<table>
<thead>
<tr>
<th>Title</th>
<th>Moral Hazard as a Question of Incentive Compatibility: A Pedagogical Exposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Murota, Takeshi</td>
</tr>
<tr>
<td>Citation</td>
<td>Hitotsubashi Journal of Economics, 31(1): 23-33</td>
</tr>
<tr>
<td>Issue Date</td>
<td>1990-07</td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Text Version</td>
<td>publisher</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://doi.org/10.15057/7825">http://doi.org/10.15057/7825</a></td>
</tr>
</tbody>
</table>
MORAL HAZARD AS A QUESTION OF INCENTIVE COMPATIBILITY
—A PEDAGOGICAL EXPOSITION—

TAKESHI MUROTA

Abstract

This paper is intended to be a simple characterization of moral hazard in a competitive insurance market. While there have been many different interpretations of moral hazard, we assert that its basic nature can be most clearly understood if we formulate it as a question of incentive compatibility in a privacy-preserving, competitive economy under uncertainty. In such a context, it will be shown that a simple scheme of allocating risk-bearings of the Arrow-Debreu type is not free from moral hazard and that some amendment is needed for its prevention. The relation of this problem to the theories of monopoly and of over-insurance will also be discussed.

I. Introduction

In contrast to the common subsidy theory of moral hazard, alternative contributions in this field show somewhat deeper analyses than what the usual competitive models present. E. Helpman and J. Laffont (1975) sets out a temporary equilibrium model where probabilities of future states and dependent on the level of the present consumption of a commodity and discussed the policies of variable insurance premium and of taxation which could remove a possible inefficiency of resource allocation arising from this dependence. M. Pauly (1974) introduces a cost of altering the probability of loss into a simple, static model and analyses such problems as overinsurance and adverse selection. Moreover, J. Marshall (1976), who uses a similar model to Pauly’s, finds an interesting result that inefficiency through moral hazard depends on an insurer’s attitude towards risk and that it vanishes if he is risk-neutral as implied by Helpman and Laffont.¹

These contributions undoubtedly mark a significant leap-forward of the economic theory of moral hazard beyond the scope of its early, informal investigation sparked by Arrow (1968 and 1971), Pauly (1968), and others. However, several important aspects of moral hazard in the daily practices of insurance still seem to be sieved off through their analyses. For example, Pauly and Marshall consider a cost to reduce the probability of

¹ See also Hirshleifer and Riley (1976, pp. 10-11) for a brief survey of the literatures directly or indirectly related to this moral hazard problem.
loss of a risk owner. But imagine a property owner who is making a contract with a fire insurance company. Instead of spending an additional money to reduce the probability of fire in his own property, he may rather increase it by a single strike of a match, which does not incur any substantive \textit{ex ante} cost, if it may result in a heavy, \textit{ex post} punishment. In doing so, he may be able to convert his would-be misfortune of fire into a wind-fall gain if overinsurance prevails. This paper, which is not intended by any means to object any of the results of these authors, aims to be a complement to them by investigating psychological situations something like this example.

Differently from the quoted models which stress the functional dependence of probabilities of future states on the cost for loss prevention or on the commodity endowment at present, we consider a simpler model of an uncertain economy where the alteration of probabilities or just a misrepresentation of them is cost-free without depending on commodity endowment prior to the insurance contract. We, then, demonstrate that the essence of some aspects of moral hazard, which does not seem to have been accurately envisaged thus far, lies in the fact that privacy- or anonymity-preserving mechanism of the competitive allocation of risk-bearing is not individually incentive compatible in the sense of L. Hurwicz (1972) and others. Strangely enough, the theories of moral hazard and of incentive compatibility have experienced parallel developments without a merging point. As a refinement of the sketchy memorandum of T. Murota (1977), this paper intends to fill this gap by considering the scheme of probability manipulation. Our discussion will be useful to understand such behaviors as internal or external arson in fire insurance practices, carelessness in medical care contexts, adverse selection, and so on. A competitive equilibrium origin of overinsurance and of monopolistic behaviors of both insurer's and insured's sides will be singled out in the same context as this.

