EGRIE Keynote Address

On Optimal Insurance in the Presence of Moral Hazard^{*}

Edi Karni

Department of Economics, Johns Hopkins University, Mergenthaler Hall 469, 3400 N. Charles Street, Baltimore, MD 21218, U.S.A.

E-mail: karni@jhu.edu

This paper expounds on the importance of identifiability of subjective probabilities in agency theory with moral hazard. An application to insurance is examined.
The Geneva Risk and Insurance Review (2008) 33, 1–18. doi:10.1057/grir.2008.7

Keywords: moral hazard; agency theory; subjective probabilities

Introduction

Over the course of the past three decades, advances in the theory of incentive contracts improved our understanding of the theory of optimal insurance in the presence of moral hazard. In particular, when monitoring of the precautions taken to reduce the risk of loss is impractical, the theory highlights the tension between the desire to attain an efficient allocation of risk bearing and the need to create incentives for the insured to take costly measures to reduce the risk of loss. Second-best optimal insurance contracts balance the gain from the reduction in the overall risk against the efficiency loss in the allocation of risk bearing and the cost of implementing the self-protections measures.

By and large, incentive contracts, including insurance policies, are analyzed as a constrained optimization problem. Specifically, the uninformed party (the principal) is supposed to design a contract that maximizes his objective function; is acceptable to the informed party (the agent); and induces the agent, in pursuit of his self-interest, to take the actions desired by the principal. This analysis is based on the assumption that the principal knows the agent's preferences on the set of contracts and actions. It is this assumption that I revisit in this paper. In particular, I argue that, in a world in which new pertinent information becomes available that requires the adjustment of the terms of the contractual agreement, attaining the second-best solution is not enough that the principal knows the agent's prior preferences, it is not even

^{*}This paper is the 19th Geneva risk Economics Lecture delivered to the 34th Seminar of the European Group of Risk and Insurance Economists (EGRIE), 2007 in Cologne, Germany.

enough that he knows that the agent is Bayesian in the sense of updating his prior beliefs using Bayes' rule, the principal must know the agent's "true" prior probabilities or his posterior preferences contingent on any relevant information that may become available.

Until recently, this aspect of the moral hazard problem had not been dealt with in the literature on agency theory.¹ In this paper, I review and discuss these issues and demonstrate, in the context of a simple insurance problem, the consequences of failing to meet these informational exigencies.

In the half century following the publication of Savage's (1954) *The Foundations of Statistics*, subjective expected utility theory provided the framework for the analysis of decision making under uncertainty. The essence of Savage's contribution is an axiomatization of choice behavior that mimics the maximization of the sum of the products of utilities of consequences and the subjective probabilities of events in which these consequences obtain. Savage's work, which builds on ideas and results by Ramsey (1927), de Finetti (1937), and von Neumann and Morgenstern (1944), includes a definition of subjective probabilities on events and utilities on consequences that are derived from the underlying preference structure and the additional convention that the utility of consequences is independent of the underlying events.

As I argue in Karni (2006, 2007b), the convention that the utilities of consequences are state-independent is not implied by Savage's axioms and, as a result, is not falsifiable by observed choice behavior. Put differently, while Savage's axioms require that the preference relation display state independence, they do not imply that the utility function is state independent.² In fact, Savage's model admits state-dependent utility functions that are linear transformations of one another.³ Consequently, the subjective probabilities in Savage's model do not necessarily measure the decision maker's degree of beliefs about the likelihoods of events. This observation is at the heart of the ensuing discussion.

¹ Grant and Karni (2005) studies a principal-agent problem in which misconstrued probabilities and utilities lead the principal to offer the agent a contract that is acceptable yet incentive incompatible. However, unlike in this paper, Grant and Karni assume that the principal does not know the agent's prior preferences (in particular, the agent's evaluation of the disutility of undertaking the effort-demanding action). Moreover, Grant and Karni did not consider the issue of updating the agent's preferences.

² Savage's axioms P3 and P4 captures the "state independence" aspect of the underlying preference relation.

³ This criticism applies to all known revealed-preference theories of choice under uncertainty that invoke Savage's notion of state-space, including the subjective expected utility model of Anscombe and Aumann (1963), as well as nonexpected utility theories, such as probabilistically sophisticated choice of Machina and Schmeidler (1992, 1995), and the maxmin expected utility model of Gilboa and Schmeidler (1989).

The second-best solution to insurance in the presence of moral hazard depends, in part, on the insured's beliefs about the likelihood of potential losses, his attitudes toward risk, and his willingness (or reluctance) to spend the resources necessary to reduce the risk of loss. New information affects the assessments of both the insured and the insurer regarding the likelihood of losses. In particular, if both parties are Bayesians, the change in their beliefs in light of new information is, in a sense, predictable. One is inclined, therefore, to suppose that if the insurer knows the insured's prior preferences and that the insured is Bayesian, then the insurer may ascribe to the insured Savage-type subjective probabilities; design an optimal incentive contract based on the prior beliefs; and, whenever new information arrives, update his own and the insured preferences by the application of Bayes' rule, and design the new contract based on their posterior preferences. This supposition is false. As a matter of fact, it is possible that if Savage-type subjective probabilities do not represent the insured's beliefs, then the posterior preferences ascribed to him by the insurer are false and the incentive contract the insurer believes should induce the insured to make the necessary effort to reduce the risk of loss is incentive incompatible.

The moral of this is that there are situations – and I think they are many – in which Savage's notion of subjective probabilities is inadequate for the study of incentive contracts. This argument also reveals the immensity of the amount of information needed to address the incentive problem posed by moral hazard.

In Karni (2006, 2007a, b) I developed an alternative analytical framework and axiomatic models that are appropriate for the analysis of the principal-agent relationship in the presence of moral hazard. My approach has the additional advantage of admitting situations analogous to state-dependent preferences. For Bayesian decision makers, this theory also pins down a unique family of action-dependent subjective probabilities that represent correctly the decision maker's beliefs.

