Insurance, Information, and Individual Action

By MICHAEL SPENCE *and* RICHARD ZECKHAUSER
Harvard University

This paper looks at some intricacies and difficulties that arise in the real world operation of contingent claims markets. Insurance contracts, the most readily observable and perhaps most important example of contingent claims markets in action, provide the focus for our discussion.

The purpose of insurance is to protect risk-averse individuals from suffering the full consequences of those actions on the part of nature which affect them unfavorably. The parties to an insurance contract agree that when the actions of nature become known, those most favorably affected will transfer resources to those who turn out to be less fortunate. If the contract is to provide protection in this way, it is essential that there be (at least substantial) independence in the actions nature takes with respect to different insured individuals. Such independence is assumed for the remainder of this paper. This enables us to look at insurance schemes from the standpoint of a single individual who is representative of the many who are insured under a contract. A further assumption is implicit in this approach—it is that all individuals have identical prospects, resources, and utility functions.[1]

The possible arguments of our representative individual's utility function, u, are his overall wealth level, w, the act of nature, n, and his individual action, a. His utility function has the familiar von Neumann-Morgenstern properties; it enables him to make choices over lotteries on all the arguments of his utility function. The individual is assumed to be risk averse with respect to lotteries on wealth.

We classify insurance schemes according to their structural characteristics in three areas. The first is (1) the presence or absence of individual choice. There is individual choice if a is a nontrivial argument in the utility function. The individual choice model is quite general; thus, the choice variable, a, might represent a level of investment, including investment in self, or the purchase of some specific good, say medical services.

Where there is room for individual choice, we will be concerned with (2) the sequencing of moves between the individual and nature, and (3) the information state, S, monitored by the insurer. This information state, which may be a vector, will be a function of the act of nature and the action of the individual.

The insurance scheme works by having the insurer determine a monetary payoff that he makes to the individual, the size of the payoff to depend on the information state he monitors. The insurance scheme is fully described by what we call the insurance payoff function, $g(S)$. It is subject to the actuarial constraint that it have a break-even financial expectation. The insurer's object is to maximize the repre-

[1] This assumption is not restrictive in the manner it might appear. If individuals do differ with regard to any of these variables, the situation can be thought of as one in which nature has already taken a move. In future efforts we hope to get into problems in which there is more than a single exchange of moves between nature and the individual. We also intend to expand the range of actions available to the individual. For example, if nature makes moves before the insurance contract is drawn up, some individuals may choose not to participate. This is the problem of adverse selection. It is not considered here.

sentative individual's expected utility subject to this constraint.

Our principal concern in this paper is to examine how (2) and (3) interact in the determination of optimal insurance schemes. As a standard of comparison, we consider first simple insurance—insurance situations in which there is no room for individual choice.

Case I: No individual Choice—Insurer Monitors n

If individual choice plays no role, the sole determinant of the individual's utility is nature's action, n. The distribution of n is given by nature's density function, $f(n)$. We treat n as a continuous variable in this paper, but with minimal modification our results hold equally well if n is discrete.

The insurer monitors the action of nature; in this case $S=[n]$. He gives the individual a payoff $g(n)$. This payoff added to the insured's initial wealth, w_0, gives the wealth argument of his utility function. The individual's expected utility under this scheme is

$$\int u(w_0 + g(n), n) f(n) dn. \tag{1}$$

The break-even constraint for this scheme is

$$\int g(n) f(n) dn = 0. \tag{2}$$

The insurer's objective is to maximize (1) subject to (2). He can employ the calculus of variations to derive the marginal efficiency condition for the optimal insurance payoff function. That condition is that there be a constant λ such that

$$u = \lambda \tag{3}$$

or $f(n)=0$, for all values of n. The optimal $g(n)$ keeps the marginal utility of income constant. Its dependence on $f(n)$ is only through the value of the parameter λ, which is determined by the constraint equation (2). Thus, the shape of $g(n)$ is only indirectly affected by nature's density function. These properties of the optimal insurance payoff function are characteristic of all cases where the insurer can monitor all available information.