II. \textit{Competitive Model of Insurance}

Instead of seeking an unnecessary generality by writing out a \(n\)-trader, \(C\)-commodity, and \(S\)-state model, let us consider a private ownership economy \(E\) of two-trader, one-commodity, and two-state given as

\[ E=\{\Omega, \pi^t, \bar{X}^t, U^t; \ i=1, 2\}, \]

where

\( \Omega=\{1, 2\} \) : an index set of two possible states \(s\) of nature commonly conceived by all traders \(i=1, 2\),

\( \pi^t=(\pi_1^t, \pi_2^t) \) : trader \(i\)'s prior, subjective probability distribution on the occurrence of each state \(s\) in \(\Omega\), where \(\pi_s^t\) is the probability that state \(s\) occurs; \(\pi_s^t \geq 0\) for all \(s=1, 2\), and \(\sum_s \pi_s^t = 1\), \(i=1, 2\),

\( \bar{X}^t: \Omega \to X(\bar{X}_1^t, \bar{X}_2^t) \) : a finite, nonnegative valued random variable given on the probability space \((\Omega, \pi^t)\) with \(Pr(\bar{X}^t=\bar{X}_s^t)=\pi_s^t\), where \(\bar{X}_s^t\) is the amount of a single extant commodity which trader \(i\) is to obtain if state \(s\) occurs; for the notational convenience, we write \(\bar{X}^t=\left(\bar{X}_1^t, \bar{X}_2^t\right) \in R^2_+\),

\( U^t: R_+ \to R \) : a von Neuman-Morgenstern utility function of trader \(i\).
As objects of exchange in this economy, we consider what follows.

**DEFINITION I:** One unit of security $s$ is defined as a certificate which entitles its bearer to claim one unit of the commodity if state $s$ occurs and nothing otherwise.

According to this Definition, trader $i$ is understood to initially have $X_{1i}$ and $X_{2i}$ units of the securities 1 and 2, respectively. In a more general context, let $X_s$ denote the amount of security $s$ and $X_i^s$ its amount privately owned by trader $i$. By $p_s$ we denote a parametric premium level of one unit of security $s$.

Suppose that a Walrasian market is organized for the exchange of these securities. When its auctioneer announces parametric premium levels $p = (p_1, p_2) \in \mathbb{R}^+$, each trader $i$ anonymously reports to the auctioneer his net demand $(X_{1i} - X_{1i}^*, X_{2i} - X_{2i}^*)$ for securities 1 and 2 after computing $(X_{1i}^*, X_{2i}^*)$ which is a solution for the problem of maximizing, with respect to $(X_1, X_2) \in \mathbb{R}^2$, his expected utility $\sum_s \pi_s U(X_s)$ subject to his budget constraint $\sum_s p_s X_s = \sum_s p_s X_i^s$; These reports can be anonymous. The auctioneer, who receives these reports, computes the sum $\sum_s (X_i^s - X_{1i})$ for each $s = 1, 2$. If such a sum of excess demand turns out to be positive (or negative) for some $s$, then the auctioneer raises (or lowers) the premium $p_s$ and reannounces a newly raised (or lowered, respectively) level. Such a process of premium adjustment, which may be called the Walrasian tatonnement, goes on until those excess demands for securities all vanish. An equilibrium of such a process is formulated in what follows.

**DEFINITION II:** Given the economy $E$, a set $X^* = (X_{11}^*, X_{12}^*, X_{21}^*, X_{22}^*)$ of security holdings of two traders is defined as an equilibrium allocation of securities through competitive insurance if and only if there exists a premium vector $p^* = (p_1^*, p_2^*) \in \mathbb{R}^2_+$ of securities 1 and 2 such that each $X_i^s = (X_{1i}^s, X_{2i}^s)$ in $X^*$ is a solution for the problem;

Given $p^*$, maximize, with respect to $(X_1, X_2) \in \mathbb{R}^2_+$, the expected utility $\sum_s \pi_s U(X_s)$ subject to the budget constraint $\sum_s p_s X_s = \sum_s p_s X_i^s$; $i = 1, 2$ and the following condition of feasibility;

$$\sum_i X_i^s = \sum_i X_{1i}^s; \quad s = 1, 2$$

are met.\(^3\)

When this equilibrium is found, trader $i$ purchases or sells $(X_{1i}^s - X_{1i}^*)$ units of securities with their premiums $p_i^s; s = 1, 2$ and he enjoys the expected utility $\sum_s \pi_s U(X_i^s)$. Provided that all traders are risk-averse in the sense that for $0 < t < 1$ and $x \neq y, tU(z) + (1 - t)U(y) < U(tx + (1 - t)y); i = 1, 2$ and some other classical conditions are satisfied in our economy.