In this paper, I show, by way of analyzing a simple insurance problem, the pitfalls of misrepresenting the beliefs of Bayesian decision makers and discuss the consequences for the theory of insurance in the presence of moral hazard. The insurance problem is stated in the next section, followed by a numerical analysis in the subsequent section. A discussion of the new approach to modeling decision making under uncertainty in the presence of moral hazard and event-dependent preferences appears in the penultimate section. The final section includes some concluding remarks.

Insurance with moral hazard

Prior beliefs, preferences, and optimal insurance

Consider a simple insurance problem. A consumer seeks to insure his property against theft. The risk of theft depends on the precautionary measures the

consumer takes (e.g., putting on the club when he parks his car, making sure to lock his house). These measures are costly in terms of effort and cannot be monitored by the insurer.

To simplify the discussion, suppose that there are two possible actions corresponding to the two types of behavior: a_0 corresponds to negligent behavior (i.e., failure to take the necessary precautions to protect the property against theft) and a_1 corresponds to diligent behavior (i.e., always taking the necessary measures to protect the property against theft). Let θ denote the event “the insured property is stolen” and denote by θ' the complementary event.

Insurance is provided by an expected profit-maximizing Bayesian insurer whose prior beliefs about the likelihood of theft, conditional on the precautions taken by the consumer, are represented by the prior probability $\pi^I(\theta; a)$, $a \in \{a_1, a_0\}$.

The insurer (principal)–consumer (agent) relationship is governed by an insurance policy (β, α) , where β denotes the net indemnity in the event of a loss due to theft and α denotes the insurance premium. In what follows, I describe the insurance policies that will be offered, taking into account the moral hazard problem and the changes in the terms of this policies if pertinent new information becomes available.

Let the consumer’s initial endowment be (w, w') , where w denotes the consumer wealth in case he sustains a loss due to theft and w' denotes his wealth otherwise. Thus $(w' - w)$ represents the financial loss in the event θ . Let $v : \{a_1, a_0\} \rightarrow \mathbb{R}$ be a function depicting the disutility (i.e., the utility cost) to the agent associated with the two actions, or modes of behavior, and assume, without essential loss of generality, that $v_1 := v(a_1) > v(a_0) = 0$.

Suppose that the consumer is a risk-averse, expected utility-maximizing, Bayesian decision maker whose prior preferences, $\succsim_a^c$, $a \in \{a_1, a_0\}$, on insurance policies contingent on his actions are represented by the utility functions

$$U^c((\beta, \alpha), a) = \gamma(\theta)u(w + \beta)\pi^c(\theta; a) + u(w' - \alpha)\pi^c(\theta'; a) - v(a), \quad (1)$$

where u denotes the consumer’s von Neumann–Morgenstern utility function; $\gamma(\theta)$ measures the, subjective, “psychic” cost to the consumer if theft occurs; and $\pi^c(\cdot; a)$ is the subjective probability distribution representing the consumer’s prior beliefs regarding the likelihoods of θ and θ' conditional on taking the action a .

Assume that the insurer knows the consumer’s prior preferences and that the consumer is Bayesian. Suppose further that the insurer ascribes to the consumer probabilities and utilities as required by Savage’s theory. In other words, the insurer assumes that the consumer’s utility function is event

independent. Then the insurer's perception of the consumer's objective function is, for $a \in \{a_1, a_0\}$,

$$\hat{U}^c((\beta, \alpha), a) = \Gamma(a)[u(w + \beta)\hat{\pi}^c(\theta; a) + u(w' - \alpha)\hat{\pi}^c(\theta'; a)] - v(a), \quad (2)$$

where $\hat{\pi}^c(\theta; a) := \pi^c(\theta; a)\gamma(\theta)/\Gamma(a)$, $\hat{\pi}^c(\theta'; a) := \pi^c(\theta'; a)/\Gamma(a)$ and $\Gamma(a) := \gamma(\theta)\pi^c(\theta; a) + \pi^c(\theta'; a)$.⁴ In other words, the insurer misrepresents the consumer's beliefs and ascribes to him a cost of action consisting of an additive component, $v(\cdot)$, and a multiplicative component, $\Gamma(\cdot)$.⁵ Note, however, that even though the insurer ascribes to the consumer incorrect probabilities and his perception of the disutility of action is incorrect, the insurer's representations of the consumer's preferences, $\succsim_a^c$, $a \in \{a_1, a_0\}$, are, nevertheless, correct.

I assume throughout that the consumer always has the option of declining to take out insurance and bearing the risk associated with theft. If the moral hazard problem is not trivial, it must be the case that, without insurance the consumer will take the action a_1 .⁶ Thus, the consumer's prior reservation utility level is:

$$U^c((0, 0), a_1) = \gamma(\theta)u(w)\pi^c(\theta; a_1) + u(w')\pi^c(\theta'; a_1) - v_1 := \bar{u}^c. \quad (3)$$

Under these assumptions, if the insurer wants to induce the consumer to implement the measures necessary to protect the insured property against theft, his problem may be stated as follows:

Program 1 Choose a policy (β, α) so as to maximize

$$-\beta\pi^I(\theta; a_1) + \pi^I(\theta'; a_1)\alpha \quad (4)$$

subject to the individual rationality constraint

$$\Gamma_1[u(w + \beta)\hat{\pi}^c(\theta; a_1) + u(w' - \alpha)\hat{\pi}^c(\theta'; a_1)] - v_1 \geq \bar{u}^c, \quad (5)$$

and the incentive compatibility constraint

$$u(w + \beta)(\Gamma_1\hat{\pi}^c(\theta; a_1) - \Gamma_0\hat{\pi}^c(\theta; a_0)) + u(w' - \alpha)(\Gamma_1\hat{\pi}^c(\theta'; a_1) - \Gamma_0\hat{\pi}^c(\theta'; a_0)) \geq v_1. \quad (6)$$

⁴ To simplify the notation, henceforth I shall write $\Gamma(a_i) = \Gamma_i$, $i = 0, 1$.

