ON MORAL HAZARD AND INSURANCE*

STEVEN SHAVELL

I. INTRODUCTION

Moral hazard refers here to the tendency of insurance protection to alter an individual's motive to prevent loss. This affects expenses for the insurer and therefore, ultimately, the cost of coverage for individuals. Beginning with Arrow [1963] and Pauly [1968], economists have discussed two partial solutions to the problem of moral hazard: (i) incomplete coverage against loss and (ii) "observation" by the insurer of the care taken to prevent loss. Incomplete coverage gives an individual a motive to prevent loss by exposing him to some financial risk; and observation of care also gives an individual a motive to prevent loss, as it allows the insurer to link to the perceived level of care either the insurance premium or the amount of coverage paid in the event of a claim.

In examining the partial solution to the problem of moral hazard afforded by incomplete coverage, it is convenient, and in some situations certainly realistic, to assume that observation of care is either impossible or too expensive to be worthwhile. Under that assumption, the degree to which it is desirable to reduce coverage and subject the insured to risk would depend on the incentive thereby created to exercise care, and such an incentive would in turn depend on the cost of taking care. This is the logic underlying the following results: as the cost of taking care falls from very high levels (at which full coverage is best), partial coverage becomes desirable; but at no point is the optimal level of coverage zero—moral hazard cannot entirely eliminate the possibilities for insurance; and as the cost of taking care approaches zero, the optimal coverage, although partial, approaches full coverage.

It is then assumed that observation of care is worthwhile. In this case attention is paid to the accuracy of the insurer's observations and to the timing of the observations, whether they are made ex ante—

* I wish to thank L. Weiss for helpful discussion; K. Arrow, P. Diamond, J. Green, L. Kotlikoff, M. Manove, J. Mirrlees, A. M. Polinsky, and A. Robson for comments; and the National Science Foundation (Grant #SOC76-20862) for financial support.

The Quarterly Journal of Economics, November 1979 0033-5533/79/0093-0541$01.00

whenever a policy is purchased—or ex post—only when a claim is presented.

If the insurer's observations are perfectly accurate, full coverage is desirable,[1] and the value of the information from the observations turns out to be the same whether they are made ex ante or ex post. Thus, if the total cost of ex post observation (which, recall, involves only those individuals presenting claims) is less than that of ex ante observation, it is best for the insurer to acquire information ex post.

If the insurer's observations are not precise, the problem arises that use of the perceived level of care imposes a new kind of risk on the insured. This is because the premium or level of coverage depends on the random factors influencing the insurer's observations. Nevertheless, if the observations convey information about changes in the level of care, it is possible to construct a policy for which the usefulness of the imperfect information as an incentive outweighs its negative effect through the imposition of risk. In contrast to the situation with perfectly accurate observations, partial coverage is generally desirable, and the value of information acquired ex ante exceeds that of information acquired ex post. This complicates the determination of the optimal timing of observations.

The present paper seems most closely related to Pauly [1974] on moral hazard when the insurer does not attempt to observe care; but the paper differs from this reference and most others[2] in that it (i) determines exactly when an insurance policy represents a compromise between no coverage and full coverage (in the case in which the insurer does not observe care), (ii) analyzes the choice concerning the timing of observation of care, and (iii) proves that imperfect information about care is valuable.

II. THE MODEL

Individuals in the model are identical. Therefore, the analysis is relevant either when differences among individuals are unimportant or when these differences are in some way recognizable. In that case, different types of individuals may be treated in an independent manner by insurers.

1. This fact (at least for the case of ex ante observation) is well-known and is noted in most of the papers cited here.

2. See Ehrlich and Becker [1972], Helpman and Laffont [1975], Mirrlees [1975], Spence and Zeckhauser [1970], Stiglitz [1974], Townsend [1976], and Zeckhauser [1970]. See also Harris and Raviv [1978] and Shavell [1979] on the principal and agent relationship.

Each individual is assumed to act so as to maximize the expected utility of wealth, is averse to risk, faces the possibility of financial loss, and is able to affect the probability distribution of loss by taking care. Although care enters formally as an expenditure (for example, on the purchase of an anti-theft device), it may be interpreted also as "effort" (for example, remembering to lock up). Let

$U(\cdot)$	be a twice-differentiable, increasing, and strictly concave function giving utility of wealth;
$y > 0$	be initial wealth;
$x \geqq 0$	be expenditure on care;
$l_i > 0$	$(i = 1, \ldots, n)$ be a possible loss;
$p_i(x) > 0$	$(i = 1, \ldots, n)$ be the probability of loss of l_i; and
$p_0(x) > 0$	be the probability of no loss.

Suppose that the expected value of losses falls with care; that is,

$$\frac{d}{dx}\left[\sum_{i=1}^{n} p_i(x) l_i\right] < 0. \tag{1}$$

However, to reduce the notational burden on the reader, the text considers only the case $i = 1$, in which there is a loss $l = l_1$ of fixed size occurring with probability $p(x) = p_1(x)$, where, by (1), $p'(x) < 0$. The Appendix discusses the general case; the three propositions of the text remain true in the general case.

The general case makes what appears to be the minimal assumption concerning the relation between care and the probability distribution of loss; that more care reduces loss on average. This allows for "self-insurance" (care that has the effect of reducing the magnitude of loss without affecting the probability of loss), for "self-protection" (care that has the effect of reducing the probability of loss without affecting the magnitude of loss), as well as for care that has more complicated effects.[3]

Insurance policies will be formally defined later. An insurance policy resulting in expected profits[4] of zero will be called a *break-even* policy. Ordinarily, different break-even policies would yield different levels of expected utility to an individual. *The problem investigated here is to determine a break-even policy that maximizes expected*

3. The terms self-insurance and self-protection are due to Ehrlich and Becker [1972]. Self-insurance is descriptive of the special case of the general model in which $\sum_{i=1}^{n} p_i(x)$ is a constant, but $p_i(x)$ falls for losses that are high and rises for losses that are low. Self-protection is descriptive of the case in which $p_i(x)$ is for each loss decreasing in x.

4. That insurers use the criterion of expected profits is based on the conventional assumptions that an insurer's risks are borne by many individuals (stockholders) and that the risks insured are numerous, small, and approximately independent.

utility. Such a break-even policy will be called the *optimal insurance policy under moral hazard*.

The optimal insurance policy under moral hazard may be interpreted in two ways: it may be regarded as the best policy that could be sold by a self-financing public insurer. Alternatively, it may be regarded as the policy that would be sold by firms in a competitive insurance industry with free entry, for then it is natural to assume that the only policies which could survive in the marketplace are those that yield expected profits of zero to insurers and, given that constraint, the highest possible expected utility to individuals.

The optimal insurance policy under moral hazard is to be distinguished from the *fully optimal insurance policy*. The latter is the policy that would be sold if an insurer could, in selecting the best break-even policy, choose *independently* an individual's care and the terms of the insurance policy. The problem of moral hazard is precisely that care is chosen by individuals and therefore *does* depend in general on the terms of the insurance policy.

