\mathcal{P}$ is of the form $\varepsilon P + (1-\varepsilon)Q$ where Q is any probability measure on $(\Omega, \mathcal{B})$ absolutely continuous with respect to P and that agents' utilities are of the form:

$$\begin{aligned} u(X) &= \inf_{R \in \mathcal{P}} E_R(U(X)) \\ &= \varepsilon E_P(U(X)) + (1-\varepsilon)\text{essinf}\, U(X) \end{aligned}$$

It may also be checked that $u(X) = E_f(U(X))$ where $E_f(X)$ is the Choquet integral of X with respect to the distortion discontinuous at one $f(x) = \varepsilon x$ for $0 \leq x < 1$ and $f(1) = 1$.

In the first example that we consider, both agents have utilities of that form but with different epsilons and have strictly concave utility index. It is shown that if the insurer takes more into account extreme losses than the insured (the insurer is " more uncertainty averse" than the insured according to Dow and Werlang (1992)), then the optimal contract has an upper limit. This holds true independently of agents' utility index. Agents' utility index however determine the policy for low values of the loss and the level of the upper limit. Let us emphazise that this type of contracts cannot be obtained if agents are expected utility maximizers with differentiable utilities, since efficient contracts are strictly increasing whenever they are strictly positive. On the other hand, if the insured takes more into account maximal losses than the insurer, then the optimal indemnity schedule equals the loss minus a deductible for high values of the loss. This result has some similarity with the case where agents are expected utility maximizers and the insurer is risk neutral and the insured risk averse.

In the second example, agents are Rank Dependent Expected Utility maximizers with strictly convex differntiable distortions and differentiable strictly concave utility index. Much richer qualitative properties than in the expected utility model with differentiable utilities may be obtained. It is shown that if the insurer's distortion has infinite slope at one (in other words if the insurer is very sensitive to the certainty effect and distorts very much very likely events and therefore overweights his maximal loss), while the insured's distortion has finite slope at one, then he offers a contract with an upper limit. Again, the phenomenon holds true independently of agents' utility index but agents' utility index determine the policy for low values of the loss and the level of the upper limit. On the contrary, if the insured's distortion has infinite slope at one, while the insurer's distortion has finite slope at one, then the insured insures himself a constant wealth for high value of the loss by buying a contract with a deductible. Similarly if the insured's distortion has slope zero at zero (respectively the insurer's distortion has slope zero at zero) which means that the insured neglects events of low probability and his maximal outcome, then an optimal contract gives no insurance (respectively full insurance) for low values of the damage. We therefore show that an optimal contract may at the same time have a deductible and an upper limit according to the relative slopes at 0 or 1 of their distortions. Again this type of contracts cannot be obtained if agents are expected utility maximizers with differentiable utilities.

The qualitative properties of the two models display common features. This raises two natural questions that are finally discussed.

- Is the first model the limit of the second?
- Suppose a RDEU agent exchange risk with an agent with an agent with epsilon-contaminated utility. What are the features of an efficient contract?

We show that if the sensitivity to the certainty effect of the RDEU agent is moderate, then we get the same qualitative properties as in the RDEU section. If it is high, assuming that the RDEU agent is the insurer, then he may not trade. If he trades, he offers a contract with a deductible for high values of the risk.

The paper is organised as follows. In section 2, we recall concepts related to convex capacities and provide a few examples of such capacities. In section 3, we introduce an insurance model and provide a class of utility functions for the insurer and the insured for which Pareto optimal contracts exist. We then show that Pareto optimal contracts have properties which are standard when agents are expected utility maximizers: they only depend on the loss and are non decreasing 1-Lipschitz function of the loss. We then get to the central part of the paper: we consider the case where agents have different sets of priors and discuss the qualitative properties of Pareto optimal contracts in two particular cases. Section 4 is devoted to the case where agents are averse to low values of wealths. Both capacities are epsilon-contaminations of a given prior (but the epsilons are different). We show that if the insured takes more into account extreme losses than the insurer, then the optimal indemnity schedule equals the loss minus a deductible for high values of the loss. On the other hand, if the insurer takes more into account extreme losses than the insured, then we show that an optimal contract has an upper limit. In section 5, we consider the case of Rank dependent expected utility maximizers agents (including the case where some agent is an expected utility maximizer) and show that optimal contracts may include a deductible or an upper limit or both if one of the agent distorts very much the probability of very likely events while the other doesn't or distorts very little the probability of very unlikely events while the other doesn't. We end the paper by comparing the two models. We first show that an epsilon contaminated utility may be obtained as the limit of RDEU utilities for well-chosen distortions. The second is that exchange of risk between a RDEU insurer and a insurer with epsilon contaminated utility may also be characterized.

2 Capacities

A capacity on a measurable space $(\Omega, \mathcal{B})$ is a set function $\nu : \mathcal{B} \to [0,1]$ such that $\nu(\emptyset) = 0$, $\nu(\Omega) = 1$, and, for all $A, B \in \mathcal{B}, A \subset B \implies \nu(A) \leq \nu(B)$. A capacity ν is convex if for all $A, B \in \mathcal{B}$, $\nu(A \cup B) + \nu(A \cap B) \geq \nu(A) + \nu(B)$.

Let X be a bounded real-valued $\mathcal{B}$ measurable function. The Choquet integral of X with respect to ν is defined by

$$E_\nu(X) = \int_{-\infty}^{0} (\nu(X \geq t) - 1)dt + \int_{0}^{\infty} \nu(X \geq t)dt$$

We shall make extensive use of the following three examples:

Example 1
Let P be a probability on $(\Omega, \mathcal{B})$, and $\varepsilon \in]0,1[$. The set function defined by $\nu_1(\emptyset) = 0$, $\nu_1(A) =$

$(1-\varepsilon)Q(A), A \neq \Omega,\ \nu_1(\Omega) = 1$ is a convex capacity called the epsilon-contamination of Q. The Choquet Integral with respect to ν_1 is:

$$E_{\nu_1}(X) = (1-\varepsilon)E_Q(X) + \varepsilon \inf_\Omega X$$

Dow and Werlang [1992] define the uncertainty aversion of a capacity ν at $B \in \mathcal{B}$ by

$$e(\nu, B) = 1 - \nu(B) - \nu(B^c)$$

It may easily be checked that in the case of $\nu_1,\ e(\nu_1, B) = \varepsilon$ for all $B, B \neq \Omega,\ B \neq \emptyset$. We shall therefore refer to ε as the uncertainty aversion coefficient of ν_1. Dow and Werlang define an agent with epsilon-contamination ε_1 to be "more uncertainty averse" than another with epsilon-contamination ε_2 if $\varepsilon_1 > \varepsilon_2$.

Example 2
Let P be a probability on $(\Omega, \mathcal{B})$, and $\varepsilon \in]0,1[$. The set function ν_2 defined by $\nu_2(\emptyset) = 0,\ \nu_2(A) = (1-\varepsilon)P(A)$ if $P(A) < 1,\ \nu_2(A) = 1$ if $P(A) = 1$ is a convex capacity. Let $X \in L^\infty(P)$. The Choquet integral with respect to ν_2 is:

$$E_{\nu_2}[X] = (1-\varepsilon)E_P[x] + \varepsilon \text{essinf } X \tag{1}$$

where we recall that

$$\text{essinf } X = \sup\{t \in \mathbb{R} \mid P(X \leq t) = 0\} = \inf\{t \in \mathbb{R} \mid P(X \leq t) > 0\}$$

Indeed to prove (1), by definition of the Choquet integral, assuming that essinf $X \leq 0$ (the case essinf $X \geq 0$ being similar), we have

$$\begin{aligned} E_{\nu_2}[X] &= \int_{-\infty}^{\text{essinf } X} (\nu_2(X \geq t) - 1) + \int_{\text{essinf } X}^{0} (\nu_2(X \geq t) - 1)dt + \int_0^\infty \nu(X \geq t)dt \\ &= \int_{\text{essinf } X}^{0} [(1-\varepsilon)(P(X \geq t) - 1) - \varepsilon]dt + \int_0^\infty (1-\varepsilon)P(X \geq t)dt \\ &= (1-\varepsilon)\ E_P[X] + \varepsilon \text{ essinf } X. \end{aligned}$$

Example 3
Let P be a probability on $(\Omega, \mathcal{B})$. Let $f : [0,1] \to [0,1]$ be increasing, convex and satisfy $f(0) = 0,\ f(1) = 1$. The set function defined by $\nu(A) = f(P(A))$ is a continuous convex capacity. The corresponding Choquet integral is simply denoted $E_f(X)$ (respectively $E_f(U(X))$) and is known as Yaari utility (respectively as Rank dependent expected utility). The differentiability of E_f is discussed in Carlier and Dana [2002].