A special subcase of interest is one where the act of nature itself is a monetary payoff so that $u(w, n)$ can be written $v(w+n)$. The optimizing condition becomes

$$v'(w_o + n + g(n)) = \lambda. \tag{4}$$

This implies that $g(n)=k-n$, where k is an arbitrary constant. The integral constraint, (2), requires that $k=\bar{n}$, the mean of n. The optimal insurance scheme when an individual has an uncertain income over which he has no control is one which always gives him his expected income.

We turn, for the remainder of this paper, to cases where the insured individual, in full knowledge of the insurance payoff function, takes an optimizing action, a.

Case II: Individual Chooses Before Nature—Insurer Monitors R *and* a

For simplicity, we consider a model in which the insured's utility function has but one argument, *ex post* wealth. This model requires that the individual's action be convertible to a monetary equivalent. It seems most reasonable to think of a as an investment the individual makes to increase an uncertain monetary return he will receive. We represent this return as $R=r(n, a)$.[2] The insurer monitors this monetary return as well as the individual's

[2] Alternatively, we may think of a as a parameter which alters the distribution of the payoff R. Since $R=r(n, a)$ and n has a distribution, $f(n)$, we may write n as a function of R and a,

$$n = h(R, a).$$

The conditional distribution of R given a is

$$t(R|a) = f(h(R, a)) | h_R(R, a) |.$$

action in determining his payoff; that is, $S=[R, a]$. The individual's wealth after he makes his investment, receives his monetary return, and obtains his payoff from the insurer is

$$(5) \qquad w = w_o + R + g(R, a) - a.$$

The insured will undertake the action that maximizes his expected utility. That is, given the insurance scheme and $f(n)$, he will pick a to maximize his expected utility

$$(6) \qquad \int u(w_o + r(n, a) + g(r(n, a), a) - a)f(n)dn.$$

Here R is written in its functional form, $r(n, a)$, to emphasize its dependence on a.

An individual who invests in his earnings future by obtaining an education, or a person who fireproofs his home to make it less likely to be damaged or destroyed, is an individual taking an action which together with the act of nature determines the monetary return he receives. The individual has some control over his fate in these circumstances, but the additional influence of the uncontrollable action of nature means he cannot determine it completely. Thus, it may be desirable to institute an insurance plan.

Let us look at the problem of the insurer who wishes to maximize the insured's expected utility. The break-even constraint is that

$$(7) \qquad \int g(r(n, a), a)f(n)dn = 0$$

for the particular a that maximizes (6) for the given $g(R, a)$. (Note that this constraint need not be satisfied for other values of a, because they will never be chosen and hence are irrelevant to real-life actuarial considerations.)

The insurer wishes to pick $g(R, a)$ to maximize (6) subject to this constraint. His problem is simplified and his performance (in terms of achieved expected utility) improved because he can monitor a and employ it as an argument of his payoff function. He can make the insured select $a=a^*$ by making $g(R, a)$ sufficiently negative for $a \neq a^*$. Given this control over the insured's action, the insurer's problem becomes to find the pair a^* together with $g(R, a^*)$ that maximizes (6) subject to (7). The optimal $g(R, a^*)$ would be determined, as before in the no-individual-choice case, by the condition

$$(8) \qquad u' = \lambda,$$

which implies that

$$(9) \qquad g_1(R, a^*) = -1.$$

The important point to realize, for this case, is that the adverse incentives problem is eliminated because the insurer can monitor the insured's action and structure the insurance payoff function so that the selected a will equal a^*.[3]

[3] If the insurer had the ability to monitor only n, the outcome that would be achieved would be identical. The insurer would select the optimal insurance payoff function, call it $g^1(n)$, for which there would be a corresponding optimal act for the insured, call it a^*. Expected utility is U^1.