⁵ See Grossman and Hart (1983) for a representation of this type.

⁶ If the consumer does not take the precautionary measures when exposed to the full risk, he will certainly not take these measures when covered by insurance and the moral-hazard issue is trivial.

Even though he ascribes the wrong beliefs to the consumer, the insurer nevertheless employs a representation of the consumer's prior preferences consistent with the consumer's choice behavior. Hence the optimal insurance policy (β^*, α^*) is the same whether the insurer solves Program 1 or the same optimization problem with the constraints representing the consumer's true beliefs and utilities as expressed in (1). Thus, insofar as the insurer is concerned, as long as the insurance policy (β^*, α^*) is in effect, the misspecification of the consumer's utilities and probabilities is of no consequence.

Remark In the literature dealing with principal-agent relationships, it is commonplace to invoke the "common prior" assumption. In the present context, this assumption is open to two interpretations: First, the "true beliefs" of the principal and the agent are in agreement (i.e., $\pi^I(\cdot; a) = \pi^c(\cdot; a)$, $a \in \{a_0, a_1\}$). Second, the principal's perception of the agent's beliefs and his own beliefs agree (i.e., $\pi^I(\cdot; a) = \hat{\pi}^c(\cdot; a)$, $a \in \{a_0, a_1\}$). In the statement of the problem above and in the analysis that follows, the insurer is assigned the role of designing the insurance policies, anticipating consumer's responses. It is natural, therefore, to analyze the problem as seen by the insurer and, consequently, to adopt the second interpretation of the common prior assumption. I assume, henceforth, that the principal's perception of the agent's beliefs agrees with his own beliefs.

Posterior beliefs, preferences, and optimal insurance

Suppose that at the time the original policy is about to expire, new information, in the form of observation x , becomes available. For instance, the number of car thefts in the neighborhood where the consumer lives has been falling. The insurer must propose a new contract taking into account the new information. What are the terms of the new contract?

To begin with, observe that the new information affects the consumer's reservation utility level. Let $\pi^c(\cdot; a, x)$ denote the subjective posterior distribution from the viewpoint of the consumer, obtained by applying Bayes' rule to the prior $\hat{\pi}^c(\cdot; a)$. Denote by $\hat{\pi}^c(\theta; a, x)$ the insurer's perceived subjective posterior distribution of the consumer, obtained by applying Bayes' rule to the prior distribution $\hat{\pi}^c(\cdot; a)$.⁷ Then, the consumer's new reservation

⁷ This is where the depiction of the consumer as Bayesian decision maker and the insurer's awareness of this kick in. Implicitly, there are likelihood functions $q(x|\theta, a)$ that are invoked by the consumer to update the beliefs. These are dealt with explicitly in the example below.

utility level is $\bar{u}^c(x) = \max\{\bar{u}^c(x, a_0), \bar{u}^c(x, a_1)\}$, where

$$\begin{aligned} \bar{u}^c(x, a_i) = & \gamma(\theta)u(w)\pi^c(\theta; a_i, x) \\ & + u(w')\pi^c(\theta'; a_i, x) - v_i, \quad i = 0, 1. \end{aligned} \quad (7)$$

However, the insurer's perception of the consumer's new reservation utility is $\hat{u}^c(x) = \max\{\hat{u}^c(x, a_0), \hat{u}^c(x, a_1)\}$, where

$$\begin{aligned} \hat{u}^c(x, a_i) = & \Gamma_i[u(w)\hat{\pi}^c(\theta; a_i, x) \\ & + u(w')\hat{\pi}^c(\theta'; a_i, x)] - v_i, \quad i = 0, 1. \end{aligned} \quad (8)$$

Suppose that the insurer still seeks to induce the agent to implement a_1 . The insurer's problem may be stated as follows:

Program 2 Choose a policy (β, α) so as to maximize

$$-\beta\pi^I(\theta; a_1, x) + \pi^I(\theta'; a_1, x)\alpha \quad (9)$$

subject to the individual rationality constraint

$$\begin{aligned} \Gamma_1[u(w + \beta)\hat{\pi}^c(\theta; a_1, x) + u(w' - \alpha)\hat{\pi}^c(\theta'; a_1, x)] \\ - v_1 \geq \hat{u}^c(x) \end{aligned} \quad (10)$$

and the incentive compatibility constraint

$$\begin{aligned} u(w + \beta)(\Gamma_1\hat{\pi}^c(\theta; a_1, x) - \Gamma_0\hat{\pi}^c(\theta; a_0, x)) \\ + u(w' - \alpha)(\Gamma_1\hat{\pi}^c(\theta'; a_1, x) - \Gamma_0\hat{\pi}^c(\theta'; a_0, x)) \geq v_1, \end{aligned} \quad (11)$$

where $\pi^I(\cdot; a, x)$ denotes the insurer's posterior probability distribution obtained by the application of Bayes' rule.

Let the solution to Program 2 be denoted by $(\beta^{**}, \alpha^{**})$.⁸

⁸ The solution $(\beta^{**}, \alpha^{**})$, to insurance Program 2, is given by the solution to the following equations:

$$\begin{aligned} u(w + \beta^{**}) \\ = \frac{\hat{u}^c(x)(\Gamma_0\hat{\pi}^c(\theta'; a_0, x) - \Gamma_1(\hat{\pi}^c(\theta'; a_1, x)) + v_1\Gamma_0\hat{\pi}^c(\theta'; a_0, x))}{\Gamma_1\Gamma_0(\hat{\pi}^c(\theta; a_1, x) - \hat{\pi}^c(\theta; a_0, x))} \end{aligned} \quad (12)$$

and

$$\begin{aligned} u(w' - \alpha^{**}) \\ = \frac{\hat{u}^c(x)(\Gamma_1(\hat{\pi}^c(\theta'; a_1, x) - \Gamma_0\hat{\pi}^c(\theta'; a_0, x)) - v_1\Gamma_0\hat{\pi}^c(\theta; a_0, x))}{\Gamma_1\Gamma_0(\hat{\pi}^c(\theta; a_1, x) - \hat{\pi}^c(\theta; a_0, x))}. \end{aligned}$$

(see Salanier, 1997).