3 An Insurance Model

Taken as primitive is a probability space $(\Omega, \mathcal{B}, P)$. A risk averse agent faces a risk $W \in L^\infty_+(P)$ with values in $[0, b]$, to his initial wealth w_0. The market provides insurance contracts for this risk. A contract is characterized by a premium Π and an indemnity schedule $\bar{I} : \Omega \to \mathbb{R}_+$. A contract is feasible if $0 \leq \bar{I} \leq W$ P-a.e.. When the insured buys the contract, he is endowed with the random wealth $X_2 = w_0 - \Pi - W + \bar{I}$, while by selling the contract the insurer gets Π and promises to pay $\bar{I}(s)$ if state s occurs. Assuming that his initial wealth is zero, we denote by $X_1 = \Pi - I$ his wealth. A pair of wealths (X_1, X_2) is feasible if $X_1 + X_2 = w_0 - W$ and furthermore $\Pi - W \leq X_1 \leq \Pi$.
We assume that agents' preferences are modeled by monotone utility functions: $u_i : L^\infty(P) \to \mathbb{R}$ that we shall further specify.
Let us first recall a pair of definitions:

Definition 1 *A random variable* X *dominates* Y *in the sense of second order stochastic dominance (respectively strictly dominates) denoted* $X \succeq_2 Y$ *(respectively* $X \succ_2 Y$*) if every* $U : \mathbb{R} \to \mathbb{R}$ *concave increasing,* $E[U(X)] \geq E[U(Y)]$*, (respectively for every* $U : \mathbb{R} \to \mathbb{R}$ *strictly concave increasing,* $E[U(X)] > E[U(Y)]$*).*

Definition 2 *A utility function* $v : L^\infty(P) \to \mathbb{R}$ *is strongly averse (respectively strictly strongly averse) if* $X \succeq_2 Y$ *(respectively* $X \succ_2 Y$*) implies* $v(X) \geq v(Y)$ *(respectively* $v(X) > v(Y)$*).*

Examples and characterizations of strongly risk averse utilities may be found in Chateauneuf et al (1997) and in Chew et al (1995). Strong risk aversion does not imply concavity or quasi-concavity of the utility function.

Definition 3 *A pair* $(X, Y) \in (L^\infty)^2$ *is comonotone if there exists a subset* $B \subset \Omega \times \Omega$ *of measure one for* $P \bigotimes P$ *such that*

$$\left[X(s) - X(s')\right]\left[Y(s) - Y(s')\right] \geq 0, \ \forall (s, s') \in B \times B$$

A pair $(X, Y) \in (L^\infty)^2$ *is anti-comonotone if the pair* $(X, -Y)$ *is comonotone.*

Let $W := X + Y$, $a :=$ essinf W and $b :=$ esssup W. An alternative characterization of comonotonicity of a pair of random variables is as follows:

Lemma 1 *A pair* $(X, Y) \in (L^\infty)^2$ *is comonotone iff there exists a pair of non decreasing continuous function* (h_1, h_2)*,* $h_i : [a, b] \to \mathbb{R}$ *such that* $h_1 + h_2 =$*Id and* $X = h_1(W)$*, a.e. and* $Y = h_2(W)$ *a.e.*

For a proof, see Denneberg (1994).

In this section, we provide a class of utility functions for which Pareto optimal contracts exist. We then show that they have properties which are standard when agents are expected utility maximizers: optimal indemnity schedule only depend on the loss and are non decreasing 1-Lipschitz function of it.

Let us introduce the following assumptions:

U1 u_i are $\sigma(L^\infty, L^1)$ upper semi-continuous, $i = 1, 2$,

U2 u_i, $i = 1, 2$ are strongly averse and some agent is strictly strongly averse.

The following result is an infinite dimension extension of Dana [2000] (see also Chateauneuf, Cohen and Kast [1997] and Landsberger and Meilijson [1994]). A similar proof for a pure exchange economy may be found in Carlier and Dana (2002).

Theorem 1 *Assume that utilities fulfill* **U1***, then we have:*

1. *there exists Pareto-optimal contracts,*
2. *if furthermore utilities fulfill* **U2***, any optimal contract* $\bar{I}^*$ *is such that* $\bar{I}^*$ *and* $W - \bar{I}^*$ *are non decreasing continuous functions of* W*.*

The proof may be found in Carlier and Dana [2002].

Theorem 1 implies that Pareto efficient allocations only depend on W. We shall from now on use capital letters to denote random variables on Ω and small letters to denote functions of W.

More precisely Pareto optimal pairs of wealths (X_1^*, X_2^*) can be written as $(x_1^*(W), x_2^*(W))$ with $x_i^* : [0, b] \to \mathbb{R},\ i = 1, 2$ nondecreasing continuous functions, solutions to the following problems:

$$\begin{cases} \max v_1(x_1) \\ v_2(x_2) \geq u_2^* \\ \Pi - t \leq x_1(t) \leq \Pi, \quad \text{for all } t \in [0, b] \\ x_1(t) + x_2(t) = w_0 - t, \quad \text{for all } t \in [0, b] \end{cases}$$

where $v_i(x_i) := u_i(x_i(W))$ and $u_2^* := u_2(x_2^*(W))$.

In the following sections, we shall further specify utilities and describe the qualitative properties of optimal solutions.

When agents have the same capacity ν, it may be shown that, under regularity conditions on ν, Pareto-optimal contracts coincide with Pareto-optimal contracts of v.N.M. agents with utilities $u_i(X) = E_\pi(U_i(X)),\ i = 1, 2$ for any probability π on $(\Omega, \mathcal{B})$ such that $E_\nu(-W) = E_\pi(-W)$ (See Dana (2001) for an exchange economy version of the result). In order to get richer qualitative behavior, we shall from now on assume that agents have different sets of priors.

4 Epsilon-contaminated capacities

In order to take into account agents' aversion to low values of wealths, we assume that agents' preferences are represented by the utility functions:

$$\begin{cases} u_1(X_1) = [(1 - \varepsilon_1)E_P[U_1(X_1)] + \varepsilon_1 \text{essinf } U_1(X_1)] \\ u_2(X_2) = [(1 - \varepsilon_2)E_P[U_2(X_2)] + \varepsilon_2 \text{essinf } U_2(X_2)] \end{cases}$$

where $U_i : \mathbb{R} \to \mathbb{R},\ i = 1, 2$ is strictly increasing, concave C^1, strictly concave for some i and $\varepsilon_i \in [0, 1[,\ i = 1, 2$. In other words, the value of a random wealth is a weighted average of its expected utility with respect to the prior P and the utility of the worst possible outcome in probability. In the case, $\varepsilon_i = 0$, agent i is an expected utility maximizer.

Proposition 1 *The utility function $u : L^\infty(P) \to \mathbb{R}$ defined by $u(X) = [(1-\varepsilon)E_P[U(X)]+\varepsilon \, essinf\, U(X)]$ with U, strictly increasing and concave is $\sigma(L^\infty, L^1)$ sequentially upper semi-continuous and is strongly averse (respectively strictly strongly averse if U is strictly concave).*

Proof. Let us prove the first assertion. Let X_n converge to $\bar{X}$ in the $\sigma(L^\infty, L^1)$ topology and define λ_n=essinf X_n, $\bar{\lambda}$=essinf $\bar{X}$. Let us show that limsup $\lambda_n \leq \bar{\lambda}$. Assume on the contrary that there exists a subsequence again denoted λ_n which converges to $\lambda > \bar{\lambda}$. Let $\mu \in]\bar{\lambda}, \lambda[$. As $\mu < \bar{\lambda}$, $P(\bar{X} \leq \mu) > 0$ while $\mu < \lambda$ implies $P(X_n \leq \mu) = 0$ for n large enough. In other words, $X_n \geq \mu$ for n large enough a.e. while $P(\bar{X} \leq \mu) > 0$. By the convergence of X_n, we have:

$$\lim_n \int_{\{\bar{X} \leq \mu\}} (X_n - \mu) dP = \int_{\{\bar{X} \leq \mu\}} (\bar{X} - \mu) dP$$

The left handside is non-negative while the right handside is strictly negative. Hence we get a contradiction. Therefore limsup $\lambda_n \leq \bar{\lambda}$.