Assume that g^1 is unique and suppose that the insurer could improve expected utility by monitoring $r(n, a)$ and a. Represent his optimal payoff function-action pair as $g^2(r(n, a), a)$ and a^{**}, with expected utility U^2. Now let $g^3(n)=g^2(r(n, a^{**}), a^{**})$. This new insurance payoff function satisfies the break-even constraint; it gives the same distribution of payoffs as did g^2. Let the insured select his optimal act for g^3, call it a^{***}. The expected utility for this pair $U^3 \geq U^2$. But g^3 is an admissible insurance payoff function which monitors only n, Because g^1 was optimal, $U^1 \geq U^3$. The previous weak inequalities imply that these expected utilities are equal. A parallel argument shows that expected utility cannot be greater when $S=[n]$ than when $S=[r(n, a), a]$. Thus we have

$$(a) \qquad g^1(n) \equiv g^2(r(n, a^{**}), a^{**}) \equiv g^3(n),$$

and

$$(b) \qquad a^* = a^{**} = a^{***}.$$

If the insurer can monitor n only, he will

$$(c) \qquad \underset{g(n)}{\text{maximize}} \int u(w_o + r(n, a) + g(n) - a)f(n)dn$$

Case III: Individual Chooses Before Nature —Insurer Monitors Only R

Frequently it is impossible or prohibitively costly for the insurer to monitor a as well as R. The term a may represent, for example, the effort an individual makes to improve his earnings opportunities or the degree of care with which he drives his car. When a cannot be observed directly, the information state contains only R, $S=[R]$, and the insurer's payoff function takes the more restricted form $g(R)$. The opportunity provided before to enforce a choice of a is no longer available.

The insurer's optimization problem becomes more complex. He selects $g(R)$ to maximize

$$\int u(w_o + r(n, a) + g(r(n, a)) - a)f(n)dn, \tag{10}$$

subject to two constraints. The first is the usual break-even condition

subject to

$$\int g(n)f(n)dn = 0 \tag{d}$$

and

$$\int (r_2 - 1)u'f(n)dn = 0, \tag{e}$$

where the second constraint represents the maximizing action of the individual. We know that u' is constant for the optimal scheme, so that the individual's maximizing a satisfies the equation

$$\int (r_2 - 1)f(n)dn = 0. \tag{f}$$

The expected net return from an additional dollar of investment is zero.

Not only does the insurance scheme spread risk, it also makes the investment decision productively efficient by overcoming the distortionary impact of the individual's risk aversion. In general, when there is no insurance, the marginal condition (f) will not be satisfied. In cases where a is a defensive investment against low values of w, as it is say with safety measures, the selected a will be too large. On the other hand, if a generates high payoffs with low probabilities, the chosen a will be too small. Perhaps investment in education for disadvantaged individuals represents the latter situation.

$$\int g(r(n, a))f(n)dn = 0. \tag{11}$$

The second is that a is selected to maximize (10) given $g(R)$. This second constraint can be given by the marginal condition[4]

$$\int [r_2 + g'r_2 - 1]u'f(n)dn = 0. \tag{12}$$

Treating the problem again as one in the calculus of variations, the marginal condition for the optimal $g(R)$ function is

$$\lambda u'' + u'\left[1 + \lambda\frac{d}{dn}\left(\frac{r_2}{r_1}\right) + \lambda\frac{r_2}{r_1}\frac{f'}{f}\right] - \Phi = 0, \tag{13}$$

where λ and Φ are parameters, determined by the constraints (12) and (11).

In this case, g depends directly on the distribution $f(n)$, not just through the parameters λ and Φ. There is a second significant difference from the previous cases. It is no longer possible to keep the marginal utility of wealth constant. The $g(R)$ function must be such that on the whole *ex post* wealth is an increasing function of the payoff received. Otherwise the insured individual would have no incentive to undertake any action at positive cost, no matter how favorably that action would affect this payoff. To achieve appropriate incentives, the insurance plan must sacrifice some of its risk-spreading capabilities.[5]

[4] This will be true, for example, if $u(w_o+r(n, a)+g(r(n, a))-a)$ is everywhere concave in a.