The main issue

Consider next the situation from the viewpoint of the consumer. He will accept $(\beta^{**}, \alpha^{**})$ if, under the corresponding best action, its expected utility exceeds his reservation utility. Formally, the policy $(\beta^{**}, \alpha^{**})$ is acceptable to the consumer (that is, it is individually rational) if

$$\begin{aligned} &\gamma(\theta)u(w + \beta^{**})\pi^c(\theta; \bar{a}, x) \\ &+ u(w - \alpha^{**})\pi^c(\theta'; \bar{a}, x) - v(\bar{a}) \geq \bar{u}^c(x), \end{aligned} \quad (13)$$

where $\bar{a} \in \operatorname{argmax}\{a_1, a_0\} \gamma(\theta)u(w + \beta^{**})\pi^c(\theta; a, x) + u(w - \alpha^{**})\pi^c(\theta'; a, x)$.

In view of the desire of the insurer to induce the consumer to implement the action a_1 , the insurance policy $(\beta^{**}, \alpha^{**})$ is incentive compatible if

$$\begin{aligned} &\gamma(\theta)u(w + \beta^{**})(\pi^c(\theta; a_1, x) - \pi^c(\theta; a_0, x)) \\ &+ u(w - \alpha^{**})(\pi^c(\theta'; a_1, x) - \pi^c(\theta'; a_0, x)) \geq v_1. \end{aligned} \quad (14)$$

The main question of concern here is the following: Is it necessarily true that if $(\beta^{**}, \alpha^{**})$ is individually rational it is also incentive compatible? Formally, is it necessarily true that if $(\beta^{**}, \alpha^{**})$ satisfies (13) it must also satisfy (14)?

In the next section I show, by way of a numerical example, that it is possible that the optimal insurance policy $(\beta^{**}, \alpha^{**})$ is individually rational but not incentive compatible. In other words, *failing to ascribe to the agent the correct prior probabilities and utilities, the principal ascribes to the agent's posterior preference relations that are incorrect and, as a result, fails to provide the agent with the incentives that would have induced him to act in the principal's best interest.*

An example

Prior beliefs, preferences, and optimal insurance

The consumer's prior beliefs and preferences

Consider a consumer facing a risk of property loss. Specifically, let the consumer's initial endowment be $(w, w') = (\$2,500; \$10,000)$, so his loss due to accident or theft is \$7,500. Let the consumer's von Neumann–Morgenstern utility functions be $u(w, \theta) = 0.5\sqrt{w}$ and $u(w, \theta') = \sqrt{w}$, and let $v_0 = 0$, $v_1 = 5$.

Suppose that the consumer's joint probability distributions, quantifying his actual beliefs regarding the likelihoods of the events and observations conditional on his actions, are as follows:

	θ	θ'		θ	θ'
x, a_0	0.02	0.68	x, a_1	0.01	0.69
$\neg x, a_0$	0.18	0.12	$\neg x, a_1$	0.09	0.21

Let the consumer's prior beliefs be represented by the marginal distributions on the events, θ and θ' , conditional on his actions. These are summarized as follows:

Actions	a_0	a_1
$\pi^c(\theta, a)$	0.2	0.1

Hence, the consumer's prior preference relations conditional on his actions, $\succsim_a^c, a \in \{a_0, a_1\}$, over the set of insurance policies (β, α) are represented by:

$$U^c((\beta, \alpha); a_0) = 0.2 \times 0.5 \sqrt{w + \beta} + 0.8 \sqrt{w - \alpha}$$

and

$$U^c((\beta, \alpha); a_1) = 0.1 \times 0.5 \sqrt{w + \beta} + 0.9 \sqrt{w - \alpha} - 5$$

Note that $\Gamma_0=0.9, \Gamma_1=0.95$,⁹ and the insurer's perception of the consumer's prior beliefs is $\hat{\pi}^c(\theta, a_0)=(0.5 \times 0.2)/0.9=0.111$, $\hat{\pi}^c(\theta, a_1)=(0.5 \times 0.1)/0.95=0.526$. Hence, *the insurer's perception of the consumer's prior is consistent with the latter actual prior preferences*. Moreover, the insurer ascribes to the consumer's the reservation utility level, $\hat{u}^c=87.5$, which is based on the correct perception that, in the absence of insurance, the consumer would take the necessary measures (i.e., the action a_1) to protect his property.¹⁰

The insurer's prior beliefs and preferences

Suppose that the insurer's joint probability distributions, representing his beliefs regarding the likelihoods of the events and observations conditional on the consumer's actions, are:

	θ	θ'		θ	θ'
x, a_0	0.020	0.680	x, a_1	0.0050	0.6950
$\neg x, a_0$	0.091	0.209	$\neg x, a_1$	0.0476	0.2524

Hence, the insurer's prior beliefs are represented by $\pi^I(\theta, a_0)=0.111$ and $\pi^I(\theta, a_1)=0.526$. Note that $\pi^I(\cdot; a)=\hat{\pi}^c(\cdot; a), a \in \{a_0, a_1\}$. Hence, the prior beliefs of the insurer and those he ascribes to the consumer are in agreement, which is consistent with the aforementioned interpretation of the common prior

⁹ By definition, $\Gamma_0 = 0.5 \times 0.2 + 0.8$ and $\Gamma_1 = 0.5 \times 0.1 + 0.9$.

