The Geneva Papers on Risk and Insurance Theory, 19: 139-152 (1994)

Optimal Insurance Contracts When Establishing the Amount of Losses Is Costly

LOUIS KAPLOW
Harvard Law School, Cambridge, MA 02138 and the National Bureau of Economic Research

Abstract

The problem of establishing the amount of losses covered by public and private insurance is often characterized by asymmetric information, in which the claimant already knows the extent of a loss but this can be demonstrated to the insurer only at a cost. It is shown that a simple arrangement, which provides greater coverage whenever individuals demonstrate high losses, gives claimants an excessive incentive to establish the amount of their losses. This paper determines what insurance claims process, consistent with the form typically employed in existing insurance arrangements, is optimal.

Key words: Insurance policy characteristics, insurance claims, asymmetric information

1. Introduction

An important feature of insurance in many contexts is that the insured will know the extent of a loss, but it is costly to establish this to the insurer.[1] For example, with disability insurance, the extent of an injury may be apparent to the insured but not costlessly observable by others. This paper considers how to design insurance contracts optimally when establishing losses is costly.[2]

A primary conclusion is that individuals may have excessive incentives to establish their losses. Suppose, for example, that an insured has suffered a loss the magnitude of which, if established, would entitle him to $1,000 of insurance coverage. He would then be willing to incur costs of up to $1,000 to demonstrate his loss, so his gain will be smaller, possibly much smaller, then the $1,000 of coverage he receives. Yet the insured pays fully, in his premium, for the prospect of receiving the $1,000 of coverage. This suggests that the insured might be better off with a contract that sometimes discourages him from later spending to establish claims. When an insured decides how much to spend to demonstrate losses, he considers only his ex post situation, but the ex post gain does not equal the ex ante benefit of insurance, which is the efficient spreading of risk.

Section 2 first shows that insureds' incentives to establish losses are excessive for a natural type of insurance contract. It then characterizes the contract, consistent with the form typically employed in existing insurance arrangements, that maximizes expected utility. Under the optimal contract, there is a range of "high" losses that, if demonstrated, do not entitle the claimant to greater coverage (even though greater coverage would be optimal taking as given that the loss has been demonstrated). Individuals suffering the highest losses—above the range just noted—do receive greater coverage (in fact, full coverage) when they demonstrate their loss. Individuals who do not demonstrate the amount of their loss (or who demonstrate a loss that is not high enough) receive a uniform, positive payment.

Section 3 discusses the generality of the results. It also addresses the extent to which important and controversial features of insurance schemes—including public programs such as Social Security Disability Insurance, which involves annually over a million claims, payments of more than $20 billion, and public administrative expenses approaching $1 billion (Social Security Administration [1991])—are consistent with the results here, particularly those concerning the problem of excessive incentives to establish claims.

The present article may be contrasted with previous investigations of administrative costs (loading) and optimal insurance contracts. Most prior work examines how a given cost structure affects optimal coverage, rather than the source of administrative costs and the problem of asymmetric information concerning losses (see Gollier [1987], Raviv [1979], Shavell [1978]).[3] In the present model, unlike most others, costs arise from establishing the amount of claims and the problem of an optimal contract is complicated by the effect of coverage provisions on the incentive to incur these costs.[4]

2. Analysis

This section states the model, analyzes a simple insurance contract of a form generally used, derives the optimal contract, and presents a numerical example.

2.1. The model

Individuals may suffer losses, the existence of which is costlessly observed but the amount of which is observed only by individuals. Individuals choose whether to demonstrate the amount of their losses to the insurer at a cost. Insurance contracts provide nonnegative coverage that is a function of the loss if the amount of the loss is demonstrated; otherwise, coverage for any loss is a single, specified amount.[5] The notation is as follows:

w = initial wealth
u = individuals' utility functions, $u' > 0$, $u'' < 0$
p = probability of a loss
ℓ = amount of loss
$f(\ell)$ = positive density of losses on $[0, w]$
c = cost of demonstrating amount of a loss
x = insurance coverage if amount of a loss is not demonstrated
$x(\ell)$ = insurance coverage if a loss in the amount of ℓ is demonstrated
π = actuarially fair insurance premium