[5] A simple example will make clear the loss of efficiency due to the insurer's inability to monitor a as well as R. The insured's utility function is $\log(w)$, with initial wealth given. The monetary payoff he receives, is functionally defined as $R=(n\cdot a)^{1/2}$. Nature's action is determined by a density function which is uniform on some interval. The form of the optimal payoff function is

$$g(R) = \alpha - R + \beta(1 + \gamma R)^{1/2},$$

where α, $\beta>0$, and $\gamma>0$ are parameters whose values

Case IV: Individual Chooses After Nature —Insurer Monitors n *and* a

We turn our attention now to situations in which the insured individual takes his maximizing action, a, after he learns nature's move. The individual who knows his wage rate (as determined in a lottery conducted by nature) and decides how much he wishes to earn, or the individual who knows his medical condition and decides how much to spend on medical care would be a man in such a situation.

The insurer's monitoring capability enables him to employ a payoff function $g(n, a)$. Insurance mechanisms that fall into this category may be thought of as unfamiliar versions of the common purchase option. Medical insurance reimbursement schemes, for example, state how much an individual must pay if he purchases a given amount of medical service, depending on his medical condition. An income tax scheme endowed with this enriched information-monitoring capability would be empowered to make an individual's tax assessment depend on his earning capability as well as his level of earnings. The individual's action in this case is to sacrifice income to purchase leisure. The tax scheme would tell him how much a given leisure purchase would cost him (income given up less taxes saved) as a function of his wage rate.

We will demonstrate below that if n can be known, the additional ability to monitor a is of no value. This means that there is no loss of efficiency if the payoff from medical insurance is made solely a function of the insured's condition, or if the imposition of an income tax is made to depend only on an individual's probabilistically-determined wage rate.

In this case, the actions of both nature and the individual enter the insured's utility function, $u(w, n, a)$. We assume that u is strictly concave in its first and third arguments. After n is announced, the individual selects a to maximize $u(w_0+g(n, a)-a, n, a)$. Given n and $g(n, a)$, the insured's efficiency condition for the choice of a is

$$u_3 + u_1(g_2 - 1) = 0. \tag{14}$$

The insurer's optimizing problem is to

$$\underset{g(n,a)}{\text{maximize}} \int u(w_o + g(n, a) - a, n, a)f(n)dn, \tag{15}$$

subject to the condition given in (14) and the break-even constraint

$$\int g(n, a)f(n)dn = 0. \tag{16}$$

Because the insurer can monitor n, he can force the insured to select whatever combination of a and w he wants. The system is in this sense controllable. Suppose that the insurer wants $a=p(n)$ and $w=w_o+g(n, a)-a=q(n)$. He may then set

$$\begin{aligned} g(n, a) = {} & \left[1 - \frac{u_3(q(n), n, p(n))}{u_1(q(n), n, p(n))}\right] \\ & \cdot [a - p(n)] \\ & + q(n) + p(n) - w_o. \end{aligned} \tag{17}$$

Since

$$g_2 = 1 - \frac{u_3(q(n), n, p(n))}{u_1(q(n), n, p(n))}, \tag{18}$$

and u is strictly concave in a and w, the insured will select $a=p(n)$ because of condition (14) above. When he does, his wealth will be $q(n)$ as required.

This controllability, which is due to the ability of the insurer to monitor n, simplifies the problem considerably. For the

---

can be found by direct search.

A loss of efficiency comes about because dg/dR is not -1, complete risk spreading is not achieved, as it is when the insurance payoff function can monitor both a and R.

insurer may now substitute the following problem:

$$\underset{p(n),q(n)}{\text{maximize}} \int u(q(n), n, p(n))f(n)dn \tag{19}$$

subject to

$$\int [p(n) + q(n) - w_o]f(n)dn = 0. \tag{20}$$

Employing the standard calculus of variations formulae, he can find the two marginal conditions

$$u_1 = \lambda, \tag{21}$$

and

$$u_3 = \lambda. \tag{22}$$

They imply that the marginal utility of income and dollars expended on services are equal and constant at the optimum. More importantly,

$$g_2 = 1 - \frac{u_3}{u_1} \equiv 0, \tag{23}$$

so that

$$g(n, a) \equiv m(n), \tag{24}$$

for some function $m(n)$. This proves our earlier contention that if the insurer can monitor the act of nature directly, there is no additional gain in being able to monitor the action of the individual. In other words, the achievable expected utility is the same when $S = [n]$ as when $S = [n, a]$.