¹⁰ This is the perceived maximum of expected utility if the consumer forgoes taking out insurance and chooses the action a_1 . Note that, $\hat{u}^c(a_1) = 0.95(0.0526 \times 50 + (1 - 0.0526) \times 100) - 5 = 87.5$, which is larger than $\hat{u}^c(a_0) = 0.9(0.111 \times 50 + (1 - 0.111) \times 100) = 85.0$.

assumption. These prior beliefs are summarized in the following table:

Actions	a_0	a_1
$\pi^I(\theta, a)$	0.111	0.0526
$\hat{\pi}^c(\theta, a)$	0.111	0.0526

The optimal prior insurance policy

Acting in his self-interest, the insurer will design a policy intended to induce the consumer to take measures to protect his property. This policy will be both individually rational (i.e., acceptable to the consumer) and incentive compatible.

If the insurer induces the consumer to implement the action a_0 , the incentive compatibility (henceforth, IC) constraint is nonbinding. Hence the insurer's problem is to design a policy (β, α) that maximizes his expected profit, $-0.111\beta + 0.889\alpha$, subject to the individual rationality (henceforth, IR) constraint

$$0.9[0.111\sqrt{2500 + \beta} + 0.889\sqrt{10000 - \alpha}] = 87.5.$$

Because the insurance is fair (no loading), the probability of loss is common to both parties, and the incentive issue does not arise, the optimal policy must provide full insurance. The solution is $\beta = \$6,952.1$, $\alpha = \$547.88$ implying that the consumer's income is \$9,452.1 regardless of the event. However, because he must meet the IR constraint, the insurer must sustain losses. In fact, the insurer's expected income is: $-0.111 \times 6952.1 + 0.889 \times 547.88 = -\284.62 .

If the insurer were to induce the consumer to implement the action a , his problem is to design a policy that maximizes $-0.0526\beta + 0.9474\alpha$ subject to the IR constraint

$$0.95[0.0526\sqrt{2500 + \beta} + 0.9474\sqrt{10000 - \alpha}] - 5 \geq 87.5$$

and the IC constraint

$$(0.95 \times 0.0526 - 0.9 \times 0.111)\sqrt{2500 + \beta} + (0.95 \times 0.9474 - 0.9 \times 0.889)\sqrt{10000 - \alpha} \geq 5.$$

As both constraints are binding, the optimal solution is given by the solution of the system of equations

$$(0.95 \times 0.0526 - 0.9 \times 0.111)\sqrt{2500 + \beta} + (0.95 \times 0.9474 - 0.9 \times 0.889)\sqrt{10000 - \alpha} = 5$$

$$0.95[0.0526\sqrt{2500 + \beta} + 0.9474\sqrt{10000 - \alpha}] - 5 = 87.5.$$

The solution is $\beta^* = \$6,429.5$, $\alpha^* = \$542.44$. The insurer's expected income is: $-0.0526 \times 6429.5 + 0.9474 \times 542.44 = \175.72 . Thus, the insurer offers the policy (β^*, α^*) anticipating, correctly, that the agent will accept the policy and implement the action a_1 .

Posterior preferences, beliefs, and insurance policy

The consumer's posterior beliefs and preferences given x

Let x be observed. The consumer's likelihood functions are $q^c(x|\theta, a_0)=0.1$, $q^c(x|\theta', a_0)=68/80=0.85$, $q^c(x|\theta, a_1)=0.1$, and $q^c(x|\theta', a_1)=69/90=0.76667$. Since the consumer is Bayesian, his posterior beliefs are represented by the posterior probabilities¹¹

Actions	a_0	a_1
$\pi^c(\theta x, a)$	0.028571	0.014286

The consumer's posterior preference relations, $\succsim_{a,x}^c$, on the set of insurance policies are represented by

$$U^c((\beta, \alpha); a_0, x) = 0.5 \times 0.028571 \times \sqrt{2500 + \beta} \\ + (1 - 0.028571) \times \sqrt{10000 - \alpha}$$

and

$$\bar{U}^c(\times)(\beta, \alpha); a_1, x) = 0.5 \times 0.014286 \times \sqrt{2500 + \beta} \\ + (1 - 0.014286) \times \sqrt{10000 - \alpha} - 5.$$

His reservation utility is $\bar{u}^c(x)=97.857$.¹²

The insurer's perception of the consumer's posterior beliefs and preferences given x

The insurer's likelihoods of x are $q^I(x|\theta, a_0)=(0.020)/0.111=0.18018$, $q^I(x|\theta', a_0)=(0.680)/(0.680 + 0.209)=0.7649$, $q^I(x|\theta, a_1)=(0.0050)/(0.0050 + 0.0476)=0.09506$ and $q^I(x|\theta', a_1)=(0.695)/(0.695 + 0.2524)=0.7336$. Thus, by Bayes' rule, the insurer's beliefs and, by assumption, the beliefs he ascribes to the consumer are represented by the posterior probabilities¹³

Actions	a_0	a_1
$\pi^I(\theta x, a)$	0.029	0.0075
$\hat{\pi}^c(\theta x, a)$	0.029	0.0075

¹¹ $\pi^c(\theta|x, a_0)=(0.1 \times 0.2)/(0.1 \times 0.2 + 0.85 \times 0.8)=0.028571$, $\pi^c(\theta|x, a_1)=(0.1 \times 0.1)/(0.1 \times 0.1 + 0.76667 \times 0.9)=0.014286$.

¹² Note that $\bar{u}^c(x, a_1)=0.5 \times 0.014286 \times 50 + (1 - 0.014286) \times 100 - 5 = 93.929$ and $\bar{u}^c(x, a_0)=0.5 \times 0.028571 \times 50 + (1 - 0.028571) \times 100 = 97.857$. Hence, $\bar{u}^c(x) = \max\{\bar{u}^c(x, a_0), \bar{u}^c(x, a_1)\} = 97.857$.

¹³ $\pi^I(\theta|x, a_0)=(0.18018 \times 0.111)/(0.18018 \times 0.111 + 0.7649 \times (1 - 0.111))=0.029$ and $\pi^I(\theta|x, a_1)=(0.09506 \times 0.0526)/(0.09506 \times 0.0526 + 0.7336 \times (1 - 0.09506))=0.0075$.