---

\(^3\) Symbols, $R$, $R_+$, and $R_S^S$ signify the set of real numbers, the set of nonnegative real numbers and the $S$-dimensional Cartesian product of $R$, respectively.

it is easy to find that the allocation of securities reached in this competitive fashion is Pareto optimal, that is to say, no trader can increase his expected utility by moving away to any allocation other than this competitive allocation without lowering other trader's expected utility. For our future convenience, we introduce

DEFINITION III: If there is some trader, say j, whose initial endowment of security is invariant under different states in that \( X_{ij} = X_{j*} \), then he is called an insurer. Trader, who is not an insurer, is called an insured.

III. Misrepresented Probabilities and a Monopoly

Due to the anonymity-preserving nature of competitive insurance described above, traders may not necessarily represent their true subjective probabilities in their computation of optimal security holdings. In order to investigate the effect of misrepresenting probabilities on an equilibrium allocation of securities, let us consider the following example of fire insurance. There is an uncertain economy where trader 2’s property may catch a fire so that his wealth position is uncertain while trader 1’s wealth position is certain, i.e., not affected by what would happen to trader 2’s property. Putting it more concretely, let state 1 imply the case where trader 2’s property does not catch a fire and state 2 the case where it does so and burns down. Suppose that trader initially holds the wealth worth 0.8 million dollars regardless of which state occurs in the future and that trader 2’s property amounts to 0.7 million dollars if there is no fire in it while this property value declines to 0.2 million dollars if there is a fire. According to Definition III, we call traders 1 and 2 the insurer and the insured, respectively.

Suppose that both of the insurer and insured believe that the chance of fire in the latter’s property is twenty five percent and that their utility functions are both logarithmic. That is to say, we have

\[
\begin{align*}
(\pi_1, \pi_2) &= (0.75, 0.25) \\
(X_{11}, X_{21}) &= (0.8, 0.8), \quad (X_{12}, X_{22}) = (0.7, 0.2) \\
U(i) &= \log(i); \quad i = 1, 2.
\end{align*}
\]

When there exists a competitive insurance market between these two persons, they will reach a unique competitive equilibrium given by

\[
\begin{align*}
(p_3/p_1) &= 2 \\
(X_{11}, X_{21}) &= (0.9, 0.6), \quad (X_{12}, X_{22}) = (0.6, 0.4).
\end{align*}
\]

This solution means that our insurance market is at equilibrium when the insurer purchases 0.1 units of security 1 and sells 0.2 units of security 2 from and to the insured.
under their premium ratio being 2. In other words, the insured is willing to pay out 0.1 million dollars to the insurer in the non-fire situation with the expectation that he will get the insurance coverage worth 0.2 million dollars when his property happens to burn. Such an insurance contract is easily shown to be Pareto optimal at the instant of its commencement.

Now, what would happen if the insured were not quite truthful about his probability representation in the process of tatonnement? Before entering into this question, we would like to dwell on what the standard monopoly theory teaches us. Let us picture to ourselves an Edgeworth Box Diagram whose horizontal side length measures the total wealth of our economy under state I and vertical side length does its total wealth under state 2. We call the southwestern and northeastern corners of the diagram the origins $O_1$ and $O_2$, respectively. Using the $x$-$y$ coordinate with the origin being $O_1$, the insurer’s offer curve is derived as

$$y = \frac{0.12}{x-0.6} + 0.2,$$

On the other hand, the true expected utility indifference map of the insured is given in terms of the $X_1^2-X_2^2$ coordinate with the origin $O_2$ as