To prove assertion 2, let us remark that $X \succeq_2 Y$ (respectively $X \succ_2 Y$) implies that $E_P[U(X)] \geq E_P[U(Y)]$ (respectively $E_P[U(X)] > E_P[U(Y)]$) and essinf $X = F_X^{-1}(0) \geq F_Y^{-1}(0)$ =essinf Y ■

Corollary 1 *Assume that agents have epsilon-contaminated utilities. Then there exists optimal contracts, and any optimal indemnity schedule $\bar{I}^*$ is such that $\bar{I}^*$ and $W-\bar{I}^*$ are non decreasing continuous functions of W.*

Since efficient contracts are non decreasing and 1-Lipschitz continuous functions of the loss and since essinf $I^*(W)$= $I^*(\text{essinf } W) = \min_{[0,b]} I^*(t) = I^*(b)$, we are therefore reduced to characterize the solutions to the following problem

$$\begin{cases} \max(1-\varepsilon_1)E_Q[U_1(\Pi - I(t))] + \varepsilon_1 \min_{t\in[0,b]} U_1(\Pi - I(t)) \\ \text{subject to :} \\ (1-\varepsilon_2)E_Q[U_2(w_0 - \Pi - t + I(t))] + \varepsilon_2 \min_{t\in[0,b]} U_2(w_0 - \Pi - t + I(t)) \geq u_2^*, \\ 0 \leq I(t) \leq t \quad \text{for all } t \in [0,b] \end{cases}$$

with $I : [0,b] \to \mathbb{R}^+$ continuous non decreasing and Q being the probability measure dF_W on $([0,b],\mathcal{B})$ and $u_2^* = (1-\varepsilon_2)E_Q[U_2(w_0 - \Pi - t + I^*(t))] + \varepsilon_2 \min_{t\in[0,b]} U_2(w_0 - \Pi - t + I^*(t))$. Hence agents behave as if they had for set of priors the epsilon contamination of the probability measure on the non-negative reals dF_W (see example 1).

We may now characterize optimal wealths or optimal contracts. We show that if the insured takes more into account maximallosses than the insurer (the insured is more uncertainty averse than the insurer according to Dow and Werlang (1992)), then the optimal indemnity schedule equals the loss minus a deductible for high values of the loss. This result has some similarity with the case where agents are expected utility maximizers and the insurer is risk neutral and the insured risk averse. On the contrary, if the insurer takes more into account maximallosses than the insured, then the optimal contract has an upper limit. This phenomenon holds true independently of agents' utility index. Agents' utility index however determine the policy for low values of the loss and the level of the upper limit. This type of contracts cannot be obtained if agents are expected utility maximizers with differentiable utility index.

In the sequel, we shall denote by δ_x the Dirac measure at point x.

Theorem 2 *Assume that $dF_W = (1-p)f_W(x)dx + p\delta_0$, with $p \in [0,1[$ and that $U_i,\ i = 1,2$ is strictly concave, then*

- *Assume that $\varepsilon_2 > \varepsilon_1$. If a contract I is Pareto-optimal, then either $I(t) = 0$ for all t or there exists $D \leq b,\ k \leq b$ such that $I(t) = t - D,\ \forall\, t \in [k,b]$. In other words, an optimal contract includes a deductible for high values of the loss. On $[0,k]$, I is the restriction of an efficient contract of Expected utility maximizers with same probability and utility index $(1-\varepsilon_i)U_i,\ i = 1,2$.*
- *Assume $\varepsilon_1 > \varepsilon_2$. If a contract I is Pareto-optimal, then either $I(t) = t$ for all t or there exists D and k with $D \leq k \leq b$ such that $I(t) = D,\ \forall\, t \in [k,b]$. In other words an optimal contract has an upper limit. On $[0,k]$, I is the restriction of an efficient contract of Expected utility maximizers with same probability and utility index $(1-\varepsilon_i)U_i,\ i = 1,2$.*

The proof may be found in the appendix.

In the case where one agent has a linear index, a more precise statement may be made.

Proposition 2 *Let the insurer have utility $u_1(X_1) = [(1-\varepsilon_1)E_P(X_1) + \varepsilon_1 essinf\, X_1]$ and the insured have utility $u_2(X_2) = [(1-\varepsilon_2)E_P U_2(X_2) + \varepsilon_2 essinf\ U_2(X_2)]$ with $\varepsilon_2 \geq \varepsilon_1$ where $U_2 : \mathbb{R} \to \mathbb{R}$ is strictly concave, strictly increasing and C^1. Then a contract $I : [0,b] \to \mathbb{R}$ is Pareto-optimal iff* *there exists D such that $I = (Id - D)^+$. If $u_1(X_1) = [(1-\varepsilon_1)E_P(U_1(X_1)) + \varepsilon_1 essinf\, U(X_1)]$ and $u_2(X_2) = [(1-\varepsilon_2)E_P(X_2) + \varepsilon_2 essinf\, X_2]$ with $\varepsilon_1 > \varepsilon_2$ where $U_1 : \mathbb{R} \to \mathbb{R}$ is strictly concave, strictly increasing and C^1, then agents do not trade.*

The proof may be found in the appendix.

Let us remark that contracts which include a deductible and an upper limit may be explained by this model. Indeed, assume that the insurer have utility $u_1(X_1) = [(1-\varepsilon_1)E_P(X_1) + \varepsilon_1 \text{essinf } X_1]$ and the insured have utility $u_2(X_2) = [(1-\varepsilon_2)E_P U_2(X_2) + \varepsilon_2 \text{essinf } U_2(X_2)]$ with $\varepsilon_1 > \varepsilon_2$ where $U_2 : \mathbb{R} \to \mathbb{R}$ is strictly concave, strictly increasing and C^1. Then according to Theorem 2, if the contract $I : [0, b] \to \mathbb{R}$ is Pareto-optimal, then for low values of the loss, it is a Pareto-optimal contract of a risk-neutral insurer and of a risk averse agent, hence, it includes a deductible and for high values of the loss, it has an upper limit.

5 The insurer and the insured are RDEU maximizers

The insurance market is as in section two but we now assume that agents' preferences are modelled by RDEU utility functions:

$$\begin{cases} u_1(X_1) = E_{f_1}(U_1(X_1)) \\ u_2(X_2) = E_{f_2}(U_2(X_2)) \end{cases}$$

where $U_i : \mathbb{R} \to \mathbb{R}_+$, $i = 1, 2$ are C^1, and strictly concave increasing, $f_i : [0,1] \to [0,1]$, $i = 1, 2$ satisfy $f_i(0) = 0$, $f_i(1) = 1$ and f_i are increasing, convex C^1. In this section, we assume that $(\Omega, \mathcal{B}, P)$ is non-atomic.

Definition 4 *Let $(\Omega, \mathcal{B}, P)$ be a probability space. A measurable set $A \in \mathcal{B}$ is an atom of P if $P(A) > 0$ and $B \subset A$ implies $P(B) = 0$ or $P(A) = P(B)$. A probability space $(\Omega, \mathcal{B}, P)$ is non-atomic if it has no atoms.*

As in the previous section, we first prove existence of Pareto optimal contracts and that they only depend on the loss. We further prove that an optimal contract is a non decreasing 1-Lipschitz function of the loss:

Proposition 3 *The utility function $u : L^\infty(P) \to \mathbb{R}$ defined by $u(X) = E_f(U(X))$ with U differentiable, strictly increasing and concave fulfills* **U1** *and is strongly averse (respectively strictly strongly averse if U is strictly concave and f strictly convex).*

The proof may be found in Carlier-Dana [2002].

Corollary 2 *Assume that agents are RDEU with convex distortions and concave utility index. Then there exists optimal contracts, and any optimal contract $\bar{I}^*$ is such that $\bar{I}^*$ and $W - \bar{I}^*$ are non decreasing continuous functions of W.*

As in the previous section, we may consider contracts which are non negative functions of the loss and we shall use the same notations.

In order to characterize optimal wealths, let us first recall a few definitions.

Definition 5 *A function $g : L^\infty \to \mathbb{R}$ is Gateaux-differentiable at Y if for all $X \in L^\infty$, the limit*

$$Dg(Y)(X) = \lim_{t \to 0^+} \frac{1}{t}(g(Y + tX) - g(Y))$$

exists and $X \mapsto Dg(Y)(X)$ is a continuous linear form in the norm topology.