Based on these beliefs, the insurer ascribes to the consumer a posterior preference relations, $\succsim_{a,x}^c$, on insurance policies and actions represented by

$$\hat{U}^c((\beta, \alpha); a_0) = 0.9[0.029\sqrt{2500 + \beta} + (1 - 0.029)\sqrt{10000 - \alpha}]$$

and

$$\hat{U}^c((\beta, \alpha); a_1) = 0.95[0.0075\sqrt{2500 + \beta} + (1 - 0.0075)\sqrt{10000 - \alpha}] - 5.$$

The insurer perceives the consumer's reservation utility level to be $\hat{u}^c(x) = 89.664$.¹⁴

The posterior insurance policy given x

If the insurer were to induce the consumer to implement the action a_1 , his problem, as he perceives it, would be the following: Choose (β, α) so as to maximize $-\beta \times 0.0075 + \alpha(1 - 0.0075)$ subject to the IR constraint

$$0.95[0.0075\sqrt{2500 + \beta} + (1 - 0.0075)\sqrt{10000 - \alpha}] - 5 \geq 89.664$$

and the IC constraint

$$(0.95 \times 0.0075 - 0.9 \times 0.029)\sqrt{2500 + \beta} + (0.95 \times 0.9925 - 0.9 \times 0.971)\sqrt{10000 - \alpha} \geq 5.$$

In this case, the two constraints are binding. The optimal insurance policy is: $\beta^{**} = \$7,249.4$, $\alpha^{**} = \$69.28$. The insurer's expected profit is: $-7249.4 \times 0.0075 + (1 - 0.0075) \times 69.28 = \14.39 .

If the insurer were to induce the consumer to implement the action a_0 , then the IC constraint would not be binding and the insurer's problem is to maximize $-0.029\beta + (1 - 0.029)\alpha$ subject to the IR constraint;

$$0.9[0.029\sqrt{2500 + \beta} + (1 - 0.029)\sqrt{10000 - \alpha}] = 89.664.$$

By the same argument, the optimal insurance policy provides for full insurance. The solution to this problem is: $\beta = \$7,425.5$, $\alpha = \$74.461$, and the

¹⁴ Note that $\hat{u}^c(x, a_0) = 0.9 \times (0.029 \times 50 + (1 - 0.029) \times 100) = 88.695$ which is smaller than $\hat{u}^c(x, a_1) = 0.95 \times (0.0075 \times 50 + (1 - 0.0075) \times 100) - 5 = 89.644$. Hence, $\hat{u}^c(x) = \max\{\hat{u}^c(x, a_0), \hat{u}^c(x, a_1)\} = 89.644$.

insurer expected profit is: $-0.029 \times \$7,425.5 + (1-0.029) \times \$74.461 = -\$143.04$. Thus, the insurer is better off with the policy $(\beta^{**}, \alpha^{**})$, believing that it induces the consumer to implement the action a_1 .

The posterior insurance policy is individually rational but incentive incompatible
Consider next the behavior of the consumer facing the contract $\beta^{**} = \$7,249.4$, $\alpha^{**} = \$69.28$. The consumer evaluates the contract according to his true posterior preferences. Thus, his posterior expected utility conditional on implementing a_1 is given by:¹⁵

$$U^c((\beta^{**}, \alpha^{**}); a_1, x) = 93.935$$

and his posterior expected utility conditional on implementing a_0 is given by:¹⁶

$$U^c((\beta^{**}, \alpha^{**}); a_0, x) = 98.216$$

Thus, the consumer is better off implementing the action a_0 . Moreover, since $98.216 = U^c((\beta^{**}, \alpha^{**}), a_0) > \bar{u}^c(x) = 97.857$, the policy $(\beta^{**}, \alpha^{**})$ is individually rational.

Hence, the optimal insurance policy perceived by the insurer is individually rational but incentive incompatible. In other words, the consumer will accept the policy $(\beta^{**}, \alpha^{**})$ but, contrary to the insurer's expectations, will make no effort to protect the insured property.

The explanation for this result is that the insurer erroneously ascribes to the consumer action-dependent probabilities that exaggerate the impact of implementing the action a_1 . In other words, given the observation x , the insurer perceives that the consumer believes that by implementing a_1 he reduces the probability of loss from 2.9 to 0.75 percent, whereas, in fact, the consumer believes that by implementing a_1 he reduces the probability of loss from approximately 2.86 to 1.43 percent. As a result, the insurer underestimates the exposure to risk necessary to produce the incentives that would induce the consumer to implement a_1 . The insurer designs an insurance policy that, once accepted, reduces the consumer exposure to risk to the point at which the cost of taking precautions to avoid the loss covered by the insurance is no longer justified.

¹⁵ $U^c((\beta^{**}, \alpha^{**}), a_1) = 0.5 \times 0.014286 \times \sqrt{2500 + 7249.4} + (1 - 0.014286) \times \sqrt{10000 - 69.280} - 5$.

¹⁶ $U^c((\beta^{**}, \alpha^{**}), a_0) = 0.5 \times 0.028571 \times \sqrt{2500 + 7249.4} + (1 - 0.028571) \times \sqrt{10000 - 69.280}$.

A new approach to modeling decision making under uncertainty

The analytical framework and the representation of the decision maker's prior preferences

In Karni (2006, 2007a, b) I developed the theory underpinning the above discussion. Departing from Savage's (1954) analytical framework, this model is composed of the following concepts:

- A finite set, Θ , of *effects* whose elements are uncertain outcomes. (In the example, these are the events θ and θ' .)
- A set of *actions*, A , representing the means by which the decision maker believes he might influence the likelihood of the effects. (In the example actions correspond to negligent and diligent behavioral modes.) Generally speaking, in addition to affecting the likelihood of the effects, actions may affect the decision maker's well-being directly. For example, taking the necessary precautions to protect one's property may require effort and other costly resources.
- A *bet*, b , is a real-valued function on Θ . (In the example, bets have the interpretation of insurance policies.) Let B denote the set of all bets.