$$0.75 \log X_1^2 + 0.25 \log X_2^2 = \text{constant},$$

or

$$(X_1^2)^3 X_2^2 = \text{constant}.$$

Transforming it to the $x$-$y$ coordinate, such an indifference map is expressed as

$$y = A + \frac{1}{(x-1.5)^3}; \quad A = \text{some constant}.$$

We can prove that the curves (2) and (3) touch to each other at the point $(x,y) = (0.87, 0.64)^5$ Looking at it from the origin $O_2$ of the insured, he can attain $(X_1^2, X_2^2) = (0.63, 0.36)$ by maximizing his expected utility with the insurer’s offer curve as a single constraint, if he has a power of setting premium rate and the insurer is a passive premium taker. It is easy to show that this monopolistic allocation of securities is realized when the insured sets the premium rate at

$$(p_1/p_2)^N = 20/9.$$

He can increase his expected utility

$$0.75 \log 0.6 + 0.25 \log 0.4 = -0.2659$$

under the competitive solution (1) up to the level

$$0.75 \log 0.63 + 0.25 \log 0.36 = -0.2628$$

by behaving as a monopolist. Accordingly, the insurer’s expected utility decreases. We also note that the Pareto optimality is violated under this solution.

What is more important than this observation, however, is the fact that our insured does not have to be an active premium setter to reach the same outcome as this. Based on the analysis of Hurwicz (1972; p. 325), we can easily prove that if the insured faultily

---

5 Exactly speaking, we have the solution $x=(10.4+\sqrt{12.16}/16$ with $y$ being computed out from (2).
uses \((\pi_1^2, \pi_3^2) = (63/79, 16/79)\) in his passive responses to the premium announcements by the auctioneer then the purely competitively achieved equilibrium yields the same allocation of securities and premium rate as the above monopolistic solution. This kind of identity of a faulty competitive equilibrium with a monopolistic equilibrium modifies the view seen in Marshall (1976) which is inclined to conceptually and operationally separate variable probabilities from a monopolistic premium setting. Our identity always holds as long as the insured's utility function is logarithmic. Moreover, a generalization of such a fact to non logarithmic cases would not be too difficult, provided that a utility func-

---

\(X\) : initial endowment  
\(X^*\) : true competitive allocation  
\(X\) : faulty competitive allocation = monopolistic allocation  
LL : offer curve of trader 1  
Pc : price line under true competitive equilibrium  
Pm : price line under faulty competitive equilibrium = monopolistic equilibrium  
EU1(X) : trader 1's true expected utility indifference curve through \(X\); \(l = 1, 2\); \(X = X^*, X\)  
EU2(X) : trader 2's false expected utility indifference curve through \(X\)  

---

\(\pi_1^2, \pi_3^2\) = (63/79, 16/79) suppose that an insured with a logarithmic utility function wants to attain \((X_1^2M, X_2^2M)\) as his security holdings at a faulty competitive equilibrium. Then, it is enough for him to reveal probabilities as \((\pi_1^2, \pi_3^2) = \left(\frac{\alpha}{1+\alpha}, \frac{1}{1+\alpha}\right)\), where \(\alpha = [X_1^2M(p_1/p_2)^M/X_2^2M]\) and \((p_1/p_2)^M = -(X_2^2M - X_2^3)/(X_1^2M - X_1^3)\). This \((p_1/p_2)^M\) gives the premium rate at a faulty equilibrium.
tion under consideration is twice differentiable and its second derivative is negative (Figure 1 will be useful to depict our entire discussion in this section.)

IV. Moral Hazard as Non Incentive Compatibility

In this section, we point out a sort of duality between a mere misrepresentation of probabilities and an active control of the occurrence of a particular state and proceed to unify these two aspects of moral hazard into a single problem of incentive compatibility.