As E_f is concave, we recall that the superdifferential $\partial E_f(Y)$ of E_f at $Y \in L^\infty$ is defined by

$$\partial E_f(Y) := \{h \in L^1 \text{ s. t. } E_f(Y) - E_f(Y') \geq \int_\Omega h(Y - Y')dP, \text{ for all } Y' \in L^\infty\}$$

We refer to Carlier and Dana [2002] for details and in particular the reason why the superdifferential is included in L^1.

Let $U : \mathbb{R} \to \mathbb{R}$ be a concave increasing C^1 function and let $E_f(U(Y))$ denote the Choquet integral of $U(Y)$ with respect to the capacity $f(P)$. Let $F_Y : \mathbb{R} \to [0,1]$ denote the distribution function of Y, $F_Y(t) = P(\{Y \leq t\})$. Let F_Y^{-1} denote the generalized inverse of F_Y defined as follows:

$$F_Y^{-1}(t) = \inf\{x \in \mathbb{R} \text{ s.t. } F_Y(x) > t\}, \text{ for } t \in]0,1]$$

$$F_Y^{-1}(0) = \operatorname{essinf} Y.$$

A measurable map $s : \Omega \mapsto [0,1]$ is a *measure-preserving map* if and only if:

$$\int_\Omega \psi(s(\omega))dP(\omega) = \int_0^1 \psi(t)dt, \text{ for all measurable function } \psi$$

for which these integrals are well defined.

Proposition 4 *Let $Y \in L^\infty$, then*

1. *$E_f(U)$ is Gateaux-differentiable at Y if and only if F_Y^{-1} is increasing and we have:*

$$DE_f(U(Y))(Z) = \int_\Omega f'(1 - F_Y(Y))U'(Y)ZdP, \ \forall\, Z \in L^\infty$$

2. *More generally, the superdifferential of $E_f(U(Y))$ is the closed convex hull for the L^1 topology of all $f'(1-s)U'(Y)$ with s measure-preserving and such that $Y = F_Y^{-1} \circ s$.*

3. *Elements of the superdifferential of $E_f(U(Y))$ coincide on the set of ω's such that $P(Y = Y(\omega)) \neq 0$ and are equal to $f'(1 - F_Y(Y(\omega)))$.*

The proof of those statements may be found in Carlier Dana [2002]. We may now characterize efficient contracts. We shall both consider the case of two RDEU agents and the case of a RDEU agent who exchanges with an expected utility maximizer. We shall show that we get much richer qualitative properties than in an expected utility model. What plays an important role in the discussion that follows is whether one of the agent has a distortion with infinite slope at one. The agent is then very sensitive to the certainty effect, he distorts very much the probability of very likely events. Let us show that the agent, in some sense, overweights his maximal loss. Consider the integral of a finite valued random variable which takes values $0 < x_1 < x_2 < \ldots < x_n$ with probabilities $p_1, \ldots, p_n$:

$$E_f(x) = \frac{(1 - f(\sum_{i=2}^n p_i))}{p_1} p_1 x_1 + \frac{f(\sum_{i=2}^n p_i) - f(\sum_{i=3}^n p_i)}{p_2} p_2 x_2 + \ldots + \frac{f(p_n)}{p_n} p_n x_n$$

If $f'(1) = \infty$, then $\frac{1-f(\sum_{i=2}^n p_i)}{p_1} \to \infty$ as $p_1 \to 0$. Hence the worst possible outcome is given much more weight than the other outcomes.

The next two results show that if the insurer is very sensitive to the certainty effect (in other words if $f_1'(1) = \infty$) while the insured is not so ($f_2'(1) < \infty$), then he offers a contract with an upper limit. On the contrary, if the insured is very sensitive to the certainty effect $f_2'(1) = \infty$ while the insurer is more moderate $f_1'(1) < \infty$, then he insures himself a constant wealth for high value of the loss by buying a contract with a deductible. The phenomenon holds true independently of agents' utility index. Agents' utility index determine the policy for low values of the loss and the level of the upper limit.

Proposition 5 *Let $f_i,\ i = 1,2$ be strictly convex and $U_i,\ i = 1,2$ be strictly concave and F_W^{-1} be strictly increasing.*

- *Assume that $f_1'(1) = \infty$ and $f_2'(1) < \infty$. If a contract I is Pareto-optimal, then there exists $D,\ k \leq b$ such that $I(t) = D,\ \forall\ t \in [k, b]$. In other words an optimal contract is constant for high values of the loss.*
- *Assume that $f_2'(1) = \infty$ and $f_1'(1) < \infty$. If a contract $I(t)$ is Pareto-optimal, then there exists $D \leq a,\ k \leq b$ such that $I(t) = t - D,\ \forall\, t \in [k, b]$. In other words an optimal contract includes a deductible for high values of the loss.*

The proof may be found in the appendix.

The result stated above also holds true in the case one agent is an expected utility maximizer:

Proposition 6 *Let $U_i,\ i = 1,2$ be strictly concave and F_W^{-1} be strictly increasing.*

- *let f_1 be strictly convex fulfilling $f_1'(1) = \infty$ and let $f_2(x) = x$. If a contract I is Pareto-optimal, then there exists $D,\ k \leq b$ such that $I(t) = D,\ \forall\ t \in [k, b]$. In other words an optimal contract has an upper limit.*
- *let f_2 be strictly convex fulfilling $f_2'(1) = \infty$ and let $f_1(x) = x$. If a contract I is Pareto-optimal, then there exists $D,\ k \leq b$ such that $I(t) = t - D,\ \forall\ t \in [k, b]$. In other words an optimal contract includes a deductible for high values of the loss.*

Let us next prove that this model allows for coexistence of a deductible for low values of the loss and an upper limit for high values of the loss. The relative slopes of the distortions at zero play an important role in the proof. The condition $f'(0) = 0$ means that the RDEU agent neglects events of low probability and hence doesn't take into account his maximal outcomes (in the discrete case, the maximal outcome is weighted by $f(p_n) = o(p_n)$).

Proposition 7 *Let $f_i,\ i = 1,2$ be strictly convex and $U_i,\ i = 1,2$ be strictly concave and F_W^{-1} be strictly increasing.*

- *Assume that $f_1'(0) > 0$ and $f_2'(0) = 0$. If a contract I is Pareto-optimal, then $I(t) = 0$ for t small enough.*
- *Assume that $f_2'(0) > 0,\ f_1'(0) = 0$. If a contract I is Pareto-optimal, then $I(t) = t$ for t small enough.*

The proof may be found in the appendix.

6 Combining the two models

6.1 Epsilon contamination as a limit case of the RDEU model

The qualitative properties of the two models display common features. This raises two natural questions that we shall discuss in this last section :

- Is it possible to unify these two models in a common framework?
- What are the characteristics of an efficient contract when an RDEU agent exchange risk with an agent with an epsilon-contaminated utility?

As far as the first question is concerned, we show that epsilon contaminated utilities are limits of Choquet expected utilities with respect to convex distortions. This result is only based on the formula $E_f(X) = \int_0^1 f'(1-t)F_X^{-1}dt$ which holds in great generality.

The following convergence result shows that an epsilon-contaminated utility is the limit of a sequence of RDEU utilities.

Let (f_n) be any increasing, strictly convex differentiable sequence of functions from $[0,1]$ into itself such that for all n, $f_n(0) = 0$, $f_n(1) = 1$ and

$$\lim_n f_n(t) = (1-\varepsilon)t, \text{ for all } t \in [0,1).$$

Proposition 8 *For all $X \in L^\infty_+$, we have:*

$$\lim_n E_{f_n}(X) = (1-\varepsilon)EX + \varepsilon essinf\, X.$$

The proof may be found in the appendix.

6.2 A RDEU agent with an agent with epsilon contaminated utility

In this section, we study the case where one agent is RDEU and the other one has an epsilon contaminated utility. As an example, we study only one case, that of a RDEU insurer who exchanges risk with an insured with epsilon contaminated utility. The symmetrical case follows from the same logic. Agents' preferences are then represented by the utility functions:

$$\begin{cases} u_1(X_1) = E_{f_1}[U_1(X_1)] \\ u_2(X_2) = [(1-\varepsilon)E_P[U_2(X_2)] + \varepsilon \text{essinf}\, U_2(X_2)] \end{cases}$$

where $U_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2$ is strictly increasing, concave C^1, strictly concave, f_i is strictly convex increasing and C^1 and $\varepsilon \in]0,1[$.