The *choice set* is the product set $\mathbb{C} := A \times B$ whose generic element, (a, b) , an action–bet pair, represents conceivable alternatives among which decision makers may have to choose (e.g., taking out insurance and, simultaneously, deciding on the precautionary measures to protect against loss). A decision maker is characterized by a (prior) preference relation, $\succsim$ on $\mathbb{C}$, that has the usual interpretation and, as usual, the strict preference relation, $\succ$, and the indifference relation, $\sim$, are the asymmetric and symmetric parts of $\succsim$, respectively.

A key idea of this approach is the notion of *constant-valuation bets*. These are bets that, once accepted, leave the decision maker indifferent among all actions. Under a constant valuation bet, the direct utility costs associated with costlier actions are precisely compensated for by the improved chances of winning better outcomes these actions afford.¹⁷ I assume that there exist constant-valuation bets b^{**} and b^* such that $b^{**} \succ b^*$ and, for every $(a, b) \in \mathbb{C}$, there is a constant-valuation bet $\bar{b}$ satisfying $(a, b) \sim \bar{b}$.

Using this analytical framework, Karni (2006) gives an axiomatization of the following representation of the decision maker's prior preferences:

$$(a, b) \mapsto f_a \left(\sum_{\theta \in \Theta} u(b(\theta); \theta) \pi(\theta; a) \right), \quad (15)$$

¹⁷ Formally, let $I(a; b) = \{b' \in B \mid (a, b') \sim (a, b)\}$. A bet $\bar{b}$ is said to be a *constant-valuation bet* if $(a, \bar{b}) \sim (a', \bar{b})$ for all $a, a' \in A$, and $\bigcap_{a \in A} I(a; \bar{b}) = \{\bar{b}\}$.

where u is a continuous utility function on the set of *consequences*, $\mathbb{R} \times \Theta$; $\{\pi(\cdot; a) | a \in A\}$ a family of action-dependent probability measures on the set of effects; and $\{f_a : \mathbb{R} \rightarrow \mathbb{R} | a \in A\}$ is a family of continuous strictly increasing functions. Moreover, u is unique; $\{f_a\}_{a \in A}$ are unique up to a common, strictly increasing, transformation; and, for each $a \in A$, the probability measure $\pi(\cdot; a)$ is unique.

Several aspects of representation (15) are worth mentioning. First, analogously to states of nature in Savage's framework, the effects in this model are resolutions of the uncertainty surrounding the payoffs of the bets. However, unlike in Savage's theory, the factor resolving the uncertainty may be of direct concern to the decision maker. For instance, if the bet is a health insurance policy, the resolution of the uncertainty is the decision maker's state of health. The representation in (15) accommodates health-dependent risk attitudes or, more generally, effect-dependent risk attitudes. Effect-independent preferences constitute a special case of this representation.¹⁸

Second, the functions f_a that figure in the representation capture the cost of action. A specific functional form is the additive separable representation of the form $\sum_{\theta \in \Theta} u(b(\theta); \theta) \pi(\theta; a) + v(a)$, often invoked in the analysis of moral hazard problems.

Third, as in Savage's model, the uniqueness of the probabilities in (15) is predicated on an arbitrary normalization of the functions u and $\{f_a\}_{a \in A}$.¹⁹ Consequently, there is no guarantee that the probabilities, $\pi(\cdot; a)$, correspond to the decision maker's beliefs.²⁰ Additional information is needed to identify the subjective probabilities that represent the decision maker's beliefs. This information may be obtained from the decision maker's posterior preferences.

¹⁸ For details, see Karni (2006).

¹⁹ The normalization consists of assigning b^{**} and b^* utilities as follows: $u(b^*(\theta), \theta) = 0$ for all $\theta \in \Theta$, and $\sum_{\theta \in \Theta} u(b^{**}(\theta), \theta) \pi(\theta; a) = 1$, $a \in A$. Then, for all $a \in A$, $f_a(1) = 1$ and $f_a(0) = 0$.

²⁰ To see this, fix a nonconstant function $\gamma : \Theta \rightarrow \mathbb{R}_{++}$ and let $\Gamma(a) = \sum_{\theta \in \Theta} \gamma(\theta) \pi(\theta; a)$, $\hat{u}(b(\theta); \theta) = u(b(\theta); \theta) / \gamma(\theta)$, $\hat{\pi}(\theta; a) = \gamma(\theta) \pi(\theta; a) / \Gamma(a)$, and $\hat{f}_a = f_a \circ \Gamma(a)$. Then, by (15), the prior preference relation $\succsim$ is represented by

$$\begin{aligned} f_a \left(\Gamma(a) \sum_{\theta \in \Theta} \frac{u(b(\theta); \theta) \gamma(\theta) \pi(\theta; a)}{\gamma(\theta)} \right) \\ = \hat{f}_a \left(\sum_{\theta \in \Theta} \hat{u}(b(\theta); \theta) \hat{\pi}(\theta; a) \right). \end{aligned} \quad (16)$$

But $\hat{\pi}(\theta; a) \neq \pi(\theta; a)$ for some a and θ . Therefore, the prior subjective probabilities are not unique.

Bayesian decision makers

Let X be a set whose elements are *observations*, representing information pertinent to the decision maker's beliefs regarding the likelihood of the alternative effects conditional on his actions. In the example, such information includes reported thefts and other evidence of property crimes in the neighborhood where the insured property is located. Decision makers are supposed to update their prior preferences if new information becomes available. Let $\succsim_x$ denote the posterior preference relation on $\mathbb{C}$ contingent on the observation $x \in X$.

A Bayesian decision maker is characterized by a prior preference relation, a set of posterior preference relations contingent on the observations, and the condition that the posterior preference relations are obtained from the prior preference relation by updating his subjective probabilities using Bayes' rule (i.e., leaving intact the other functions that figure in the representation, such as the functions u and $\{f_a\}_{a \in A}$ in the representation (15)). To model the behavior of Bayesian decision makers, it is necessary to formally incorporate new information into the analytical framework.