III. Moral Hazard When Care Is Not Observed by the Insurer

If it is either too expensive or impossible for the insurer to observe care, the terms under which insurance is sold obviously cannot depend on care. An insurance policy is therefore described simply by a premium $\pi \geqq 0$ and level of coverage $q \geqq 0$.[5] If an individual decides for some reason to buy a policy (π, q), he then selects x to maximize expected utility:

$$(2) \quad EU = (1 - p(x))U(y - \pi - x) + p(x)U(y - \pi - x - l + q).$$

It is assumed that the x chosen (which will sometimes be written $x(\pi, q)$) is unique,[6] and therefore, if it is positive, it is identified by the first-order condition,

5. It is assumed here that π and q are not random. It is, however, possible (but, I would say, only a theoretical curiosity) that randomness might be desirable: Consider the function $e(r)$, where $e(r)$ is the highest expected utility that can be provided to individuals by insurers selling nonrandom policies and earning a return of r. Let $(\pi(r),q(r))$ be a policy that gives individuals expected utility $e(r)$ when the return to the insurer is r. There is no reason to suspect that $e(r)$ is not locally convex at 0. If it is locally convex at 0, choose ϵ small and let the insurer offer the policy $(\pi(-\epsilon),q(-\epsilon))$ with probability ½ and $(\pi(\epsilon),q(\epsilon))$ with probability ½. Then the insurer will break even, and individual expected utility will be $\frac{1}{2}e(-\epsilon) + \frac{1}{2}e(\epsilon) > e(0)$.

6. Expected utility may not be concave in x; there does not seem to be any simple condition on the function p that would guarantee concavity. Mirrlees [1975] contains an interesting disscussion of this problem.

$$(3)\quad p'(x)[U(y-\pi-x-l+q)-U(y-\pi-x)] \\ = (1-p(x))U'(y-\pi-x)+p(x)U'(y-\pi-x-l+q).$$

The left-hand side is the marginal benefit of taking care and the right-, the marginal cost. If (3) does not hold, then $x(\pi, q) = 0$ and

$$(4)\quad p'(0)[U(y-\pi-l+q)-U(y-\pi)] \\ < (1-p(0))U'(y-\pi)+p(0)U'(y-\pi-l+q).$$

A break-even policy must satisfy

$$(5)\qquad \pi = p(x(\pi,q))q,$$

since expected profits of the insurer must be zero given that individuals choose care. It is assumed that given the coverage q, there is a unique premium $\pi(q)$ such that the insurer breaks even.[7] Writing $x(q)$ for $x(\pi(q),q)$, we may express expected utility as a function of q:

$$(6)\quad EU(q) = (1-p(x(q)))U(y-\pi(q)-x(q)) \\ + p(x(q))U(y-\pi(q)-x(q)-l+q).$$

The optimal insurance policy under moral hazard is found by maximizing (6) over q; this policy will be denoted $(\bar{\pi}, \bar{q})$ and will be assumed unique. Differentiate (6) with respect to q to obtain (noting that $\pi' = x'p'q + p$):

(7)

$$EU'(q) = x'p'[U(y-\pi-x-l+q)-U(y-\pi-x)] \\ -x'[(1-p)U'(y-\pi-x)+pU'(y-\pi-x-l+q)] \\ -x'p'q[(1-p)U'(y-\pi-x)+pU'(y-\pi-x-l+q)] \\ -p[(1-p)U'(y-\pi-x)+pU'(y-\pi-x-l+q)] \\ +pU'(y-\pi-x-l+q).$$

The five terms in this expression reflect, respectively, the following changes that would accompany a small increase in coverage, with the premium adjusting so as to allow the insurer to break even:

(a) a change in the probability of loss

(b) a change in the level of care

(c) a change in the premium due to a change in the premium *rate* per dollar of coverage

(d) a change in the premium due to an increased *level* of coverage

7. That there exists some such π is clear: Assuming that $x(\pi,q)$ is continuous, we note that the function $f(\pi) = p(x(\pi,q))q$ is continuous and maps $[0,q] \rightarrow [0,q]$. Therefore $f(\cdot)$ has a fixed point, at which $\pi = p(x(\pi,q))q$.

(e) a change in the level of coverage.

The changes in expected utility due to (a) and (b) are offsetting (if $x(q) > 0$), since the individual adjusts the level of care to equate them. Thus, (7) reduces to the last three terms, and (c), (d), and (e) are all that one needs to think about. It is in fact only (c), the change in the premium attributable to a change in the rate per dollar of coverage, that reflects moral hazard; (d) and (e) correspond to the benefits and costs of purchasing additional coverage at an actuarially fair rate in the absence of moral hazard. This will help to motivate the following result.

PROPOSITION 1. When care is not observed by the insurer, the optimal insurance policy under moral hazard

(a) always offers positive coverage—moral hazard alone cannot eliminate possibilities for insurance,

(b) offers partial rather than full coverage if the cost of taking care is sufficiently low, but the level of coverage approaches full coverage as the cost of taking care tends to zero.

Note. Starting from a position of no coverage, the term corresponding to (c) must be zero, since no premium is being paid. This is the explanation for the first part of the proposition.

The interpretation of the cost of care, denoted by r, is as follows. If one is thinking of care as an actual expenditure on a good, then r should be regarded as the price of the good. Thus, if x is the level of expenditure, x/r is the amount purchased, and $p(x/r)$ is the probability of loss. On the other hand, if one is thinking of care as effort, $1/r$ should be regarded as the efficiency of effort. Thus, if x is the level of effort, x/r is a measure in "efficiency units," and $p(x/r)$ is the probability of loss. We shall assume here that the probability of loss is bounded above zero.

The explanation for the second part of the proposition is straightforward:[8] If it is not very costly to take care, then the incentive effect due to partial coverage should be strong. Therefore, the advantage of the incentive effect should outweigh the disadvantage of partial coverage, namely, the imposition of risk. Accordingly, we would

8. With regard to the use of partial coverage in a competitive setting, Pauly [1974] brings up the problem that private insurers may not know an individual's purchase of coverage from all sources, so that nothing would prevent an individual from arranging for full coverage by getting partial coverage from different insurers. To the extent that this is a problem, it would be advantageous to have coverage sold by only a single source, such as a public insurer. On the other hand, private insurers often have specific provisions concerning collection from multiple policies. For this and a variety of other reasons, individuals frequently choose to deal with a single company for related lines of insurance.