We show that if the sensitivity to the certainty effect of the RDEU agent is moderate, then we get the same qualitative properties than in the RDEU section. If it is high, than the insurer may refuse to sell a contract. If not, he offers a contract with a deductible for high values of the risk.

We assume that P is non atomic and that F_W^{-1} is strictly increasing.

Proposition 9

- *Assume that $f_1'(1) < \infty$. If a contract $I^*(t)$ is Pareto-optimal, then there exists $D \leq b$, $k \leq b$ such that $I^*(t) = t - D$, $\forall\, t \in [k, b]$. In other words an optimal contract includes a deductible for high values of the loss.*
- *Assume that $f_1'(1) = \infty$. If a contract $I^*(t)$ is Pareto-optimal, then, either $I^*(t) = 0$ or there exists $D \leq b$, $k \leq b$ such that $I^*(t) = t - D$, $\forall\, t \in [k, b]$.*

The proof may be found in the appendix.

7 Appendix: Proofs

Proof of Theorem 2:

Assume that $\varepsilon_2 > \varepsilon_1$ and let $I^* : [0,b] \to \mathbb{R}$ be a Pareto optimal contract and let $x_i^* : [0,b] \to \mathbb{R}$, $i = 1,2$ be the associated wealths, $x_1^*(t) = \Pi - I^*(t)$ and $x_2^*(t) = w_0 - \Pi - t + I^*(t)$. There exists $\lambda \geq 0$ such that I^* maximizes the functional F_λ defined below subject to the constraint $0 \leq I(t) \leq t$

$$F_\lambda(I) = (1-\varepsilon_1)E_Q(U_1(\Pi - I)) + \varepsilon_1 \min_{t\in[0,b]} U_1(\Pi - I)$$

$$+\lambda[(1-\varepsilon_2)E_Q(U_2(w_0 - \Pi - \mathrm{Id} + I)) + \varepsilon_2 \min_{t\in[0,b]} U_2(w_0 - \Pi - \mathrm{Id} + I)]$$

If $\lambda = 0$ then $I = 0$ and there is nothing to prove we therefore assume $\lambda > 0$. First order condition can be written as:

$$(1-\varepsilon_1)U_1'(x_1^*)dF_W + \varepsilon_1 U_1'(x_1^*(b))m_1 - \nu_1 = \lambda[(1-\varepsilon_2)U_2'(x_2^*)dF_W + \varepsilon_2 U_2'(x_2^*(b))m_2] - \nu_2 \quad (2)$$

where m_i is a probability measure for $i = 1,2$ that is supported by $\{t \in [0,b] \text{ s.t. } x_i^*(t) = x_i^*(b)\}$ and ν_i is a nonnegative measure such that ν_1 (respectively ν_2) is supported by $\{I^* = 0\}$ (respectively $\{I^*(t) = t\}$).

If both x_1^* and x_2^* achieve their minimum only at b, then $m_1 = m_2 = \delta_b$. Furthermore, this implies that $0 < I^*(b) < b$ and since I^* is continuous that $0 < I^*(t) < t$ for t close to b. Hence the measures ν_i vanish in a neighbourhood of b. In such a neighbourhood, it follows from (2) that:

$$(1-\varepsilon_1)U_1'(x_1^*(t))dF_W(t) = \lambda[(1-\varepsilon_2)U_2'(x_2^*(t))dF_W(t)] \quad (3)$$

hence

$$(1-\varepsilon_1)U_1'(x_1^*(t)) = \lambda(1-\varepsilon_2)U_2'(x_2^*(t)) \quad (4)$$

By continuity, we obtain

$$(1-\varepsilon_1)U_1'(x_1^*(b)) = \lambda(1-\varepsilon_2)U_2'(x_2^*(b)) \quad (5)$$

Furthermore, it also follows from (2) that

$$\varepsilon_1 U_1'(x_1^*(b)) = \lambda\varepsilon_2 U_2'(x_2^*(b)) \quad (6)$$

Dividing (5) by (6), we obtain $\frac{\varepsilon_1}{1-\varepsilon_1} = \frac{\varepsilon_2}{1-\varepsilon_2}$, hence $\varepsilon_1 = \varepsilon_2$ a contradiction. Hence x_1^* and x_2^* cannot both achieve their minimum only at b. Hence one of those functions has to be constant in a neighborhood of b.

Assume that x_1^* is constant in a neighborhood of b. We denote by $[k,b]$ the largest interval on which $x_1^* = x_1^*(b)$. There are two cases : either $I^*(b) = 0$ so that $I^*(t) = 0$ for all t or $0 < I^*(b) = I^*(k) \leq k$. In that case we have $m_2 = \delta_b$, $\nu_1([k,b]) = 0$ and $\nu_2(]k,b]) = 0$. Then using (2), we first get:

$$\varepsilon_1 U_1'(x_1^*(b)) \geq \varepsilon_1 U_1'(x_1^*(b))m_1(\{b\}) = \lambda\varepsilon_2 U_2'(x_2^*(b)) \quad (7)$$

Similarly we have

$$0 \geq -\nu_2(\{k\}) = \varepsilon_1 U_1'(x_1^*(b))m_1(\{k\})$$

so that $\nu_2(\{k\}) = m_1(\{k\}) = 0$.

Integrating (2) between k and b, we get:

$$\lambda[(1-\varepsilon_2)\int_k^b U_2'(x_2^*(s))dF_W + \varepsilon_2 U_2'(x_2^*)(b)] = [(1-F_W(k))(1-\varepsilon_1) + \varepsilon_1]U_1'(x_1^*(b)) \quad (8)$$

Since U_2' and x_2^* are decreasing, and using (7) we get:

$$\begin{aligned} \lambda(1-\varepsilon_2)\int_k^b U_2'(x_2^*(s))dF_W &< \lambda(1-F_W(k))(1-\varepsilon_2)U_2'(x_2^*(b)) && (9)\\ &\leq (1-F_W(k))(1-\varepsilon_2)\frac{\varepsilon_1}{\varepsilon_2}U_1'(x_1^*(b)) && (10)\end{aligned}$$

then using (8) and (7), we obtain:

$$\begin{aligned}[(1-F_W(k))(1-\varepsilon_1)+\varepsilon_1]U_1'(x_1^*(b)) &< [\tfrac{\varepsilon_1}{\varepsilon_2}(1-\varepsilon_2)(1-F_W(k))+\varepsilon_1]U_1'(x_1^*(b))\\ &= [(1-F_W(k))\tfrac{\varepsilon_1}{\varepsilon_2}+\varepsilon_1 F_W(k)]U_1'(x_1^*(b))\end{aligned}$$

After simple computations, this last inequality yields $\varepsilon_1 > \varepsilon_2$, hence a contradiction and x_1^* cannot be constant in a neighborhood of b. Finally this proves that either I^* is the zero function or x_2^* is constant on some interval $[k,b]$. In that case, it follows from (2) that the restriction of I^* to $[0,k[$ is a Pareto optimal contract of expected utility maximizers with utilities $(1-\varepsilon_i)U_i$ ■

Proof of Proposition 2:

Let $I:[0,b]\to\mathbb{R}$ be a feasible contract and let (x_1,x_2) be the associated wealths. Let us prove that I is dominated by a contract $(\mathrm{Id}-D)^+$ for some D. From Corollary 1, we can assume that $I(W)$ and $W-I(W)$ are non decreasing functions of W. If I is not of the form $(\mathrm{Id}-D)^+$ for some D, let D be such that $v_1(x_1)=v_1(\Pi-(\mathrm{Id}-D)^+)$. Equivalently $E_{Q'}(x_1)=E_{Q'}(\Pi-(\mathrm{Id}-D)^+)$ with $Q'=(1-\varepsilon_1)Q+\varepsilon_1\delta_b$. Let $x_{2D}(t)=w_0-\Pi-t+(t-D)^+=w_0-\Pi-\min(t,D)$. As U_2 is strictly concave,

$$E_{Q'}(U_2(x_{2D})-E_{Q'}(U_2(x_2))>U_2'(w_0-\Pi-D)[(E_{Q'}(x_1)-E_{Q'}(\Pi-(\mathrm{Id}-D)^+))=0 \qquad (11)$$

In other words,

$$(1-\varepsilon_1)E_Q\left(U_2((x_{2D}))+\varepsilon_1U_2((x_{2D}(b))>(1-\varepsilon_1)E_Q\left(U_2(x_2)\right)+\varepsilon_1U_2(x_2(b)\right) \qquad (12)$$

Let us show that $D<b-I(b)$ or equivalently that $x_{2D}(b)>x_2(b)$. Indeed $D\geq b-I(b)$ implies $D\geq t-I(t)$ for all $t\in[0,b]$, hence $\min(t,D)\geq t-I(t)$ for all $t\in[0,b]$ or $x_{2D}(t)\leq x_2(t)$ for all t and $E_{Q'}(U_2(x_{2D}))\leq E_{Q'}(U_2(x_2))$ contradicting (13). Hence $D<b-I(b)$. Therefore $x_{2D}(b)=w_0-\Pi-D>x_2(b)$.