Let $\mathcal{L}$ denote the set of losses for which individuals do not demonstrate the amount of their loss and $\sim\mathcal{L}$ the set for which losses are demonstrated. Expected utility is then given by the following expression:

$$EU = (1 - p)u(w - \pi) + p \int_{\mathcal{L}} u(w - \pi - \ell + x)f(\ell)d\ell + p \int_{\sim\mathcal{L}} u(w - \pi - \ell + x(\ell) - c)f(\ell)d\ell. \quad (1)$$

The problem is to choose $x(\ell)$ and x to maximize (1) subject to the actuarial requirement that

$$\pi = px \int_{\mathcal{L}} f(\ell)d\ell + p \int_{\sim\mathcal{L}} x(\ell)f(\ell)d\ell \tag{2}$$

and an incentive compatibility constraint, which indicates that an individual will choose to demonstrate the amount of his loss (i.e., $\ell \notin \mathcal{L}$) if and only if

$$x(\ell) - c \geq x. \tag{3}$$

In the familiar case in which it is implicitly assumed that $c = 0$, the optimal insurance contract involves full coverage ($x(\ell) = \ell$ for all ℓ), each individual's loss is demonstrated, and the insurance premium equals the expected loss.

2.2. *Excessive incentives under simple insurance contracts*

Consider individuals' incentives to establish their losses under contracts in which individuals who demonstrate losses receive coverage $x(\ell)$ that is continuously increasing in the amount of the loss ℓ that is demonstrated (such as a full coverage contract). It is now proved that such contracts are inefficient because insured individuals sometimes have excessive incentives to demonstrate their losses.

Proposition 1. *Suppose that under an insurance contract $x(\ell)$ rises continuously with ℓ and there exists an $\ell < w$ that would be demonstrated. Then, expected utility is higher under an alternative contract under which fewer individuals demonstrate their losses.*

Proof: First, consider the case in which there exists an $\ell_1 < w$ such that the incentive compatibility constraint (3) holds as an equality. Because $x(\ell)$ is continuously increasing, all individuals for whom $\ell \geq \ell_1$ demonstrate their losses and others do not. Compare an alternative contract in which $x(\ell) = x$ for all $\ell \in [\ell_1, \ell_2]$, where $\ell_2 \geq \ell_1$. It is clear that individuals in this interval will not demonstrate their losses, so that

$$EU = (1 - p)u(w - \pi) + p \int_0^{\ell_2} u(w - \pi - \ell + x)f(\ell)d\ell$$
$$+ p \int_{\ell_2}^{w} u(w - \pi - \ell + x(\ell) - c)f(\ell)d\ell. \tag{4}$$

To find the optimal modified contract of this type, examine

$$\frac{dEU}{d\ell_2} = pf(\ell_2)[u(w - \pi - \ell_2 + x) - u(w - \pi - \ell_2 + x(\ell_2) - c)$$
$$+ \bar{u}' \cdot (x(\ell_2) - x)], \tag{5}$$

where $\bar{u}'$ is the expected marginal utility of wealth. (The final term is the effect on EU from the change in π.) At $\ell_2 = \ell_1$, $x(\ell_2) - c = x$, so the first two terms combined equal zero and the third term is positive. (Raising ℓ_2 reduces coverage in states with losses of ℓ_2 by an amount that just equals the savings in demonstration costs, and it also reduces the premium.) Therefore, the optimal modified contract has $\ell_2 > \ell_1$ and produces higher expected utility than the original contract.

Second, consider the case in which there does not exist an $\ell_1 < w$ such that (3) holds as an equality. Then either no individuals demonstrate their ℓ (a case ruled out by the assumption of the proposition) or all individuals demonstrate their ℓ. The latter case can occur only if $x(0) - c \geq x$. Consider the alternative contract under which $x = x(0)$, with $x(\ell)$ as before. Then individuals with losses such that $x(\ell) \geq x(0) + c$ will demonstrate their losses and receive $x(\ell)$, as under the original contract. Individuals with losses sufficiently small that $x(\ell) < x(0) + c$ will not demonstrate their losses and will receive $x(0)$. Each such individual is better off ex post than under the original contract because $x(0) > x(\ell) - c$. Moreover, because these individuals receive less from the insurer ($x(0) < x(\ell)$ for $\ell > 0$), the insurance premium must be lower. Thus, expected utility must be higher under the alternative contract in which individuals with low losses are discouraged from demonstrating their losses. ■

2.3. *The optimal insurance contract*

Proposition 2. *Under the optimal insurance contract, there exists a critical loss level $\hat{\ell}$, $x < \hat{\ell} \leq w$, such that*

$$x(\ell) = \begin{cases} x, \text{ for } \ell \leq \hat{\ell} \\ \ell + c, \text{ for } \ell > \hat{\ell} \end{cases}$$

and individuals demonstrate their losses if and only if $\ell > \hat{\ell}$.