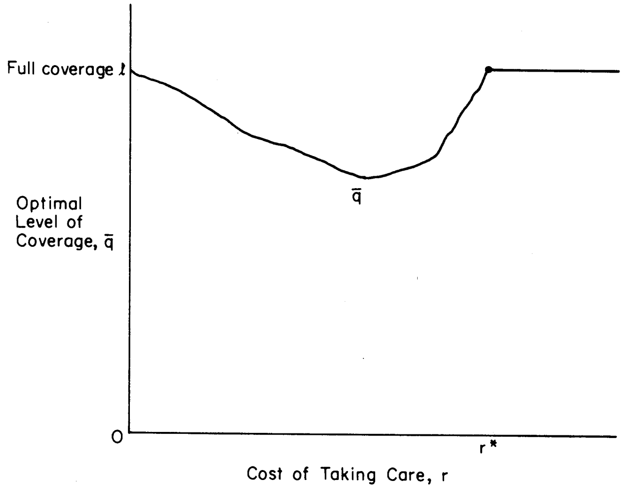

FIGURE I
Optimal Insurance Coverage Under Moral Hazard

expect partial coverage to be optimal if the cost of taking care is below some critical level. However, if it is very cheap to take care, little exposure to risk is needed to induce an individual to take care. Thus, we would expect nearly complete coverage to be optimal if the cost of taking care is close to zero. Figure I illustrates the second part of the proposition. Below a critical level r^*, optimal coverage $\bar{q}$ is partial, but $\bar{q}$ tends to l as r tends to 0. (As drawn, the graph falls and then rises between 0 and r^*, but it could look more complicated.)

Proof. To show that $\bar{q} > 0$, we want to verify that (7) is positive at $q = 0$. If at $q = 0$ the first-order condition (3) is satisfied, then (7) reduces to

$$(8) \quad -p[(1-p)U'(y-\pi-x)+pU'(y-\pi-x-l)] \\ +pU'(y-\pi-x-l) \\ = p(1-p)[U'(y-\pi-x-l)-U'(y-\pi-x)] > 0.$$

On the other hand, if at $q = 0$ (3) is not satisfied, (4) holds, and $x(0) = 0$. In this case, it is easy to show that for all q sufficiently small, $x(q) = 0$.[9] Therefore, $x'(0) = 0$, and (7) again reduces to (8).

9. Since $p'(0)[U(y-l)-U(y)] < (1-p(0))U'(y)+p(0)U'(y-l)$, we must have for q sufficiently small $p'(0)[U(y-qp(0)-l+q)-U(y-qp(0))] < (1-p(0))U'(y-qp(0))+p(0)U'(y-qp(0)-l+q)$.

Let $\bar{q}(r)$ be optimal coverage when the cost of taking care is r. To show that $\bar{q}(r) = l$ if r is sufficiently high, note first that the first-order condition (3), which is appropriate only when $r = 1$, must be rewritten as

$$(3') \quad (1/r)p'(x/r)[U(y - \pi - x - l + q) - U(y - \pi - x)] \\ = (1 - p(x/r))U'(y - \pi - x) + p(x/r)U'(y - \pi - x - l + q).$$

Since this equation cannot hold if r is sufficiently high (the maximum over x and q of the left-hand side tends to 0 as r grows large, but the right-hand side is bounded above 0), we must have for such r that $x(q) \equiv 0$. Hence $x'(q) = 0$, so that $EU'(q)$ reduces to $p(0)(1 - p(0))[U'(y - \pi - l + q) - U'(y - \pi)]$, which is positive for $q < l$. Thus, for such r, $\bar{q}(r) = l$.

To prove that $\bar{q}(r) < l$ if r is sufficiently low, select $\hat{p} < p(0)$ (and in the range of the function $p(\cdot)$), $\hat{q} < l$, and $\hat{x} > 0$ such that[10]

$$(9) \quad (1 - \hat{p})U(y - \hat{p}\hat{q} - \hat{x}) + \hat{p}U(y - \hat{p}\hat{q} - x - l + \hat{q}) \\ > U(y - p(0)l).$$

Note that the right-hand side is the utility of the policy giving complete coverage. If the cost of care r is sufficiently low, the care, say x^0, chosen by an individual with the policy $(\hat{p}\hat{q}, \hat{q})$ is clearly such that $p(x^0/r) < \hat{p}$. Hence an insurer would make profits selling that policy. On the other hand, r can also be chosen low enough so that if $\tilde{x}$ is such that $p(\tilde{x}/r) = \hat{p}$, then $\tilde{x} < \hat{x}$. Thus, as x^0 is the *optimal* choice of the individual with the policy $(\hat{p}\hat{q}, \hat{q})$,

$$(10) \\ (1 - p(x^0/r))U(y - \hat{p}\hat{q} - x^0) + p(x^0/r)U(y - \hat{p}\hat{q} - x^0 - l + q) \\ \geqq (1 - p(\tilde{x}/r))U(y - \hat{p}\hat{q} - \tilde{x}) + p(\tilde{x}/r)U(y - \hat{p}\hat{q} - \tilde{x} - l + q) \\ > (1 - \hat{p})U(y - \hat{p}\hat{q} - \tilde{x}) + \hat{p}U(y - \hat{p}\hat{q} - \hat{x} - l + q) > U(y - p(0)l).$$

In other words, if r is sufficiently small, the policy $(\hat{p}\hat{q}, \hat{q})$ affords greater expected utility than that giving complete coverage and, as constructed, makes profits. Therefore, the policy giving complete coverage cannot be optimal.[11]

To show that $\bar{q}(r) < l$ implies $\bar{q}(r') < l$ for $r' < r$, consider an in-

10. This is clearly possible. Given the choice of $\hat{p}$, choose a small $\hat{x}$ and a $\hat{q}$ close to l.

11. It is easy to rule out the possibility of greater than complete coverage: If $q > l$, then (since the optimal $x = 0$) expected utility is $(1 - p(0)U(y - p(0)q) + p(0)U(y - p(0)q - l + q))$, which (by concavity of U) is less than $U(y - p(0)l)$. In any event, if $q > l$, there would be a motive for fraud.

dividual who is offered $(\pi(\overline{q}(r)), \overline{q}(r))$—the optimal policy if the cost of taking care is r—but who in fact faces the cost r'. He would choose an x leading to a lower probability of loss than his counterpart who faces the cost r and who buys the policy. (This follows from (3′).) Therefore, the insurer would make profits selling the policy. Also, the individual would clearly be better off than his counterpart. But his counterpart would, by assumption, be better off with the policy than if he had full coverage. Therefore, the individual who faces the cost r' would be better off with the policy than with full coverage. Since an insurer selling the policy would make profits and since the individual would be better off with the policy than with full coverage, certainly the optimal policy $(\pi(\overline{q}(r')), \overline{q}(r'))$ cannot involve full coverage.

To show that $\overline{q}(r)$ tends to l as r tends to 0, fix $q < l$ and consider what happens under the policy $(\pi(q),q)$ as r tends to 0. (Note that $\pi(q)$, the break-even premium, depends on r.) First, we claim that x tends to 0 as r tends to 0. This follows from the first-order condition (3′).[12] Second, we assert that $p(x/r)$ tends to p^* as r tends to 0. (Let $p^* > 0$ be the limit of p as its argument tends to infinity.) This is obvious, and the details of the argument are left to the reader. These two facts directly imply that $EU(q)$ tends to $(1 - p^*)U(y - p^*q) + p^*U(y - p^*q - l + q)$ as r tends to 0. Thus, $\lim\limits_{r \to 0} EU(q)$ is increasing in q for $q < l$, and so $\overline{q}(r)$ must approach l as r tends to 0.