At first, let us change our previous numerical example in the following manner. Suppose that the insured's initial beliefs in the occurrence of states 1 and 2 are expressed by

$$(\pi_1, \pi_2) = (0.797, 0.203)$$

and that his intended carelessness in favor of fire makes him feel that the probabilities of non-fire and of fire are now

$$(\pi_1^*, \pi_2^*) = (0.75, 0.25).$$

This time differently from the previous occasion, we assume that he reveals his true initial beliefs in the successive computation of optimal security holding during the tatonnement process. He then evaluates his expected utility in terms of his controlled (through his intended carelessness) probabilities. In this case, he enjoys his expected utility

$$0.75 \log 0.63 + 0.25 \log 0.36 = -0.2628.$$ 

This expected utility is higher than what he would attain ($-0.2628$) if he revealed his controlled probabilities during the tatonnement process. In our discussion in the previous section, he obtained an extra benefit by faultily representing probabilities without any substantive control over states. But now he obtains an extra benefit with the intended control over states while hiding such a behavior. Hence, we can conclude that a mere, non-substantive manipulation of probabilities and an active control over states are something like the positive and negative of a same picture photographed. Moral hazard can be seen from these two seemingly different points of view.

To capture the essence of the active control over states in a more drastic manner, let us consider the following example. With states 1 and 2 being interpreted as non-fire and fire situations as before, we have

$$(\pi_1^1, \pi_2^1) = (0.9, 0.1)$$

such economy will generate a unique competitive equilibrium given by

$$(p_1/p_2)^* = 33/5$$

$$(X_1^1*, X_2^1*) = (57/55, 19/25), \quad (X_1^2*, X_2^2*) = (51/110, 17/50).$$
Now, what would happen if the insured committed arson without being noticed by the insurer? Suppose that he can set his property afire by a strike of a match so that there is no cost involved here. A part of the data (4) should be rewritten as

\[(\pi_1, \pi_2) = (0.9, 0.1), \quad (\pi_1^2, \pi_2^2) = (0, 1).\]

Then, this economy's competitive equilibrium exists but is not unique. It is given as

\[0 < \frac{p_1}{p_2} < 1.5 \]

\[X_1^{1*} = 1.5, \quad 0.25 < X_2^{1*} < 1\]

\[X_1^{2*} = 0, \quad X_2^{2*} = 1.1 - X_2^{1*}\]

Hence, it is possible under competition that such a point as \((X_1^{1*}, X_2^{1*}) = (1.5, 0.25)\) with \((X_1^{2*}, X_2^{2*}) = (0, 0.85)\) is picked up as an equilibrium allocation of securities. Initially, the insured was to own the wealth worth 0.5 million dollars without fire. This means that his active arson may possibly increase his wealth from 0.5 to 0.85 million dollars at the maximum. This is a general-equilibrium-theoretic interpretation of overinsurance.\(^7\)

Of course, the insurer would not appreciate such a situation to happen. What he would do then may be described as follows. That is to say, he will try to make his client's potential level of wealth not exceed the latter's initial wealth under non-fire situation. To make it possible, the intersection of his offer curve \(y = \frac{\pi_2 X_2}{(x - \pi_1 X_1)}\) with the right hand edge of the Edgeworth Box does not go down below the point \((X_1, X_2) = (1.5, 0.6)\). In order for the offer-curve to satisfy this condition, we should have the condition;

\[
\frac{1.5 \pi_2}{1.5 - \pi_1} \geq 0.6.
\]

From here, we obtain \(\pi_1 \leq 2/3\), or equivalently \(\pi_2 \geq 1/3\). (See Figure 2.)

This observation brings us to a recognition of the fact that a privacy-preserving competitive insurance itself generates incentives towards the violation of a privacy principle. In our present example, the insurer is motivated to look into the property value of his client to exclude the possibility of overinsurance. And this estimation of that property value, in turn, leads him to modify his initial belief in the occurrence of fire in the property. Summarizing all what we have considered so far, we maintain that there are incentives for both the insurer and the insured to change their probabilities from their truthful, initial beliefs in their attempt to exploit an extra benefit or to protect oneself from an unexpected loss by being cheated.\(^8\) Finally, a natural question would arise: Are there any set of probability distributions of both the insurer and insured from which they start off and with which they end up to their own satisfactions? If there is an insurance system which yields as affirmative answer to this question, we may say that such a system is free from moral hazard. Such a system, if exists, will discourage people to manipulate probabilities.

In order to formalize our discussion, let us introduce the following notations. Given

---

\(^7\) For an insightful discussion on overinsurance, see Pauly (1974).