Remarks: The optimal contract provides full coverage for sufficiently high losses, which will be demonstrated. (Full coverage includes the cost of demonstrating the losses.) Low losses will not be demonstrated; it will be shown that the level of coverage x is that which would be optimal taking as given that those losses are not demonstrated. Individuals for whom $\ell \in [x, \hat{\ell})$ would (inefficiently) demonstrate their losses if $x(\ell) = \ell + c$—that is, if they were fully covered. The optimal contract sets $x(\ell) = x$ when $\ell \leq \hat{\ell}$ to discourage this.

Observe that the optimal contract characterized in Proposition 2 involves a function $x(\ell)$ that is discontinuous: it equals x for $\ell \leq \hat{\ell}$ and a value exceeding $x + c$ for all higher ℓ. This is consistent with Proposition 1, which states that it is not optimal for $x(\ell)$ to be continuously increasing in ℓ.

Proof: The proof consists of three steps: (1) deriving optimal coverage taking as given which losses are demonstrated; (2) showing that the same coverages are optimal when the incentive compatibility constraint is taken into account; and (3) determining which losses are demonstrated under the optimal contract.

Step 1: Optimal coverage taking as given which losses are demonstrated. It will be useful initially to characterize optimal coverage taking as given which losses are demonstrated. Consider optimal coverage for losses in $\mathcal{L}$, those that are not demonstrated. (This assumes that $\mathcal{L}$ is not a set of measure zero; if it is, let $x = 0$.)

$$\frac{dEU}{dx} = -\frac{d\pi}{dx} \cdot \bar{u}' + p \int_{\mathcal{L}} u'(w - \pi - \ell + x) f(\ell) d\ell = 0. \tag{6}$$

From (2), $d\pi/dx = p\int_{\mathcal{L}} f(\ell)d\ell$, so this condition reduces to

$$\bar{u}'_{\mathcal{L}} = \bar{u}', \tag{7}$$

where $\bar{u}'_{\mathcal{L}}$ is the expected marginal utility of wealth for states in which there is a loss the amount of which is not demonstrated.[6]

For each loss ℓ that is demonstrated, I now show that optimal coverage is $x(\ell) = \ell + c$.[7] Before proceeding, observe that when coverage is at this level,

$$u'(w - \pi - \ell + x(\ell) - c) = \bar{u}'. \tag{8}$$

That is, coverage of $\ell + c$ equates the marginal utility of wealth when the loss is demonstrated with the expected marginal utility of wealth. The reason is that this coverage results in a marginal utility of wealth equal to $u'(w - \pi)$, which is also the marginal utility of wealth when there is no loss. From (7), it therefore follows that this level of marginal utility must equal the expected marginal utility.[8]

If optimal coverage is not given by $x(\ell) = \ell + c$, there must exist some set of losses in $\sim\mathcal{L}$ with higher or lower coverage.[9] Suppose that there exists a set of losses H that optimally have higher coverage. Consider a modified contract in which all coverages are the same except that coverage of losses in H is changed by e. Expected utility is

$$\begin{aligned} EU = {} & (1 - p)u(w - \pi) + p \int_{\mathcal{L}} u(w - \pi - \ell + x) f(\ell) d\ell \\ & + p \int_{\sim\mathcal{L} \cap \sim H} u(w - \pi - \ell + x(\ell) - c) f(\ell) d\ell \\ & + p \int_{H} u(w - \pi - \ell + x(\ell) + e - c) f(\ell) d\ell. \end{aligned} \tag{9}$$

To find the optimal modified contract, examine

$$\frac{dEU}{de} = -\bar{u}' \cdot p \int_H f(\ell) d\ell + p \int_H u'(w - \pi - \ell + x(\ell) + e - c) f(\ell) d\ell. \tag{10}$$

At $e = 0$, this derivative is negative. (From (8) and the assumption that $x(\ell) > \ell + c$ for $\ell \in H$, it follows that $u'(w - \pi - \ell + x(\ell) - c) < \bar{u}'$ for $\ell \in H$.) Therefore, a modified contract with negative e increases expected utility, contradicting the hypothesized optimality of the original contract. The same argument shows that there cannot exist a set of losses with lower coverage than $\ell + c$ when coverage is optimal.