Case V: Individual Chooses After Nature —Insurer Monitors Only a

We conclude with an examination of the case where the insurer can monitor only a, that is where $S = [a]$. An income tax scheme where the taxing authority can monitor income, but not earning opportunities, or a medical insurance plan that relates only to amounts spent, but not medical condition, would represent such a situation.

Here a is a signal for n. With the appropriate monotonicity properties, it may even be a perfect signal. But due to the problem of adverse incentives, an insurance payoff scheme that monitors a as a signal for n will perform less well than one which can monitor n directly.

This problem is familiar in the context of purchase-option schemes. Income tax plans distort incentives for work; medical insurance reimbursement programs lead individuals to overexpend on medical services. This problem also has interesting ramifications in cases where a is not most appropriately interpreted as a dollar expenditure.

Assume that the amount of sleep an individual chooses is a perfect signal for his medical condition, the more sleep the more serious the condition. If an insurer could not monitor medical condition directly, nor anything else but the individual's chosen quantity of sleep, he would be forced to employ sleep as the sole argument of his insurance payoff function. This would give the insured an incentive to select a supraoptimal amount of sleep no matter what his condition. (We are assuming the insurer would like to give a higher monetary payoff the more serious the condition.) The sleep amount chosen might even remain a perfect signal for medical condition, but there would be no way to overcome the problem that differential payoffs for different amounts of sleep would introduce a new and inappropriate factor into the sleep-quantity decision.

The insurer in Case V recognizes the problems that confront him. He proceeds to optimize in a second-best world, one where the information he would like to monitor is not available. He will choose his insurance payoff function $g(a)$ to achieve the optimal tradeoff between the conflicting goals of furthering risk spreading and providing appropriate incentives. To define this function he can employ the Hamiltonian method with U, achieved utility, serving as the state variable, and g serving as the control variable.[6] For fixed n, a is defined implicitly as a function of U and g by the equation

$$U = u(w_o + g(a) - a, n, a). \tag{25}$$

The optimizing marginal condition for the insured's choice of a is

$$u_3 + u_1(g' - 1) = 0. \tag{26}$$

This may be written in the equivalent form

$$\frac{dU}{dn} = [u_3 + u_1(g' - 1)]\frac{da}{dn} + u_2 = u_2. \tag{27}$$

We let $\Phi(n)f(n)$ be the multiplier corresponding to the state variable U, and λ be the multiplier corresponding to the break-even constraint

$$\int g(a)f(n)dn = 0. \tag{28}$$

With this notation, the Hamiltonian becomes

$$H = [U - \lambda g + \Phi u_2]f. \tag{29}$$

The optimizing equations are therefore

$$\lambda g' = \Phi[u_{23} + u_{12}(g' - 1)], \tag{30}$$

$$\frac{d\Phi}{dn} = -\frac{f'}{f}\Phi - 1, \tag{31}$$

and

$$u_3 + u_1(g' - 1) = 0. \tag{32}$$

We can interpret these results in the concrete context of the medical insurance example. Let n represent medical condition, with larger values representing greater illness. To put the problem in a more tractable form, so as to provide us with further insight into the properties of the optimal payoff function, we consider the case where wealth enters the utility function in an additive way. For this case, $u_{12} \equiv u_{13} \equiv 0$. Thus, condition (30) becomes

$$g' = \frac{\Phi}{\lambda} u_{23}. \tag{33}$$

Typically, u_{23} will be positive, the marginal utility of medical expenditure will increase with an increase in the severity of illness. This implies that g' will be positive as well, in contrast to the situation where the insurer was able to monitor n directly and $\partial g/\partial a \equiv 0$. When only a can be monitored, the optimal insurance plan will not require the insured to pay the full marginal cost of his purchases; he will purchase too much.[7]