Q.E.D.

Of course, the fully optimal insurance policy always involves full coverage.[13] The level of care associated with the fully optimal policy could be either above or below that associated with the optimal policy

12. We only sketch the argument here. Suppose on the contrary that

$$\lim_{r \to 0} x = x^* > 0$$

and consider

$$l^* = \lim_{r \to 0} (1/r)p'(x/r) = \lim_{r \to 0} (1/r)p'(x^*/r)$$

(these limits are assumed to exist). If $l^* < 0$, then, using $x^* > 0$, it can be shown that

$$\lim_{t \to \infty} \int_0^t p'(\tau)d\tau = \lim_t p(t)$$

is unbounded from below, a contradiction, since $p(t)$ is a probability. On the other hand, if $l^* = 0$, then the left-hand side of (3′) tends to zero. But the right-hand side of (3′) is bounded away from zero, a contradiction.

13. This is well-known. To prove it, maximize expected utility given by (2) (and subject to $\pi = p(x)q$) with respect to q and x to determine that $q = l$ at the optimum.

under moral hazard. The latter situation may arise when the optimal policy under moral hazard involves relatively little coverage, in which case the insured is induced to take relatively much care. However, moral hazard does result in too little care in a restricted sense: given the level of coverage, a marginal increase in care (from the level $x(\pi(q), q)$ raises expected utility.[14] This obviously helps to explain why insurers may find it profitable to exhort individuals to take more care even though the cost of such effort by the insurer must be reflected in premiums.

IV. Moral Hazard When Care Is Observed by the Insurer

Both the cost and the potential usefulness of observations of care may depend on when the observations are made.[15] As noted before, it is assumed that observations are made, if at all, either ex ante, when a policy is purchased, or ex post, when a claim is presented.[16] The total cost incurred by the insurer in making observations depends on their timing for two reasons: ex ante observation requires that all policyholders are investigated, while ex post observation requires only that those who make claims are investigated; and the costs of making an individual ex ante versus an individual ex post observation may differ. The potential usefulness of observations also depends on their timing: the quality of ex ante versus ex post observations may differ; and, the quality of ex ante and ex post observations held equal, ex ante observations turn out to be at least useful.

In thinking about the observation of care by insurers, it will be helpful to consider several examples.

(a) *Fire insurance.* Evidence of care (alarms and smoke detectors, absence of oily rags and other hazards) taken to prevent or reduce loss might itself be partially or completely destroyed in a fire, so that an ex ante observation might have an advantage over an ex post observation in quality or cost.

14. Expected utility is $(1 - p(x))U(y - p(x)q - x) + p(x)U(y - p(x)q - x - l + q)$, where $x = x(q)$. The derivative of this with respect to x is $p'(x)[U(y - p(x)q - x - l + q) - U(y - p(x)q - x)] + (-1 - p'(x)q)[(1 - p(x))U'(y - p(x)q - x) + p(x)U'(y - p(x)q - x - l + q)]$, which is positive when evaluated at a point at which (3) holds.

15. The interest here is of course in the potential usefulness of observations as an incentive to take care, not as a means of separating high- from low-risk individuals (for individuals are assumed to be identical).

16. It is also assumed that observations are not made at both times and are not for only a random sample of individuals. Random sampling is discussed in Townsend [1976].

(b) *Theft insurance.* Unless a burglar destroyed or removed evidence of care (burglar alarms, locks), the quality or cost of an observation made after a theft might be equal to that of one made ex ante.

(c) *Automobile collision insurance.* It might be more difficult to evaluate an individual's driving behavior when a policy is purchased than it would be after a collision. An individual would typically be able to alter his driving behavior if evaluated (say, in a driving test) when buying a policy, but if evaluated (using testimony of witnesses or police) after a collision this might not be as much of a problem.

Indeed, for most examples that come to mind, the quality of an ex ante observation is at least as good as—or the cost at least as low as—that of an ex post observation when care is an expenditure on assets that are fixed over the period of insurability, whereas the opposite is true for risk-reducing activity that can be varied to some extent over the period.

Now let

z be the observed level of care (a random variable, the distribution of which depends on the true care x);
$F(;x)$ be the cumulative probability distribution function of z given x; and
c be the cost of making an observation.

The variables z and c and the function F all depend on the timing of observations, but this dependence is not made explicit in the notation, since the timing will usually be clear from context or else it will not matter.

There are two slightly different ways of interpreting a discrepancy $z - x$. First, it may represent an error in observation. For example, an insurance company's inspector may forget to note how many burglar alarms are installed or may not be able to judge adequately their effectiveness. Second, a discrepancy may reflect a random element in the momentary level of care actually taken by an individual. This interpretation, made by Diamond [1974] in a different context, is illustrated by the example of an automobile owner with collision insurance. His true level of care may be thought of as his usual or average driving behavior. His behavior at a particular instant might be due to factors beyond his control (he may be temporarily blinded by the sun). Of course, if the actual driving behavior at the time of an accident is observed with error, both interpretations of a discrepancy might apply.

It is assumed that observations convey information about care in the sense that

$$z = x + \eta, \tag{11}$$

here η is "noise," a random variable (with a distribution that may depend on x) that has a mean of zero.[17] A special case is $\eta = 0$, observations are perfectly accurate.

An insurance policy is a pair (π, q) where, obviously, the premium π may be a function of the observation z only if it is made ex ante, and where the level of coverage q may be a function of the observation whether it is made ex ante or ex post.[18]

If observation of care is made ex ante, an insured individual maximizes over x[19]

$$(1 - p(x))\int U(y - \pi(z) - x)dF(z; x) + p(x)\int U(y - \pi(z) - x - l + q(z))dF(z; x). \tag{12}$$

Thus, the individual takes into account the effect of x on the probability of an accident and on the probability distribution of z. The break-even constraint for the insurer is

$$\int \pi(z)dF(z; x) = c + p(x) \int q(z)dF(z; x). \tag{13}$$

Similarly, if observation of care is made ex post, the individual's problem is to maximize over x

$$(1 - p(x))U(y - \pi - x) + p(x)\int U(y - \pi - x - l + q(z))dF(z; x), \tag{14}$$

and the break-even constraint is

$$\pi = p(x)(c + \int q(z)dF(z; x)). \tag{15}$$

Two cases are now considered: one in which observations are perfectly accurate and one in which they are not.

17. However, it is clear from Shavell [1979] that Proposition 3 holds under the general assumption that the probability distribution of z is different for different x.