There are two cases: either

$$E_Q\left(U_2(x_{2D})\right)\geq E_Q\left(U_2(x_2)\right). \qquad (13)$$

then obviously we have $v_2(x_{2D})>v_2(x_2)$ which proves that $(\mathrm{Id}-D)^+$ dominates I, or:

$$E_Q\left(U_2(x_{2D})\right)<E_Q\left(U_2(x_2)\right) \qquad (14)$$

$$\begin{aligned}\text{As}\quad \varepsilon_2[U_2(x_{2D}(b))-U_2(X_2(b))] &> \varepsilon_1[U_2(x_{2D}(b))-U_2(X_2(b))]\\ &> (1-\varepsilon_1)[E_Q\left(U_2(X_2)\right)-E_Q\left(U_2(x_{2D})\right)]\\ &> (1-\varepsilon_2)[E_Q\left(U_2(X_2)\right)-E_Q\left(U_2(x_{2D})\right)]\end{aligned}$$

(the first inequality follows from $x_{2D}(b)>x_2(b)$ and $\varepsilon_2\geq\varepsilon_1$ and U_2 increasing, the second from (12)) and the third from (14) and $\varepsilon_2\geq\varepsilon_1$. We therefore get

$$(1-\varepsilon_2)E_Q\left(U_2(x_{2D})\right)+\varepsilon_2U_2(x_{2D}(b))>(1-\varepsilon_2)E_Q\left(U_2(x_2)\right))+\varepsilon_2U_2(x_2(b))$$

which proves that $(\mathrm{Id}-D)^+$ dominates I ■

Proof of Proposition 5:

We prove the result in the case : $f_1'(1)=+\infty$, $f_2'(1)<+\infty$. The other case is obviously obtained by permuting the two agents.

Let $I^*:[0,b]\to\mathbb{R}$ be a Pareto optimal contract and let $x_i^*:[0,b]\to\mathbb{R}$, $i=1,2$ be the associated wealths, $x_1^*(t)=\Pi-I^*(t)$ and $x_2^*(t)=w_0-\Pi-t+I^*(t)$. Since I^* is Pareto-Optimal, there exists $\lambda\geq 0$ such that I^* maximizes

$$F_\lambda(I):=E_{f_1}(U_1(\Pi-I(W)))+\lambda E_{f_2}(U_2(w_0-\Pi+I(W)-W))$$

subject to the constraint $0\leq I\leq W$.

Let us first prove that there exists a neighbourhood of b on which $I^*(t)<t$. If it was not the case, as $t\mapsto I^*(t)-t$ is nonincreasing, one would have $I^*(t)=t$ for all $t\in[0,b]$. Let us therefore show that $I^*\neq\mathrm{Id}$. For $\varepsilon>0$ small enough, let

$$I_\varepsilon(t):=\min(t,b-\varepsilon).$$

We have

$$E_{f_2}(U_2(w_0-\Pi+I_\varepsilon(W)-W))-E_{f_2}U_2(w_0-\Pi)\geq\int_\Omega h(I_\varepsilon(W)-W)U_2'(w_0-\Pi+I_\varepsilon(W)-W)dP \quad (15)$$

for any $h\in\mathrm{core}(f_2(P))$. Since $h\leq f_2'(1)$ and $I_\varepsilon(W)-W\leq 0$, we get

$$\begin{array}{rcl} E_{f_2}(U_2(w_0-\Pi+I_\varepsilon(W)-W))-E_{f_2}U_2(w_0-\Pi) & \geq & \int_{b-\varepsilon}^b f_2'(1)U_2'(w_0-\Pi+b-\varepsilon-t)(b-\varepsilon-t)dF_W(t)\\ & \geq & U_2'(w_0-\Pi-\varepsilon)f_2'(1)\int_{b-\varepsilon}^b(b-\varepsilon-t)dF_W(t) \end{array} \quad (16)$$

Similarly we have

$$E_{f_1}U_1(\Pi-I_\varepsilon(W))-E_{f_1}U_1(\Pi-W)\geq\int_\Omega h(W-I_\varepsilon(W))U_1'(\Pi-I_\varepsilon)dP \quad (17)$$

for any $h\in\partial E_{f_1}(U_1(\Pi-I_\varepsilon))$. Taking $h=f_1'(F_W(W))$ in (17), we obtain

$$\begin{array}{rcl} E_{f_1}(U_1(\Pi-I_\varepsilon(W)))-E_{f_1}U_1(\Pi-W) & \geq & \int_{b-\varepsilon}^b f_1'(F_Wt)U_1'(\Pi-b+\varepsilon)(t-b+\varepsilon)dF_W(t)\\ & \geq & f_1'(F_W(b-\varepsilon))U_1'(\Pi-b+\varepsilon)\int_{b-\varepsilon}^b(t-b+\varepsilon)dF_W(t) \end{array} \quad (18)$$

Since $f_1'(F_W(t))$ tends to $+\infty$ as t tends to b, it follows from (16) and (18) that

$$F_\lambda(I_\varepsilon)-F_\lambda(\mathrm{Id})\geq\left(f_1'(F_W(b-\varepsilon))U_1'(\Pi-b+\varepsilon)-\lambda U_2'(w_0-\Pi-\varepsilon)f_2'(1)\right)\int_{b-\varepsilon}^b(t-b+\varepsilon)dF_W(t)>0$$

for $\varepsilon>0$ small enough. Hence $I^*\neq\mathrm{Id}$.

Let us show now that x_1^* and x_2^* cannot be both strictly decreasing. Since x_i^* strictly decreasing implies that $F_{x_i^*}^{-1}(t)=x_i^*\circ F_W^{-1}(1-t)$ is strictly increasing, we may therefore apply proposition 4. As $F_{x_i^*}(x_i^*)=1-F_W(w)$, the first order conditions are,

$$\lambda U_2'(x_2^*(W(\omega)))f_2'(F_W(x_2^*(W(\omega))))=U_1'((x_1^*(W(\omega)))f_1'(F_W(W(\omega))),\ \forall\omega \quad (19)$$

Equivalently,

$$\lambda U_2'(x_2^*(t))f_2'(F_W(t)) = U_1'(x_1^*(t))f_1'(F_W(t)),\ \forall\ t \in [0,b] \tag{20}$$

Let $t \to b$. Since U_i' is strictly decreasing and x_i^* is decreasing, we have $U_i'(x_i^*(b)) > 0,\ i = 1,2$. As $f_1'(F_W(t)) \to \infty$, $f_2'(F_W(t)) \to \infty$ contradicting the hypothesis. Hence $x_i^*,\ i = 1,2$ cannot be both strictly decreasing. Let us now show that x_1^* is constant in a neighbourhood of b. By proposition 6, the first order conditions are:

$$f_1'(s_1(\omega))U_1'(x_1^*(W(\omega))) = \lambda f_2'(s_2(\omega))U_2'(x_2^*(W(\omega))) \tag{21}$$

with $s_i(\omega) = F_W(W(\omega))$ for all ω which does not belong to a level set of $x_i^*(w)$ with positive measure. If x_1^* is not constant in a neighbourhood of b, then there exists an increasing sequence w_n with limit b and such that for almost every $\omega \in \{W = w_n\}$, $s_1(\omega) = F_W(w_n)$. Hence for every n, we have:

$$\lambda f_2'(s_2(\omega))U_2'(x_2^*(W(\omega))) = U_1'(x_1^*(w_n))f_1'(F_W(w_n)),\ \forall\ \omega \in \{W = w_n\} \tag{22}$$

Passing to the limit, we have a contradiction the right hand side being unbounded while the left hand side is bounded. Hence x_1^* is constant in a neighbourhood of b ∎

Proof of Proposition 6:

The proof of the first case is as in the previous proposition. One first shows that there exists a neighborhood $[b-\varepsilon, b]$ of b such that $I(t) < t$, for all $t \in [b-\varepsilon, b]$. One then shows that x_1^* and x_2^* cannot be both strictly decreasing since the first order conditions would be:

$$\lambda f_1'(F_W(t))U_1'(x_1^*(t)) = U_2'(x_2^*(t)),\ \forall\ t \in [0,b]$$

leading to a contradiction as $t \to b$ since the left hand side tends to ∞ and the right hand side to a finite value. The end of the proof is similar to that of proposition 5 ∎

Proof of Proposition 7:

Let $I^* : [0,b] \to \mathbb{R}$ be a Pareto optimal contract. Let us first show that there does not exist $\alpha > 0$ such that $I^*(t) = t$ on $[0,\alpha]$.