\(^8\) Our analysis of probability misrepresentation or control is somewhat comparable with the problem of manipulation of voting schemes such as the one in Gibbard (1973).
MORAL HAZARD AS A QUESTION OF INCENTIVE COMPATIBILITY

FIGURE 2

some mechanism of securities allocation, let $X_i(\pi^1, \pi^2)$ denote trader $i$'s optimal holding of security $s$ at the equilibrium of that allocation mechanism when traders 1 and 2 reveal (perhaps untruthfully) $\pi^1 = (\pi^1_1, \pi^1_2)$ and $\pi^2 = (\pi^2_1, \pi^2_2)$ during a specific process towards the equilibrium. Using such a notation, we define

$$\begin{align*}
EU^1(\pi^1, \pi^2; \pi^1) &= \pi^1_1 U^1[X_1(\pi^1_1, \pi^2)] + \pi^2_1 U^1[X_2(\pi^1_1, \pi^2)] \\
EU^2(\pi^1, \pi^2; \pi^2) &= \pi^1_2 U^2[X_1(\pi^1_2, \pi^2)] + \pi^2_2 U^2[X_2(\pi^1_2, \pi^2)].
\end{align*}$$

Let $\Pi$ denote the set given by

$$\Pi = \{(\pi_1, \pi_2) \in R_+^2; \pi_1 + \pi_2 = 1\}.$$ 

We then introduce

DEFINITION IV: Given some insurance system, if Nash equilibrium probabilities $\hat{\pi}^1$ and $\hat{\pi}^2$ which satisfy the conditions

$$EU^1(\hat{\pi}^1, \hat{\pi}^2; \hat{\pi}^1) \geq EU^1(\pi^1, \hat{\pi}^2; \hat{\pi}^1) \text{ for all } \pi^1 \in \Pi,$$
EU^2(\hat{\pi}^1, \hat{\pi}^2; \hat{x}^3) \geq EU^2(\hat{\pi}^1, \pi^2; \hat{x}^3) \quad \text{for all } \pi^2 \in \Pi

uniquely exist and turn out to be

\hat{\pi}^1 = \hat{\pi}^2 \quad \text{and} \quad \hat{x}^3 = \hat{x}^3,

where \hat{\pi}^1 and \hat{\pi}^2 denote the initial, truthful probabilities of both traders, then such an insurance system is defined to be free from moral hazard.

With regard to the competitive system of insurance formulated in our Section 2, we have an unfortunate result:

PROPOSITION: A competitive system of insurance is not free from moral hazard in general.

The point of proof: In the context of a certainty economy, Hurwicz (1972) obtained a general result that a privacy-preserving competitive mechanism of resource allocation is not individually incentive compatible. An application of this result to our uncertain economy is enough to prove the statement.

This general outcome does not necessarily mean that a competitive system of insurance always yields motivations for moral hazard. There may be some special cases of economy where the enforcement of a competitive system to it brings about an individually incentive compatible result. However, such cases seem to be fairly rare. Both in practice and theory, therefore, competitive insurance systems generally ask some supplementary devices against moral hazard, of which Helpman-Laffont's proposal is an example. Moreover, the difficulty in finding and/or enforcing such devices within the scope of partially amending competitive mechanism calls for a quest for efficient methods of public provision of insurance.

V. Conclusion

Recent discussions on moral hazard have opened a new perspective in the theory of insurance beyond the realm of the Walras-Arrow-Debreu model of general equilibrium allocation of risk-bearing. To sharpen the economic analysis in this direction, we attempted to set out the moral hazard problem into the framework of the theory of incentive compatibility. As a result, such economic or psychological issues as carelessness, overinsurance, monopolistic behaviors and so on have been shown to have their roots in an anonymity-preserving nature of competitive insurance, in the sense that it is not individually incentive compatible. This type of situations seem to correspond to our daily observation that under the nominal guise of competition, insurance companies search into the privacies of their clients on the one hand and insurance businesses are, in fact, very oligopolistic on the other.

These findings urge us to go further into the investigation towards alternative provision of insurance and efficient methods to amend the defects of competitive system.
REFERENCES