18. There is, however, a constraint of sorts on the use of less than perfectly accurate ex ante observations: whereas it is assumed here that an individual decides whether to purchase a policy before an ex ante observation is made, in fact an individual would usually have the opportunity to refuse to buy a policy if its terms—determined after an ex ante observation—were to turn out to be sufficiently unfavorable. The results of this subsection would not be changed if account were taken of the constraint, but a study attempting a detailed characterization of insurance policies would have to recognize it.

19. In the case of a probability distribution with a density function $f(z;x)$, the notation $\int U(y - \pi(z) - x)dF(z;x)$ is a shorthand for $\int U(y - \pi(z) - x)f(z;x)dz$; in the case of a discrete probability distribution, it is a shorthand for $\Sigma_j U(y - \pi(z_j) - x)f(z_j;x)$, where z_j is a possible value of z and $f(z_j;x)$ is the probability of z_j given x.

A. Care Is Observed with Perfect Accuracy by the Insurer

In this simple case, by linking the terms of insurance to the level of care, the insured may be given an appropriate incentive to take care. In particular, there is no need to provide him with an incentive by use of partial coverage. Since this is true whether observations are made ex ante or ex post and since fewer ex post observations are made, ex post observation is superior to ex ante other things equal.

PROPOSITION 2. Suppose that the insurer observes care and that he does so with perfect accuracy. Then[20]

(a) an optimal insurance policy under moral hazard involves full coverage,

(b) care is observed ex post (and the amount of coverage depends on care) unless the relative cost of an ex ante observation is sufficiently small.

Note. The insurer will observe care if the cost of an observation is sufficiently low.

Proof. Let us first verify (a). Suppose that it is optimal to observe care ex ante and consider the policy $(c + p(x)l, l)$. This policy must be optimal given moral hazard and ex ante observation as it breaks even and results in the fully optimal expected utility given ex ante observation.[21] Therefore, any optimal policy given moral hazard and ex ante observation must result in the fully optimal expected utility. But, as shown in the previous footnote, this requires that coverage is full.

Similarly, suppose that it is optimal to observe care ex post and consider the policy $(\pi, \pi/p(x) - c)$, which, given π and x, yields expected utility,

$$(16) \quad (1 - p(x))U(y - \pi - x) + p(x)U(y - \pi - x - l + \pi/p(x) - c).$$

Let $\bar{\pi}$ and $\bar{x}$ maximize (16). Then if the insurer offers the policy $(\bar{\pi},\bar{\pi}/p(x) - c)$, the individual will select $\bar{x}$. On the other hand, maximizing (16) over π and x is equivalent to maximizing (17) over q and x (make the substitution $q = \pi(x) - c$):

20. Part (a) of this proposition is, as previously remarked, well-known, at least with regard to ex ante observation.

21. Under this policy, the individual selects x to maximize $U(y - c - p(x)l - x)$. On the other hand, to find fully optimal expected utility given ex ante observation, it is necessary to maximize over q and x $(1 - p(x))U(y - c - p(x)q - x) + p(x)U(y - c - p(x)q - x - l + q)$. It is easy to check that the optimal $q = l$, so that the problem becomes maximize $U(y - c - p(x)l - x)$ over x.

$$(17)\quad (1 - p(x))U(y - p(x)(q + c) - x) + p(x)U(y - p(x)(q + c) - x - l + q).$$

But this is the problem of determining fully optimal expected utility given ex post observation. Therefore, this case is complete by the logic of the previous paragraph.

The claim of the note to Proposition 2 is true, since we have shown that an optimal insurance policy under moral hazard results in the fully optimal expected utility given that care is observed.

Part (b) is also clear. Let c_a and c_p be the costs of an ex ante and an ex post observation. If care is observed ex ante, expected utility is the maximum over x of $U(y - c_a - p(x)l - x)$ and if care is observed ex post, it is the maximum of $U(y - p(x)(c_p + l) - x)$. Therefore, if $c_a = c_p$, ex ante observation cannot be optimal (since then $y - c_a - p(x)l - x > y - p(x)(c_p + l) - x$ for all x). Also it is easy to see that if it is optimal to observe care ex ante given c_a and c_p, it must be optimal to observe care ex ante given c_a' and c_p for $c_a' < c_a$. Furthermore, if c_a is low enough to make ex ante observation superior to no observation and if, in addition, $c_a < p(l)c_p$, ex ante observation must be optimal.[22] Q.E.D.

B. Care is Observed with Less Than Perfect Accuracy by the Insurer

The motivation for the proof that it is possible to design an insurance policy for which the usefulness as an incentive of imperfect information about care outweighs any negative effect due to the imposition of risk is as follows (and is probably easiest to understand after a look at the general argument of the proof): First, suppose the contrary, that the premium and the amount of coverage are fixed. Now alter the policy by making the amount of coverage depend very slightly on observed care. There will be no first-order effect on the individual's expected utility that can be attributed to the imposition of risk because, initially, his coverage and thus his final wealth were fixed, *conditional* on there being a loss.[23] However, if the policy is

22. The optimal x must certainly be bounded by the loss l and if $x < l$ and $c_a < p(l)c_p$, then $y - c_a - p(x)l - x < y - p(x)(c_p + l) - x$.

23. To illustrate this idea, let us ignore the role of care and assume that the probability of loss is exogenous. Now suppose that a scalar multiple t of a random variable z, which has zero mean, is added to coverage if there is a loss. Then expected utility as a function of t is $EU(t) = (1 - p)U(y - \pi) + p\int U(y - \pi - l + q + tz)dG(z)$. Thus, $EU'(t) = p\int zU'(y - \pi - l + q + tz)dG(z)$ so that $EU'(0) = pU'(y - \pi - l + q)\int zdG(z) = 0$, since z has mean zero. (Here G is the cumulative distribution function of z.) In other words, the first-order effect of the imposition of risk is zero. This is as expected, since a differentiable function is by definition linear, and therefore displays risk neutrality, in the small.

altered in the appropriate way, there will be a positive first-order effect on the care taken by the individual, lowering the probability of loss and therefore allowing a reduction in the premium. The motivation for the proof that imperfect information is more valuable when acquired ex ante is analogous: Suppose that care is observed ex post; that is, suppose that only the amount of coverage can depend on information about care. In this case conditional on there *not* being a loss, final wealth is constant. But the previous logic then suggests that it would be advantageous to make the premium as well depend to some degree on observed care; this requires that care be observed ex ante.

PROPOSITION 3. Suppose that the insurer's observations are made without cost and convey only imperfect information about care (see (11)). Then

(a) either ex ante or ex post observations are of positive value—the terms of the insurance policy will depend to some extent on them;

(b) ex ante observations are more valuable than ex post, at least when the quality of the two types of observations is the same (more precisely, when the probability distribution of z given x is the same in the ex ante and ex post cases and is not degenerate).