Step 2: The same coverages are optimal when the incentive compatibility constraint is taken into account. It will now be demonstrated that, under the optimal contract, the level of coverage is $x(\ell) = \ell + c$ for all losses that are demonstrated and is determined by (7) for all losses that are not demonstrated. Step 1 showed that these coverages are optimal *taking as given* whether losses are demonstrated. For such coverage to be incentive compatible, there are two requirements.

First, for all losses that are not demonstrated under the optimal contract, it must be that (3) does not hold—that is, that insureds do not have an incentive to demonstrate such losses. This requires that $x(\ell) - c < x$. To guarantee this result, set $x(\ell) = x$ for any ℓ that it is not optimal to demonstrate. Then insureds with such losses would get the same coverage whether or not they demonstrate their losses, so they would never spend c to do so.[10]

Second, for all losses that are demonstrated, condition (3) must hold. Suppose that for demonstrated losses $x(\ell) = \ell + c$, which is optimal taking as given that ℓ is demonstrated. Then the left side of (3) equals ℓ, so individuals would demonstrate their losses whenever $\ell \geq x$. Therefore, the incentive compatibility constraint prevents demonstration of losses only when $\ell < x$. It will now be shown that the optimal contract does not involve demonstrating any such losses.

Suppose that a set of losses less than x are demonstrated under the optimal contract. For this to be incentive compatible, it must be either that $x(\ell) > \ell + c$ or that x is less than the value given by (7). With regard to the former possibility, it has been shown that reducing $x(\ell)$ for all such losses will increase expected utility, as long as the losses continue to be demonstrated. Therefore, if $x(\ell) > \ell + c$, it must be that the incentive compatibility constraint is binding, which implies that $x(\ell) = x + c$ for all such ℓ. The argument proving Proposition 1, however, indicates that it is not optimal to demonstrate such losses: if instead they are not demonstrated, individuals' ex post utility is unaffected, but premiums are reduced by c multiplied by the probability that there is a loss ℓ in the specified set.

Alternatively, x may be below the value given by (7). But reducing x will decrease expected utility as well. The derivative in (6) took as given which losses were demonstrated. Allowing for the effect on which losses are demonstrated, the derivative becomes

$$\frac{dEU}{dx} = -\frac{d\pi}{dx} \cdot \bar{u}' + p \int_{\mathcal{L}} u'(w - \pi - \ell + x) f(\ell) d\ell$$
$$+ pf(\tilde{\ell})[u(w - \pi - \tilde{\ell} + x) - u(w - \pi - \tilde{\ell} + x(\tilde{\ell}) - c) + (x(\tilde{\ell}) - x)\bar{u}'], \quad (11)$$

where $\tilde{\ell}$ is the value of ℓ for which the incentive compatibility constraint is binding at x. Because $x(\ell) = \ell + c$, $\tilde{\ell} = x$. Thus, the terms involving the difference in utility when an individual for whom (3) is binding changes groups have a combined value of zero. Because premiums fall due to the marginal type not demonstrating losses, the final term is positive $(x(\tilde{\ell}) - x = c)$. The first two terms of (11) taken together are positive (zero) if x is less than (equal to) the value given by (7). Therefore, when x is less than the value given by (7), expected utility is increasing in x, so a lower value of x than that given by (7) cannot be optimal.

Step 3: Which losses are demonstrated under the optimal contract. It remains to characterize the optimal set $\mathcal{L}$ (the losses that are not demonstrated). First, $\mathcal{L}$ is not empty. Such a contract is dominated by one in which $x = x(0) = c$ and individuals for whom $\ell < c$ do not demonstrate their ℓ. Such a modified contract is feasible, as individuals who demonstrate such ℓ can be given coverage of $x(\ell) = x = c$. In states with $\ell < c$, individuals' wealth exceeds that under the contract involving ℓ always being demonstrated. (For the latter contract, coverage is $\ell + c$, but individuals spend c, leaving them with a net of ℓ.) Moreover, the insurance premium is lower under the modified contract.