From equation (31), we note that $\Phi(n)$ depends entirely on $f(n)$, the distribution of the act of nature. We wish now to get some feeling for the shape of the optimal function $g(a)$. Differentiating (33) with respect to n and solving for the g'' gives

$$g'' = \frac{\Phi'}{\lambda}\frac{u_{23}}{\dot{a}} + \frac{\Phi}{\lambda}u_{223} + \frac{\Phi}{\lambda}\frac{u_{233}}{\dot{a}}. \tag{34}$$

The second order conditions for the insured's optimizing decision imply that $\dot{a} > 0$. Moreover, it seems reasonable to assume that $u_{233} > 0$, meaning that the rate at which diminishing returns to medical services set in decreases with ill health. Similarly $u_{223} > 0$ is reasonable.

The sign of g'' therefore depends upon both the sign and magnitude of Φ' which

[6] A mathematically similar problem in the context of income taxation is treated more completely and rigorously by James A. Mirrlees, "An Exploration in the Theory of Optimal Income Taxation," to appear in *Econometrica*. He uses variational methods in large part, but indicates that the Hamiltonian method can be used in this way. The problems are structurally different in that n, the act of nature, does not appear in the control function $g(a)$ in our problem, whereas it does in his income tax problem.

[7] If the insurer can monitor only the action of the individual and if there is to be any risk spreading at all, the insured must be reimbursed somewhat as his level of expenditure increases. He will not be paying the full marginal cost of his purchases.

in turn depends on the distribution $f(n)$. If $\Phi' > 0$, then $g'' > 0$ and g is convex at that point. However, if Φ' is sufficiently negative, then $g'' < 0$ and g is concave at the relevant point.

We can posit one situation in which we would expect $g(a)$ to be concave, at least over a limited range of values. From equation (31), Φ' will be very large and negative if f'/f is very large and positive. In those regions of a corresponding to the regions of n where f'/f is very great, there will be a sharp increase in the frequency with which a is chosen as its value gets larger. If this frequency increase is sufficiently great, the relative importance of providing appropriate incentives as opposed to furthering risk spreading may increase as a itself increases. This will require that $g(a)$ flatten out somewhat and thus be concave over this range of values for a.[8]

For the most part, however, we would expect $g(a)$ to be convex.[9] The risk spreading objective will receive relatively greater emphasis when expenditures are already high, for the greater is the individual's a, the less is his wealth, and the greater the utility cost to him of further expenditure.

[8] By squishing together the upper values of the n scale, it will always be possible to have f'/f increasing rapidly. However, this will have the effect of increasing the rate of change of a with respect to n, and the two changes will cancel out, having no net effect on the shape of the optimal $g(a)$.

[9] We note that if $g(a)$ is convex over most of its domain, then deductible policies, which are a common form of insurance for automobiles and health, may be a reasonable approximation to the optimum.

Conclusion

Some general principles emerge from our brief examination of different cases of insurance. If the insurer can monitor n directly, as he can with Cases I and IV, the insurance scheme can operate like a traditional contingent claims market. Full risk spreading can be achieved, and there will be no need to worry about adverse incentives.

Similarly, if the insurer can monitor the individual's action taken in advance of nature's act (Case II), the adverse incentives problem can be avoided by structuring the insurance payoff function to enforce the choice of the appropriate a.

These hopeful results do not hold, unfortunately, for Cases III and V. In these cases a signal which depends in part or completely on the insured individual's action is employed as the sole argument of the insurance payoff function. The insured will be induced to alter his natural maximizing action somewhat in order to influence this signal and thus increase his payoff from the insurer. The insurer can be cognizant of this adverse incentives problem, but he cannot overcome it. Given his limited information-monitoring capability, his selection of the optimal insurance payoff function is a second-best exercise. Neither complete risk spreading nor appropriate incentives for individual action will be achieved. To find the optimal mixture of these two competing objectives is a difficult problem, here as in the real world.