Note. It follows from (a) that if the cost of either ex ante or ex post observation is sufficiently low, the insurer will observe care. Moreover, it follows from (b) that if the cost of ex ante observation is sufficiently low, the insurer will observe care ex ante no matter how low the cost of ex post observation (given that the quality of the two types of observation is the same and that the observations are not perfectly accurate).

It can also be shown that the optimal policy typically involves less than complete coverage.

Proof. Recall that $(\bar{\pi}, \bar{q})$ is the optimal policy under moral hazard when care is not observed and let $\bar{x}$ be the associated level of care. Define a new policy by

$$(18) \qquad (\pi, q(z)) = (\bar{\pi} - \alpha\epsilon, \bar{q} + \epsilon(z - \bar{x})),$$

where ϵ and α are greater than zero and will be determined below. This policy has a lower premium than $\bar{\pi}$. It also appears to give an additional incentive to take care, since if x is raised above $\bar{x}$, z will exceed $\bar{x}$ on average, increasing coverage on average. To prove (a), we shall show that if α and ϵ are properly chosen, expected utility will be higher under $(\pi, q(z))$ than under $(\bar{\pi}, \bar{q})$, and that the insurer will at least break even under $(\pi, q(z))$.

The individual now maximizes over x

$$(19)\quad g(x, \epsilon) = (1 - p(x))U(y - \bar{\pi} + \alpha\epsilon - x) + p(x) \times \int U(y - \bar{\pi} + \alpha\epsilon - x - l + \bar{q} + \epsilon(z - \bar{x}))dF(z; x),$$

and it is assumed that

$$(20)\quad g_x(x, \epsilon) = 0$$

determines the optimal x. Let $G(\epsilon) = \max_x g(x, \epsilon)$. Using (20) and noting that at $\epsilon = 0$, $(\pi, q(z)) = (\bar{\pi}, \bar{q})$, we have

$$(21)\quad G'(0) = g_x(\bar{x}, 0)\frac{dx}{d\epsilon} + g_\epsilon(\bar{x}, 0) = g_\epsilon(\bar{x}, 0).$$

To prove (a), we need to show that $G'(0) > 0$ and that α can be chosen so that the premium covers the insurer's expected expenses. Now

$$\begin{aligned}(22)\quad g_\epsilon(\bar{x}, 0) &= (1 - p(\bar{x}))\alpha U'(y - \bar{\pi} - \bar{x}) \\ &\quad + p(\bar{x}) \int (\alpha + z - \bar{x})U'(y - \bar{\pi} - \bar{x} - l + \bar{q})dF(z; \bar{x}) \\ &= \alpha[(1 - p(\bar{x}))U'(y - \bar{\pi} - \bar{x}) + p(\bar{x})U'(y - \bar{\pi} - \bar{x} - l + \bar{q})] > 0,\end{aligned}$$

since $\int (z - \bar{x})dF(z; \bar{x}) = 0$. To show that the insurer remains solvent, write his net revenues R as a function of ϵ,

$$(23)\quad R(\epsilon) = \bar{\pi} - \alpha\epsilon - p(x)[\bar{q} + \epsilon\int (z - \bar{x})dF(z; x)],$$

where x is understood to be a function of ϵ, determined implicitly by (20). Differentiating (23) gives

$$\begin{aligned}(24)\quad R'(\epsilon) &= -\alpha - p'(x)\frac{dx}{d\epsilon}[\bar{q} + \epsilon\int (z - \bar{x})dF(z; x)] \\ &\quad - p(x)\left[\int (z - \bar{x})dF(z; x) + \epsilon\frac{dx}{d\epsilon}\int (z - \bar{x})dF_x(z; x)\right]\end{aligned}$$

so that

$$(25)\quad R'(0) = -\alpha - p'(\bar{x})\left(\frac{dx(0)}{d\epsilon}\right)\bar{q}.$$

We need to show that $R'(0) > 0$ if α is chosen appropriately. To do this, we shall show that $dx(0)/d\epsilon$ is greater than some $\delta > 0$ whenever

α is less than some $\gamma > 0$, for then we need only choose α less than min $(\gamma, -p'(\bar{x})\delta\bar{q})$. Let us write (20) in explicit form:

$$(26) \quad \begin{aligned} 0 = p'(x)[\textstyle\int U(y - \bar{\pi} + \alpha\epsilon - x - l + \bar{q} + \epsilon(z - \bar{x}))dF(z; x) \\ - U(y - \bar{\pi} + \alpha\epsilon - x)] - [(1 - p(x))U'(y - \bar{\pi} + \alpha\epsilon - x) \\ + p(x)\textstyle\int U'(y - \bar{\pi} + \alpha\epsilon - x - l + \bar{q} + \epsilon(z - \bar{x}))dF(z; x)] \\ + p(x)\textstyle\int U(y - \bar{\pi} + \alpha\epsilon - x - l + \bar{q} + \epsilon (z - \bar{x}))dF_x(z; x). \end{aligned}$$

On the other hand, differentiating (20) and solving for $dx(0)/d\epsilon$, we get

$$(27) \qquad \frac{dx(0)}{d\epsilon} = \frac{-g_{x\epsilon}(\bar{x}; 0)}{g_{xx}(\bar{x}; 0)}.$$

As $g_{xx}(\bar{x}; 0) < 0$ (this is the second-order sufficiency condition for the optimal choice of x) and is independent of α, it is enough to show that $g_{x\epsilon}(\bar{x}, 0) > \delta > 0$ whenever α is less than some $\gamma > 0$. Differentiating the right-hand side of (26) with respect to ϵ and evaluating at $(\bar{x}, 0)$, gives

$$(28) \quad \begin{aligned} g_{x\epsilon}(\bar{x}, 0) = \alpha\{p'(\bar{x})[U'(y - \bar{\pi} - \bar{x} - l + \bar{q}) - U'(y - \bar{\pi} - \bar{x})] \\ - [(1 - p(\bar{x}))U''(y - \bar{\pi} - \bar{x}) + p(\bar{x})U''(y - \bar{\pi} - \bar{x} - l + \bar{q})]\} \\ + p(\bar{x})U'(y - \bar{\pi} - \bar{x} - l + \bar{q})\textstyle\int z dF_x(z; \bar{x}). \end{aligned}$$

Use was made here of $\int dF_x(z; \bar{x}) = 0$, which follows from the identity $\int dF(z; x) = 1$. Now $\int z dF_x(z; x) = d(\int z dF(z; x))/dx$. But by (11), $\int z dF(z; x) = x$. Thus, $\int z dF_x(z; x) = 1$, so that the right-hand side of (28) is of the form $\alpha K_1 + K_2$ with $K_2 > 0$. Therefore, if $\delta = K_2/2$ and $\gamma = K_2/(|K_1|2)$, then $g_{x\epsilon}(\bar{x}, 0) > \delta$ for all $\alpha < \gamma$, which completes the proof of (a).