Second, if $\ell < x$, $\ell \in \mathcal{L}$. This follows from the incentive compatibility constraint and the demonstration that $x(\ell) = \ell + c$.

Third, let $\hat{\ell}$ denote the infimum of $\ell \notin \mathcal{L}$; if $\ell > \hat{\ell}$, then $\ell \notin \mathcal{L}$. That is, the optimal rule is such that all losses above a threshold and none below it are demonstrated. Suppose otherwise. Then there exists a set of losses N that are not demonstrated and a set of losses D that are demonstrated such that $\inf(N) \geq \sup(D)$ (with strict inequality everywhere except possibly for a boundary point, which is ignored in the analysis to follow). Let N' and D' denote any subsets of N and D, respectively, that have equal (nonzero) measure.

Consider the alternative contract in which losses in D' are not demonstrated and thus receive coverage of x instead of $\ell + c$, and losses in N' are demonstrated and receive coverage equal to $\tilde{x}(\ell)$ rather than x, where $\tilde{x}(\ell) = \ell + c - k$, and k is such that total insurance payments for losses in D' and N' are the same as under the original contract. To ensure incentive compatibility, set $x(\ell) = x$ for all $\ell \in D'$, so losses in D' will not be demonstrated. It need not be true, however, that all losses in N' will be demonstrated. If some would not be, simply redefine N' by raising its lower boundary until the incentive compatibility constraint is satisfied for all losses in N'.[11]

Insurance premiums are unaffected by this change in the contract because the alternative contract was constructed so that total insurance payments are the same. In addition, because N' has the same measure as D' (or less, if N' was modified as described in the preceding paragraph), total expenditures on demonstration are the same (or less) under the alternative contract. The available net compensation—insurance payments minus demonstration costs—beyond x for individuals with losses in the identified subsets is thus the same (or greater) under the alternative contract. This additional net compensation beyond x is distributed entirely to those with losses in D' under the original contract and entirely to those with losses in N' under the alternative contract. (In both cases, the distribution among those with losses in the subset is optimal, as marginal utility is equalized within the subset.) When compensation is only x, utility is lower for all $\ell \in N'$ than for all $\ell \in D'$ (losses are higher in N'); hence, the marginal utility of wealth is higher in N'. As a result, utility increases more if the additional net compensation goes entirely to those with losses in N' than if it goes to those with losses in D'.[12] Thus, expected utility is higher under the alternative contract, contradicting the possibility that the hypothesized sets N and D exist under the optimal contract.

Fourth, the optimal threshold $\hat{\ell}$ may be characterized. Because the set $\mathcal{L}$ has the simple form just demonstrated, one can take the derivative of expected utility (1) with respect to $\hat{\ell}$ (substituting for π from (2) and using the facts that x is set optimally (7) and $x(\ell) = \ell + c$ for demonstrated losses) to obtain

$$\frac{dEU}{d\hat{\ell}} = pf(\hat{\ell})[u(w - \pi - \hat{\ell} + x) - u(w - \pi) + (\hat{\ell} + c - x)\bar{u}']. \tag{12}$$

The first two terms in brackets indicate the change in utility when individuals with losses of $\hat{\ell}$ no longer demonstrate their losses, and the third term is the change in the insurance premium, weighted by the expected marginal utility of wealth. If there is an interior solution, the optimal $\hat{\ell}$ will be the value of ℓ for which expression (12) equals zero and, as already established, $\hat{\ell} \geq x$. An extreme solution in which all ℓ are demonstrated has been ruled out. If (12) is negative for all $\hat{\ell}$, in which case demonstration is never desirable, let $\hat{\ell} = w$.