Indeed, assume by contradiction that I^* is Pareto-Optimal and $I^*(t) = t$ on $[0,\alpha]$ for some $\alpha > 0$. Since I^* is Pareto-Optimal, there exists $\lambda \geq 0$ such that I^* maximizes

$$F_\lambda(I) := E_{f_1}(U_1(\Pi - I(W))) + \lambda E_{f_2}(U_2(w_0 - \Pi + I(W) - W))$$

subject to the constraint $0 \leq I \leq W$.

For $\varepsilon \in (0,\alpha)$, let

$$I_\varepsilon(t) := I(t)1_{t \geq \varepsilon}$$

Let $x_{1\varepsilon} = \Pi - I_\varepsilon$ and $x_{2\varepsilon} = w_0 - \Pi + I_\varepsilon - \mathrm{Id}$ be the associated wealths. As $x_{2\varepsilon} = w_0 - \Pi - \mathrm{Id}$ in a neighborhood of 0, it is strictly decreasing in that neighborhood. By proposition 5, the function $U_2'(w_0 - \Pi - W)f_2'(F_W(W)$ is the only element in the restriction of $\partial E_{f_2}(U_2(w_0 - \Pi + I_\varepsilon(W) - W)$ to $\{W \leq \varepsilon\}$ (see Carlier and Dana (2002)). As $I_\varepsilon = I$ for $t \geq \varepsilon$, we thus have:

$$E_{f_2}(U_2(w_0 - \Pi + I_\varepsilon(W) - W)) - E_{f_2}(U_2(w_0 - \Pi + I(W) - W)) \geq -\int_0^\varepsilon tU_2'(-t + w_0 - \Pi)f_2'(F_W(t))dF_W(t)$$

hence

$$E_{f_2}(U_2(w_0 - \Pi + I_\varepsilon(W) - W) - E_{f_2}(U_2(w_0 - \Pi + I(W) - W))) \geq -f_2'(F_W(\varepsilon))\int_0^\varepsilon tU_2'(w_0 - \Pi - t)dF_W(t) \tag{23}$$

On the other hand, let h be any density in $\partial E_{f_1} U_1(\Pi - I_\varepsilon(W))$. By Carlier and Dana [2002] lemma 1, $h \geq f_1'(0)$ a.e.. Hence

$$E_{f_1}(U_1(\Pi - I_\varepsilon(W)) - E_{f_1}(U_1(\Pi - I(W)) \geq U_1'(\Pi) f_1'(0) \int_0^\varepsilon t dF_W(t). \tag{24}$$

It follows that from (23) and (24) that

$$F_\lambda(I_\varepsilon) - F_\lambda(I) \geq \int_0^\varepsilon t \left(U_1'(\Pi) f_1'(0) - \lambda U_2'((w_0 - \Pi - t) f_2'(F_W(\varepsilon)) \right) dF_W(t)$$

Since $f_2'(F_W(\varepsilon)) \to 0$, $U_1'(0) > 0$ and $f_1'(0) > 0$, we obtain that $F(I_\varepsilon) > F_\lambda(I)$ for $\varepsilon > 0$ small enough, hence a contradiction.

If I^* is interior, by proposition 5, the first order conditions are:

$$f_1'(s_1(\omega)) U_1'(x_1^*(W(\omega))) = \lambda f_2'(s_2(\omega)) U_2'(x_2^*(W(\omega))) \tag{25}$$

with $s_i(\omega) = F_W(W(\omega))$ for all ω which does not belong to a level set of $x_i^*(W)$ with positive measure. It follows in particular from the previous paragraph that x_2^* is not constant in a neighborhood of 0. hence there exists an decreasing sequence w_n with limit 0 and such that for almost every $\omega \in \{W = w_n\}$, $s_2(\omega) = F_W(w_n)$. Hence for every n, we have:

$$\lambda f_2'(F_W(w_n)) U_2'(x_2^*(W(\omega))) = U_1'(x_1^*(w_n)) f_1'(s_1(\omega)), \ \forall\, \omega \in \{W = w_n\} \tag{26}$$

Passing to the limit, we have a contradiction the left hand side going to 0 while the right hand side being bounded below. Hence I^* is not interior in a neighborhood of 0. In other words, $I(t) = 0$ for t small enough ■

Proof of Proposition 8:

Let us first recall that given a function f from $[0,1]$ to $\mathbb{R}$, the Legendre-Fenchel Transform of f denoted f^* is the convex function defined for all $x \in \mathbb{R}$ by:

$$f^*(x) = \sup_{t \in [0,1]} xt - f(t).$$

$x \in \mathbb{R}$ is a subgradient of f at $t \in (0,1)$ if and only if $f(t) + f^*(x) = xt$.

Let us also recall that if a sequence of continuous nondecreasing function f_n converges pointwise to a continuous function f on $[0,1)$, then the convergence is uniform on compact subsets of $[0,1)$ (Dini's Theorem).

Let (f_n) be any increasing, strictly convex differentiable sequence of functions from $[0,1]$ into itself such that for all n, $f_n(0) = 0$, $f_n(1) = 1$ and

$$\lim_n f_n(t) = (1-\varepsilon)t, \text{ for all } t \in [0,1).$$

We now proceed to the proof of the proposition. Let us first remark that f_n' converges to $1-\varepsilon$ pointwise on $[0,1[$. Indeed for any $t \in]0,1[$ and h sufficiently small, we have

$$\frac{f_n(t) - f_n(t-h)}{h} \leq f_n'(t) \leq \frac{f_n(t+h) - f_n(t)}{h}$$

Taking the limit, we obtain the pointwise convergence of f_n' converges to $1-\varepsilon$ on $[0,1[$. By Dini's theorem, the convergence is uniform on compact subsets of $[0,1[$.

Let us now define for all n, ε_n such that

$$f_n'(1-\varepsilon_n) = 1.$$

Let us prove that ε_n converges to 0. If not some subsequence again denoted ε_n would converge to $\varepsilon_0 > 0$ and $f_n'(1-\varepsilon_n)$ would converge to $1-\varepsilon$ contradicting the definition of ε_n. Hence ε_n tends to 0.

Let us next prove that

$$\lim_n f_n(1-\varepsilon_n) = 1-\varepsilon. \tag{27}$$

Since, each f_n is increasing, it follows from Dini's Theorem that f_n converges to $t \mapsto (1-\varepsilon)t$ uniformly on compact subsets of $[0,1)$. Therefore, there exists δ_n tending to 0 such that $\delta_n \geq \varepsilon_n$ and $f_n(1-\delta_n)$ tends to $1-\varepsilon$. Since f_n is increasing, we get:

$$\liminf_n f_n(1-\varepsilon_n) \geq 1-\varepsilon. \tag{28}$$

On the other hand, let f_n^* is be the Legendre-Fenchel Transform of f_n. By definition of ε_n, we have:

$$f_n^*(1) := \sup_{0\leq t\leq 1} t - f_n(t) = 1-\varepsilon_n - f_n(1-\varepsilon_n).$$

Hence

$$f_n(1-\varepsilon_n) = 1-\varepsilon_n - f_n^*(1) \tag{29}$$

For all $t \in [0,1)$, we have:

$$\liminf f_n^*(1) \geq \liminf (t - f_n(t)) = \varepsilon t$$

so that $\liminf f_n^*(1) \geq \varepsilon$. By (29), we obtain:

$$\limsup f_n(1-\varepsilon_n) \leq 1-\varepsilon. \tag{30}$$

which proves (27).