We shall prove (b) by contradiction. Thus, suppose that z is observed ex ante and that $(\pi, q(z))$ is the best policy under moral hazard. We shall construct a new policy $(\hat{\pi}(z), \hat{q}(z))$ under which expected utility is higher and the insurer at least breaks even. Let

$$(29) \qquad \hat{\pi}(z) = \pi - \alpha\epsilon - \epsilon\beta(z - x^*),$$

where α, β, and ϵ are positive and x^* is the optimal x given $(\pi, q(z))$. Also, let

$$(30) \qquad \hat{q}(z) = q(z) - \epsilon\beta(z - x^*).$$

Note that $(\hat{\pi}(z), \hat{q}(z))$ is designed so that if there is a loss, the stochastic component of the premium in the new policy is exactly offset by the new stochastic component of coverage. Thus, additional risk in the new policy is imposed only when there is not loss. The individual maximizes over x

$$(31)\quad h(x, \epsilon) = (1 - p(x))\int U(y - \pi + \alpha\epsilon + \epsilon\beta(z - x^*) - x)dF(z; x) + p(x)\int U(y - \pi + \alpha\epsilon - x - l + q(z))dF(z; x),$$

and it is assumed that

$$(32)\qquad h_x(x, \epsilon) = 0$$

determines the optimal x. Let $H(\epsilon) = \max_x h(x, \epsilon)$, so that

$$(33)\qquad H'(0) = h_\epsilon(x^*, 0).$$

To prove (b), we need to show that $H'(0) > 0$ and that α and β can be chosen so that the premium covers the insurer's expected expenses. Now

$$(34)\quad h_\epsilon(x^*, 0) = \alpha[(1 - p(x^*))U'(y - \pi - x^*) + p(x^*)\int U'(y - \pi - x^* - l + q(z))dF(z; x^*)] > 0.$$

To show that the insurer remains solvent, write his net revenue S as a function of ϵ:

$$(35)\qquad S(\epsilon) = \pi - \alpha\epsilon - (1 - p(x))\epsilon\beta\int(z - x^*)dF(z; x) - p(x)\int q(z)dF(z; x),$$

where x is understood to be a function of ϵ determined by (32). Differentiating (35), we see that

$$(36)\quad S'(\epsilon) = -\alpha - (1 - p(x))\beta\int(z - x^*)dF(z; x) + \frac{dx}{d\epsilon}p'(x)[\epsilon\beta\int(z - x^*)dF(z; x) - \int q(z)dF(z; x)] - \frac{dx}{d\epsilon}[(1 - p(x))\epsilon\beta\int(z - x^*)dF_x(z; x) + p(x)\int q(z)dF_x(z; x)].$$

Therefore,

$$(37)\quad S'(0) = -\alpha + \frac{dx(0)}{d\epsilon}[-p'(x^*)\int q(z)dF(z; x^*) - p(x^*)\int q(z)dF_x(z, x^*)].$$

The two terms in brackets represent two effects of an increase in the level of care: a reduction in expected payments by the insurer due to a decline in the probability of an accident and a change in expected payments by the insurer due to a change in the distribution of z.

Suppose that the net effect of these terms is postive. Then, as in the proof of (a), $S'(0) > 0$ holds if $dx(0)/d\epsilon$ is greater than some positive δ whenever α is sufficiently small and β is chosen appropriately. The latter is shown to hold by steps analogous to those used in the proof of (a): The first-order condition determining individual behavior is

$$\begin{aligned} (38) \quad h_x(x, \epsilon) = p'(x)[&-\textstyle\int U(y - \pi + \alpha\epsilon + \epsilon\beta(z - x^*) - x) \\ &\times dF(z; x) + \textstyle\int U(y - \pi + \alpha\epsilon - x - l \\ &+ q(z))dF(z; x)] \\ &- \{(1 - p(x))\textstyle\int U'(y - \pi + \alpha\epsilon \\ &+ \epsilon\beta(z - x^*) - x)dF(z; x) \\ &+ p(x)\textstyle\int U'(y - \pi + \alpha\epsilon - x - l + q(z))dF(z; x)\} \\ &+ (1 - p(x))\textstyle\int U(y - \pi + \alpha\epsilon \\ &+ \epsilon\beta(z - x^*) - x)dF_x(z; x) \\ &+ p(x)\textstyle\int U(y - \pi + \alpha\epsilon - x - l \\ &+ q(z))dF_x(z; x) = 0. \end{aligned}$$

Differentiating $h_x(x, \epsilon) = 0$ with respect to ϵ and solving for $dx/d\epsilon$, we get

$$(39) \qquad \frac{dx}{d\epsilon} = \frac{-h_{x\epsilon}(x, \epsilon)}{h_{xx}(x, \epsilon)}.$$

As $h_{xx}(x^*, 0) < 0$ (the second-order sufficiency condition for a maximum) and is independent of α and β, it suffices to show that $h_{x\epsilon}(x^*, 0) > \delta > 0$ as long as α is sufficiently small. But

$$\begin{aligned} (40) \quad h_{x\epsilon}(x^*, 0) = \alpha\{&p'(x^*)[-U'(y - \pi - x^*) \\ &+ \textstyle\int U'(y - \pi - x^* - l + q(z))dF(z; x^*)] \\ &- [(1 - p(x^*))U''(y - \pi - x^*) \\ &+ p(x^*)\textstyle\int U''(y - \pi - x^* - l \\ &+ q(z))dF(z; x^*)] \\ &+ p(x^*)\textstyle\int U'(y - \pi - x^* - l + q(z))dF_x(z; x^*)\} \\ &+ (1 - p(x^*))U'(y - \pi - x^*)\beta\textstyle\int z dF_x(z; x^*), \end{aligned}$$

which is of the form $\alpha K_1 + \beta K_2$ where $K_2 > 0$. Thus, if α is less than or equal to, say, ω, (40) will be positive if β is chosen to be a least $|K_1\omega|/K_2$.

On the other hand, suppose that the terms in brackets in (37) have a negative sum. Then redefine $(\hat{\pi}(z), \hat{q}(z))$ as follows:

$$\hat{\pi}(z) = \pi - \alpha\epsilon + \epsilon\beta(z - x^*) \tag{41}$$

$$\hat{q}(z) = q(z) + \epsilon\beta(z - x^*). \tag{42}$$

Then the proof that was just used for $(\hat{\pi}(z), \hat{q}(z))$ as initially defined may be carried out, the only change being that $dx(0)/d\epsilon$ is shown to be negative rather than positive.