Moreover, $\hat{\ell} > x$. The third term of (12) is positive because $\hat{\ell} \geq x$. For an interior solution, the first two terms together must be negative, so $\hat{\ell} > x$. (If (12) is negative for all ℓ, so no losses are demonstrated, $\hat{\ell} = w$. In this case, it is trivial that $\hat{\ell} > x$.) ■

Remarks: In principle, it may be desirable for low values of ℓ to be demonstrated. In particular, unless c is too high there will exist a critical value of $\hat{\ell}$ that is less than $x - c$ such that (12) equals zero, and it would increase expected utility if all lower values of ℓ were demonstrated.[13] Step 2 of the proof shows, however, that the only incentive-compatible ways to induce individuals to demonstrate low losses would involve coverages that reduce expected utility.

There remains the option of demonstrating all losses.[14] Step 3 of the proof demonstrates, however, that this is not optimal. Nonetheless, if the insurance company were permitted to demonstrate all losses, it might have an incentive to do so if coverage for all losses would then be $x(\ell) = \ell + c$. Such an incentive exists if the resulting expected payments to individuals who do not choose to demonstrate ℓ (the mean of $x(\ell)$ over the set $\mathcal{L}$) is less than x.[15] An illustration in the footnote suggests that this is possible.[16]

2.4. *An example*

To illustrate how the cost of establishing the amount of losses affects the optimal form of an insurance contract, consider the following simple example. Individuals have a 1 percent chance of suffering a loss. If they suffer a loss, it has an equal probability of being \$1,000 or \$3,000. Individuals have constant-absolute-risk-aversion utility functions of the form $u(y) = 1 - e^{-\eta y}$, where η is the coefficient of risk aversion. The discussion will consider cases in which η is .0001 and .0005 (corresponding to an individual being indifferent between a gamble involving an equal probability of gaining and losing \$1,000 and suffering a certain loss of \$50 and \$241, respectively).

If the contract takes the simple form in which individuals receive optimal coverage taking as given whether the loss is demonstrated, coverage would be \$1,000 for those who are silent and \$3,000 + c for those who spend c to demonstrate a loss of \$3,000. If the contract forbids individuals from demonstrating the amount of their loss (that is, if they receive the optimal amount for those who are silent even if they demonstrate the high loss), all individuals who suffer a loss (whether \$1,000 or \$3,000) would receive coverage of \$2,050 when $\eta = .0001$ and \$2,240 when $\eta = .0005$. A third option to be considered is that individuals do not purchase any insurance.

For the case in which $\eta = .0001$, individuals would prefer the contract in which the high loss is demonstrated rather than that in which it is not as long as the cost does not exceed \$100. That is, when the cost of demonstrating a high loss exceeds \$100, the optimal form of the contract involves individuals being deterred from demonstrating that their loss is high (with the result that they are undercompensated, and that since the same coverage applies to high and low losses, individuals with a low loss are overcompensated). In addition, a contract in which high losses are demonstrated is dominated by no insurance coverage at all when the cost of demonstration exceeds \$546. (No insurance never dominates the contract in which high losses are not demonstrated because no administrative cost is incurred under such a contract.)

For the case in which $\eta = .0005$, individuals would prefer the contract in which the high loss is demonstrated rather than that in which it is not when the cost does not exceed \$485. In addition, a contract in which high losses are demonstrated is dominated by no insurance coverage at all only when the cost of demonstration exceeds \$4,179.

Further analysis reveals, not surprisingly, that allowing demonstration of high losses is more likely to be desirable (that is, remains preferable for higher levels of the demonstration cost) the greater is individuals' risk aversion (and thus the more valuable is finely tuned insurance coverage) and the greater is the range of possible losses.[17] (When there are more than two possible levels of loss, the optimal contract, recall, generally involves demonstration of the highest losses but not others.)

3. Discussion

3.1. Generality of the results

The model employs the assumption that the existence of losses is determined costlessly while the extent of losses can be demonstrated only at a cost. This assumption roughly characterizes many losses covered by first-party insurance, such as disabilities from workplace or automobile accidents and property damage caused by fires or storms.