To end the proof, recall that: $E_f(X) = \int_0^1 F_X^{-1}(t) f'(1-t)dt$ and that $F_X^{-1}(0) = \text{essinf} X$. Write $E_{f_n}(X) = I_n + J_n$ where

$$I_n := \int_0^{\varepsilon_n} F_X^{-1}(t) f_n'(1-t)dt \text{ and } J_n := \int_{\varepsilon_n}^1 F_X^{-1}(t) f_n'(1-t)dt$$

We first have

$$F_X^{-1}(0)[f_n(1) - f_n(1-\varepsilon_n)] \leq I_n \leq F_X^{-1}(\varepsilon_n)[f_n(1) - f_n(1-\varepsilon_n)]$$

so that with (27) and the convergence of $F_X^{-1}(\varepsilon_n)$ to $F_X^{-1}(0)$, we get

$$\lim_n I_n = \varepsilon F_X^{-1}(0) = \varepsilon \text{essinf} X. \tag{31}$$

Note now by convexity of f, we have for all $t \in [\varepsilon_n, 1]$,

$$0 \leq f_n'(1-t) \leq f_n'(1-\varepsilon_n) = 1$$

As f_n' converges pointwise to $(1-\varepsilon)$ on $[0,1)$, by Lebesgue's Dominated Convergence Theorem, we get:

$$\lim_n J_n = (1-\varepsilon) \int_0^1 F_X^{-1}(t)dt = (1-\varepsilon)EX. \tag{32}$$

The desired result follows from (31) and (32) ■

Proof of Proposition 9:

Since I^* is Pareto-Optimal, there exists $\lambda \geq 0$ such that I^* maximizes

$$F_\lambda(I) := E_{f_1}(U_1(\Pi - I(W))) + \lambda\left((1-\varepsilon)E_P[U_2(w_0 - \Pi + I(W) - W)] + \varepsilon \min_{\Omega} U_2(w_0 - \Pi + I(W) - W)\right)$$

subject to the constraint $0 \leq I \leq W$. The first order condition $0 \in \partial F_\lambda(I^*)$ can be written as the equality of two measures:

$$hU_1'(\Pi - I^*(W))P = \lambda\left(\varepsilon m U_2'(w_0 - \Pi + I^*(b) - b) + (1-\varepsilon)U_2'(w_0 - \Pi + I^*(W) - W)P + \mu + \nu\right) \tag{33}$$

where $h \in$core$f(P)$ is a strictly positive integrable function, m is supported by $\{I^*(W) = I^*(b)\}$, μ is supported by $\{I^*(W) = 0\}$ and ν is supported by $\{I^*(W) = W\}$.

Assume by contradiction that $I^*(t) - t$ is not constant in any neighbourhood of b. Since $I^* - \mathrm{Id}$ is non increasing, $I^*(t) - t < I^*(0) = 0$ for t close to b. Hence there exists a neighbourhood $\mathcal{V}$ of b such that ν vanishes on $\{I^*(W) \in \mathcal{V}\}$ and furthermore m is supported by $\{W = b\}$

Let us now prove that similarly μ vanishes on $I^*(W) \in \mathcal{V}$. If not, we would have, since I^* is nondecreasing, $I^*(t) = 0$ for all $t \in [0, b]$. Let us prove then that $I^* = 0$ cannot be optimal. For $\delta > 0$ small enough, consider the contract

$$I_\delta(t) := \max(0, t - b + \delta)$$

We first have,

$$\min_{t\in[0,b]} U_2(w_0 - \Pi - \mathrm{Id} + I_\delta) - \min_{t\in[0,b]} U_2(w_0 - \Pi - \mathrm{Id}) = U_2(w_0 - \Pi - b + \delta) - U_2(w_0 - \Pi - b) \geq c_0\delta \tag{34}$$

for some constant $c_0 > 0$. Moreover, we have for any $h \in \partial E_{f_1}(U_1(\Pi - I_\delta(W)))$

$$\begin{aligned} E_{f_1}(U_1(\Pi - I_\delta(W))) - E_{f_1}(U_1(\Pi)) &\geq -\int_\Omega I_\delta(W)U_1'(\Pi - I_\delta(W))hdP \\ &\geq -f_1'(1)\int_\Omega I_\delta(W)U_1'(\Pi - I_\delta(W))dP \\ &= -f_1'(1)\int_{b-\delta}^{b} I_\delta(t)U_1'(\Pi - I_\delta(t))dF_W \\ &\geq -c_1\delta f_1'(1)(1 - F_W(b-\delta)) \end{aligned} \tag{35}$$

for some constant $c_1 > 0$. The second inequality follows from the fact that for every $h \in$core$f(P)$, $0 \leq h(\omega) \leq f_1'(1)$ (see Carlier-Dana (2002), lemma 1) and the last inequality $0 \leq I_\delta \leq \delta$. It follows from (34) and (35) that

$$F_\lambda(I_\delta) - F_\lambda(0) \geq \delta(c_0\varepsilon\lambda - c_1 f_1'(1)(1 - F_W(b-\delta)))$$

hence for δ small enough, $F_\lambda(I_\delta) > F_\lambda(0)$ which shows that $I^* = 0$ cannot be optimal and that μ vanishes for t close to b.

Finally, we may rewrite (33) as

$$\alpha P = \beta m$$

for some function α and some constant $\beta > 0$ contradicting the assumption that P is non atomic and that $\widetilde{m}$ is concentrated on the set $\{W = b\}$ which has zero measure since F_W^{-1} is strictly increasing. Hence $I^*(t) - t$ is constant in any neighbourhood of b as was to be proven.
In the case $f_1'(1) = \infty$, the proof is identical except that it may not be proven that $I^* = 0$ cannot be optimal ■

References

Arrow K.J. "Uncertainty and the Welfare of Medical Care", *American Economic Review*, **53** pp 941-973.

Carlier G. and Dana R.A. (2002) "Core of a convex distortion of a probability on a non atomic space"

Chateauneuf A., Cohen M. and Kast M. (1997), "A review of some results related to comonotonicity", Cahier eco-maths 97.32. Université Paris 1.

Chateauneuf A., Cohen M. and Meilijson I. I (1997), "New tools to better model behavior under risk and uncertainty : An overview", Finance 18, 1, pp 25-46.

Chateauneuf A., Dana R.A. and Tallon J.M. (2000), "Risk sharing rules and Equilibria with non-additive expected utilities" *Journal of Mathematical Economics*,**34**, pp.191-214.

Chew S.H., Karni E. and Safra E. (1987), "Risk aversion in the theory of expected utility with Rank dependent probabilities", *Journal of Economic Theory*, **42**, pp.370-381.

Chew S.H. and Mao M.H. (1995) "A Shur Concave Characterization of Risk Aversion for Non Expected Utility Preferences", *Journal of Economic Theory*, **67**, pp.402-435.

Dana R.A. (2001) "On Equilibria When Agents Have Multiple Priors" To appear in *Annals of Operation Research.*

Denneberg D. (1994) "Non-additive Measures and Integral", Kluwer Academic Publishers, Holland.

Dow, J. and Werlang S. (1992), " Uncertainty aversion, risk aversion, and the optimal choice of portofolio", *Econometrica*, **60**, pp.197-204.

Epstein L. and Wang T. (1994), " Intertemporal Asset Pricing under Knightian Uncertainty", *Econometrica*, **62**, pp.283-322.

Hubermann G., Mayers D. and Smith C.W. (1983) *Optimal insurance policy indemnity schedules*, Bell Journal of Economics, 415-426

Landsberger M. and Meilijson I. (1994), "Comonotone allocations, Bickel Lehmann dispersion and the Arrow-Pratt measure of risk aversion ", *Annals of Operation Research*, **52**, pp.97–106.

Machina M. (1995),"Non expected utility and the robustness of the classical insurance paradigm ", in "Non expected utility and Risk management" edited by Gollier C. and Machina M., Kluwers editor.

Safra Z. and Zilcha I. (1988), "Efficient sets with and without the expected utility hypothesis", *Journal of Mathematical Economics*, **17**, pp. 369-284.

Schmeidler D. (1972), "Cores of exact games", *Journal of Mathematical Analysis and applications*, **40**, pp.214–225.

Schmeidler D. (1989), "Subjective Probability and expected utility without additivity", *Econometrica*, **57**, pp.571-587.

Winter R. D. (1992), “Moral Hasard and Insurance Contracts”, Contributions to Insurance Economics, Kluwer Academic Publishers, Holland.