Consider a decision maker who, before choosing an action–bet pair, receives information that he deems pertinent to his assessment of the likelihoods of alternative effects conditional on his choice of action. In general, for every $a \in A$, the decision maker is supposed to have action-dependent subjective joint distributions $\pi(\cdot, \cdot | a)$ on $\Theta \times X$. Let $\pi(\theta | a)$, $\pi(x | a)$, and the likelihood functions $q(x | \theta, a)$ be determined from $\pi(\cdot, \cdot | a)$ by marginalization and conditioning.²¹

A decision maker whose prior preference relation $\succsim$ on $\mathbb{C}$ is represented by (15) is said to be *Bayesian* if his posterior preference relations $\{\succsim_x\}_{x \in X}$ are represented by

$$(a, b) \mapsto f_a \left(\sum_{\theta \in \Theta} u(b(\theta); \theta) \pi(\theta | x, a) \right), x \in X, \quad (17)$$

where, for all $x \in X$ and $a \in A$, $\pi(\theta | x, a)$ is determined by the application of Bayes' rule, namely,

$$\pi(\theta | x, a) = \frac{q(x | \theta, a) \pi(\theta | a)}{\sum_{\theta' \in \Theta} q(x | \theta', a) \pi(\theta' | a)}. \quad (18)$$

²¹ That is, for every $a \in A$, $\pi(\theta | a) = \sum_{x \in X} \pi(\theta, x | a)$, $\pi(x | a) = \sum_{\theta \in \Theta} \pi(\theta, x | a)$, and $q(x | \theta, a) = \pi(\theta, x | a) / \pi(\theta | a)$.

In Karni (2007b) I show that, *if the decision maker is Bayesian then the representation of his prior preferences admits a unique set of action-dependent probability measures representing the decision maker's true beliefs.*

Concluding remarks

This analysis shows that predicting the choice of a Bayesian decision maker when new information becomes available requires knowing his true beliefs, not just his prior preferences. Failure to ascribe to the agent his true prior probabilities and utilities, even if the representation of the agent's prior preferences is correct, may result in attributing to him the wrong posterior preferences. In these kind of situations, as the example clearly shows, recontracting under new information may be individually rational and yet incentive incompatible.

The results of this analysis have several implications for the study of insurance markets. First, if the empirical distributions of losses contingent on the precautionary measures taken by the insured is known and is believed, by both the insurer and the insured, to be a correct representation of the risk, the problems raised here may be largely avoided. The probabilities ascribed to the agents are the empirical probabilities, and the utility functions may be effect-dependent.

Second, when the evidence does not permit the construction of a reliable empirical distributions, the insurer does not know the probabilities that represent the insured true beliefs. This means that the set of consumers that purchases a particular policy pools individuals for whom the incentives provided are insufficient to induce them to take the precautionary measures desired by the insurer, thus becoming high-risk types, and individuals for whom the incentive are sufficient, representing low-risk types. The result is that "pooling" contracts implicitly involve cross-subsidization among different risk groups. Moreover, because the insurer lacks the means to control the composition of individuals in the pool, if the insurance market is competitive, there will always be firms that live "on the edge," because they unknowingly draw a "bad pool." Insurers can reduce this risk through reinsurance. In general, in these kinds of situations, the incentives written into the insurance policies are excessive for some individuals. These individuals bear too much risk compared with the second-best risk allocation. By the same logic, other individuals, for whom the incentives are insufficient, bear too little risk relative to the second-best risk allocation.

Third, because the fundamental issue here is the hidden characteristics of the consumer – namely, his beliefs – it seems useful to analyze the situation as a special adverse selection problem. Such analysis, however, is beyond the scope of this paper.

References

- Anscombe, F. and Aumann, R.J. (1963) 'A definition of subjective-probability', *Annals of Mathematical Statistics* **34**: 199–205.
- de Finetti, B. (1937) 'La prévision: Ses lois logiques, ses sources subjectives', *Annals de l'Institut Henri Poincaré* **7**: 1–68. (English translation, by H.E. Kyburg, appears in H.E. Kyburg and H.E. Smokler (eds) (1964) *Studies in Subjective Probabilities*. New York: John Wiley and Sons.).
- Gilboa, I. and Schmeidler, D. (1989) 'Maxmin expected utility with non-unique prior', *Journal of Mathematical Economics* **18**: 141–153.
- Grant, S. and Karni, E. (2005) 'Why does it matter that beliefs and valuations be correctly represented?', *International Economic Review* **46**: 917–934.
- Grossman, S. and Hart, O. (1983) 'An analysis of the principal-agent problem', *Econometrica* **51**: 7–45.
- Karni, E. (2006) 'Subjective expected utility theory without states of the world', *Journal of Mathematical Economics* **42**: 325–342.
- Karni, E. (2007a) 'Agency theory: Choice-based foundations of the parametrized distribution formulation', *Economic Theory*, online publication (doi:10.1007/s00199-007-0277-9).
- Karni, E. (2007b) 'Subjective expected utility theory and the representation of beliefs', unpublished manuscript.
- Machina, M.J. and Schmeidler, D. (1992) 'A more robust definition of subjective probability', *Econometrica* **60**: 745–780.
- Machina, M.J. and Schmeidler, D. (1995) 'Bayes without Bernoulli: simple conditions for probabilistically sophisticated choice', *Journal of Economic Theory* **67**: 106–128.
- Ramsey, F.P. (1927) 'Truth and probability', in R.B. Braithwaite and F. Plumpton (eds) (1931) *The Foundations of Mathematics and Other Logical Essays*, London: K. Paul, Trench, Truber and Co.
- Salanier, B. (1997) *The Economics of Contracts*, Cambridge, MA: MIT Press.
- Savage, L.J. (1954) *The Foundations of Statistics*, New York: John Wiley and Sons.
- von Neumann, J. and Morgenstern, O. (1944) *Games and Economic Behavior*, Princeton, NJ: Princeton University Press.