The model can be extended to the case in which it is also costly to establish the existence of a loss. Suppose that c is the cost of establisihng both the existence and amount of a loss. Using the proof technique of Proposition 2, it can be shown that the optimal contract has the following characteristics. First, without loss of generality, $x = 0$. (The amount x is received in all states in which a loss is not demonstrated; by setting $x = 0$, reducing π by x, and reducing $x(\ell)$ by x, net insurance payments and receipts are the same in all states as when $x \neq 0$.) Second, the optimal contract would involve coverage for demonstrated losses of $\ell + c - k$ rather than $\ell - c$. (It is not optimal to provide complete coverage for demonstrated losses because undemonstrated losses are uncompensated.[18]) Third, the condition for the optimal threshold becomes

$$\frac{dEU}{d\hat{\ell}} = pf(\hat{\ell})[u(w - \pi - \hat{\ell}) - u(w - \pi - k) + (\hat{\ell} + c - k)\bar{u}']. \qquad (12')$$

Finally, $\hat{\ell} > k$, so that individuals with losses in the interval $\ell \in [k, \hat{\ell})$ would have an excessive incentive to demonstrate their losses if coverage would be $\ell + c - k$ rather than zero.

Another difference between practice and the model is that the process of establishing losses often involves matters of degree. An insurance company may use a claims adjuster to provide an inexpensive appraisal. The claimant may then have the option of accepting this initial judgment or invoking a more expensive appeals process in the hope of receiving more favorable treatment. Alternatively, within a dispute resolution process, an adjudicator must determine how extensive a demonstration it is willing to consider.

The model can be extended to such cases by adding the assumption that there is a fixed cost of claims that is initially incurred; insureds then decide whether to spend c to demonstrate that their loss is higher than the initial estimate. An optimal contract would involve insured individuals being entitled to a payment after an initial determination, while providing them with the opportunity to demonstrate that their losses are significantly higher. This suggests that an insurance claims process would have relatively cheap initial assessments, with the option of more elaborate consideration prompted by individual appeals. At the same time, appeals should not freely be permitted for all individuals who may in principle be entitled to coverage exceeding their initial awards, even if claimants bear the full costs of such appeals, because the ex post private gain exceeds the ex ante benefit of improved risk allocation. In the optimal contract, inefficient appeals are discouraged by a rule under which the tribunal essentially refuses to hear small disagreements, or, equivalently, refuses to consider a claimant's evidence unless it establishes that the initial award involved significant error.

3.2. Existing insurance arrangements

Private insurance contracts appear to have at least some properties of the optimal contract. It is common for claims to be established initially through inexpensive procedures (when compared, say, to tort suits when the extent of damages is contested). Claimants, then, are often permitted an appeal, perhaps to an arbitrator who will make a binding determination. What is less certain is how coverage is determined at each stage. The analysis here suggests two features. First, the coverage awarded at the initial stage should be the optimal coverage for those who will not appeal (and whether a claimant will appeal cannot be observed at the time of the determination); thus, such payments should be less than what would be optimal if no appeal were permitted. Second, when an appeal is taken, claimants should receive higher payments only when their demonstrated loss exceeds the initial award by a nontrivial, and perhaps substantial, amount, so as to avoid what would otherwise be an excessive incentive to appeal claims. Neither feature may be explicit in existing insurance arrangements, although it is possible that claims adjusters and arbitrators behave in this manner in any event. If these features are not present in practice, expenditures to establish claims are likely to be excessive.[19]

Procedures for public insurance programs, such as Social Security Disability Insurance, have been quite controversial. In particular, unsuccessful claimants have challenged arrangements in court on the ground that available procedures provide insufficient opportunity to challenge initial determinations (see, for example, *Mathews v. Eldridge*, 424 U.S. 319 (1976)). While the present analysis does not indicate how much accuracy is best or how it can best be produced, it suggests that, even in an optimally designed system, there will be individuals with valid claims who would wish to demonstrate their validity but who would not be induced or permitted to do so.

Acknowledgments

I am grateful to A. Mitchell Polinsky, Steven Shavell, Kathryn Spier, and the referees for comments.