The terms in brackets in (37) cannot sum to zero if z is not a degenerate random variable: Otherwise, assume that

$$-p'(x^*)\int q(z)dF(z; x^*) - p(x^*)\int q(z)dF_x(z; x^*) = 0. \tag{43}$$

We shall show that this allows us to construct a new policy that breaks even and improves expected utility, in contradiction to the supposed optimality of $(\pi, q(z))$. To do so, select z_1 and z_2 such that $q(z_1) > q(z_2)$. (This can be done by part (a) of the Proposition.) Hence

$$U'(y - \pi - x - l + q(z_1)) < U'(y - \pi - x - l + q(z_2)). \tag{44}$$

Suppose that z_1 and z_2 occur with positive probabilities $p(z_1; x)$ and $p(z_2; x)$ (it will be clear that the argument can be easily modified if the distribution of z is not discrete). Then define a new policy $(\tilde{\pi}, \tilde{q}(z); \lambda, \delta(\lambda))$ by

$$\begin{aligned} \tilde{\pi} &= \pi + \delta(\lambda) \\ \tilde{q}(z_1) &= q(z_1) - \lambda \\ \tilde{q}(z_2) &= q(z_2) + \frac{p(z_1; x^*)}{p(z_2; x^*)}\lambda \\ \tilde{q}(z) &= q(z) \qquad \text{otherwise,} \end{aligned} \tag{45}$$

where $\lambda > 0$ and $\delta(\lambda)$ is chosen to satisfy the break-even constraint (13) given that x is chosen optimally. Let expected utility as a function of x be $f(x, \lambda, \delta(\lambda))$. Since the individual chooses x optimally, his expected utility is $F(\lambda) = \max_x f(x; \lambda, \delta(\lambda))$. Thus, since $\delta(0) = 0$,

$$\frac{dF(0)}{d\lambda} = f_\lambda(x^*; 0,0) + \delta'(0)f_\delta(x^*; 0,0). \tag{46}$$

But

$$\begin{aligned} f_\lambda(x^*; 0,0) = (&U'(y - \pi - x^* - l + q(z_2)) \\ &- U'(y - \pi - x^* - l + q(z_1))) \\ &\times p(z_1; x^*)p(x^*) > 0. \end{aligned} \tag{47}$$

Consequently, it suffices to show that $\delta'(0) = 0$, for then (46) must be positive. Now profits of the insurer as a function of λ and δ are

$$T(\lambda, \delta) = \pi + \delta - p(x)\int \tilde{q}(z)dF(z; x), \tag{48}$$

where x is understood to have been chosen optimally. Since by definition of $\delta(\lambda)$, $T(\lambda, \delta(\lambda)) = 0$, we have $\delta'(\lambda) = -T_\lambda(\lambda, \delta)/T_\delta(\lambda, \delta)$. But if we use (43),

$$(49)\quad T_\lambda(0,0) = \frac{dx}{d\lambda}\left[-p'(x^*)\int q(z)dF(z;x^*) - p(x)\int q(z)dF_x(z;x^*)\right] = 0$$

so that $\delta'(0) = 0$ as required.

Q.E.D.

In summary of Propositions 2 and 3 and of the earlier discussion in this section, three factors may be identified as influencing the optimal timing of observations: (1) The number of individuals that have to be checked. Consideration of this factor works in favor of ex post observation. (2) The value of imperfect information. Consideration of this factor works in favor of ex ante observation. And (3) the cost and quality of the two types of observation. Consideration of this factor usually works in favor of ex ante observation when care is an expenditure on a good that is fixed over the period of insurability, whereas it may favor ex post observation for risk-reducing activity that can be varied over the period.

APPENDIX

In the general case the proofs to the propositions are virtually identical to what was given in the text. All that we shall write here are the analogs of several equations of the text. The analog of (2) is

$$(2')\quad EU = p_0(x)U(y-\pi-x) + \sum_{i=1}^{n} p_i(x)U(y-\pi-x-l_i+q_i);$$

that of (7),

$$(7')\quad \frac{\partial}{\partial q_j}EU(q_1,\ldots,q_n) = \frac{\partial x}{\partial q_j}\sum_{i=1}^{n} p_i'(U(y-\pi-x-l_i+q_i) - U(y-\pi-x)) - \frac{\partial x}{\partial q_j}\left[p_0U'(y-\pi-x) + \sum_{i=1}^{n} p_iU'(y-\pi-x-l_i+q_i)\right] - \frac{\partial x}{\partial q_j}\sum_{i=1}^{n} p_i'q_i\left[p_0U'(y-\pi-x) + \sum_{i=1}^{n} p_iU'(y-\pi-x-l_i+q_i)\right] - p_j\left[p_0U'(y-\pi-x) + \sum_{i=1}^{n} p_iU'(y-\pi-x-l_i+q_i)\right] + p_jU'(y-\pi-x-l_j+q_j);$$

that of (12),

$$(12') \quad p_0(x)\int U(y - \pi(z) - x)dF(z; x) + \sum_{i=1}^{n} p_i(x)\int U(y - \pi(z) - x - l_i + q_i(z))dF(z; x);$$

that of (18),

$$(18') \quad \pi = \bar{\pi} - \alpha\epsilon$$
$$q_i(z) = \bar{q}_i + \epsilon(z - \bar{x}),$$

and so forth. The only difference in the general case is with regard to the first proposition: in the general case "positive coverage" means that $q_i \geqq 0$ with strict inequality for some i and "partial coverage" means that $q_i \leqq l_i$ with strict inequality for some i.

HARVARD UNIVERSITY

REFERENCES

Arrow, Kenneth J., "Insurance, Risk and Resource Allocation," in *Essays in the Theory of Risk-Bearing* (Chicago: Markham, 1971).
Diamond, P., "Single Activity Accidents," *Journal of Legal Studies,* III (Jan. 1974), 107–64.
Ehrlich, I., and G. Becker, "Market Insurance, Self-insurance, and Self-protection," *Journal of Political Economy,* LXXX (1972), 623–48.
Harris, M., and A. Raviv, "Optimal Incentive Contracts with Imperfect Information," Carnegie-Mellon University, mimeo, 1978.
Helpman, E., and J. Laffont, "On Moral Hazard in General Equilibrium Theory," *Journal of Economic Theory,* X (Feb. 1975), 8–23.
Mirrlees, J., "On Moral Hazard and the Theory of Unobservable Behavior," Oxford, mimeo, 1975.
Pauly, M., "The Economics of Moral Hazard: Comment," *American Economic Review,* LVIII (June 1968), 531–36.
——, "Overinsurance and Public Provision of Insurance," this *Journal,* LXXXVII (1974), 44–62.
Shavell, S., "Risk Sharing and Incentives in the Principal and Agent Relationship," *Bell Journal of Economics,* X (1979), 55–73.
Spence, A. M., and R. Zeckhauser, "Insurance, Information, and Individual Action," *American Economic Review,* LXI (1971), 380–87.
Stiglitz, J. E., "Incentives and Risk Sharing in Sharecropping," *Review of Economic Studies,* LXI (1974), 219–56.
Townsend, R., "Optimal Contracts and Competitive Markets with Costly State Verification," Carnegie-Mellon University, mimeo, 1976.
Zeckhauser, R., "Medical Insurance: A Case Study of the Trade-off Between Risk-Spreading and Appropriate Incentives," *Journal of Economic Theory,* II (March 1970), 10–26.