Notes

1. The problem of verifying claims does not always have this feature. For example, with medical insurance, coverage is usually for expenditures that are reported both to the individual and to the insurer by the provider of medical care. For other work involving ex post verification, see, e.g., Mookherjee and Png [1989], Spier [1992], and Townsend [1979].
2. The discussion abstracts from insurer risk aversion, moral hazard, and other sources of administrative cost (see, e.g., Arrow [1963], Raviv [1979], Shavell [1979]). Incorporating these previously studied dimensions obviously would affect the optimal contract, but the core features identified in this paper would remain. For example, with moral hazard, optimal coverage when the amount of the loss is known would be incomplete, but it still would be desirable for the level of coverage to increase with the amount of the loss. This paper also ignores the possible effects of claims procedures on the incentive to file claims, including frivolous claims.
3. This work determines optimal coverage when there are fixed costs either for obtaining coverage or for all claims, or when costs are a continuous function of the amount of the insurance payment.
4. Blazenko [1985], building on Townsend [1979], considers the optimality of a deductible when there are verification costs but does not address the issue of incentives to make claims. Closest to the present result is that of Huberman, Mayers, and Smith [1983], who show that a vanishing deductible is optimal when administrative costs exhibit economies of scale. For an informal discussion in related contexts of the issues addressed in this article, see Kaplow [1994].
5. More complicated contracts, such as the type that might be designed using the revelation principle, see Myerson [1979] and the articles cited in note 1, are not considered. Attention is confined to examining optimality within the general framework of claims processes—public and private—that actually are used. (The optimal contract described in Proposition 2 is, however, analogous to the best contract that could be implemented with a deterministic audit policy of the sort examined in Reinganum and Wilde [1985] and Townsend [1979].)
6. $$\bar{u}'_{\pounds} = \frac{\int_{\pounds} u'(w - \pi - \ell + x) f(\ell) d\ell}{\int_{\pounds} f(\ell) d\ell}.$$
7. It will be apparent that if c were borne by the insurance company, $x(\ell)$ would be lower by c and expected utility would be unaffected.
8. $$\begin{aligned}\bar{u}' &= (1 - p)u'(w - \pi) + p \int_{\pounds} u'(w - \pi - \ell + x) f(\ell) d\ell \\ &\quad + p \int_{\sim\pounds} u'(w - \pi - \ell + x(\ell) - c) f(\ell) d\ell \\ &= (1 - p)u'(w - \pi) + p\bar{u}' \int_{\pounds} f(\ell) d\ell + p \int_{\pounds} u'(w - \pi) f(\ell) d\ell.\end{aligned}$$

   It is immediate that $u'(w - \pi) = \bar{u}'$.

9. Here and below, it will be assumed that sets that are considered are not of measure zero; how losses in sets of measure zero are treated does not affect expected utility.
10. The optimal contract specified in Proposition 2 is not unique, because $x(\ell)$ could be set at any value less than $\ell + c$. Since insureds will never demonstrate any such ℓ, they never will receive $x(\ell)$, so its precise value is unimportant.
11. As the lower boundary of N' is raised, k will decrease because the same total amount of insurance payments are available for fewer states. As the lower boundary approaches the upper boundary, coverage for losses will approach infinity, so there will exist a lower boundary for which the incentive-compatibility constraint is satisfied as an equality at the lower boundary.
12. Let z denote the amount of additional net compensation available for D'. (The amount available for N' may be greater, but this additional amount can be ignored because the result holds even if it is not.) The value of dEU/dz must be higher under the alternative contract than under the original contract for all z. (Under both contracts, z is allocated to equalize marginal utility for those with demonstrated losses in the relevant subset. When z is the same for both contracts, therefore, it must be that marginal utility is higher for all $\ell \in N'$ than for all $\ell \in D'$—recall that k is positive so that when all of z is allocated, there is undercompensation in N' but not in D'. Hence, any increase in z raises utility more under the alternative contract.) Therefore, the integral of dEU/dz from zero to the amount of additional net compensation available for D' is greater under the alternative contract. This implies that

$$\int_{N'} u(w - \pi - k)f(\ell)d\ell - \int_{N'} u(w - \pi - \ell + x)f(\ell)d\ell >$$
$$\int_{D'} u(w - \pi)f(\ell)d\ell - \int_{D'} u(w - \pi - \ell + x)f(\ell)d\ell.$$

As a result,

$$p\int_{D'} u(w - \pi - \ell + x)f(\ell)d\ell + p\int_{N'} u(w - \pi - k)f(\ell)d\ell >$$
$$p\int_{N'} u(w - \pi - \ell + x)f(\ell)d\ell + p\int_{D'} u(w - \pi)f(\ell)d\ell.$$

The left side is expected utility when there is a loss in either subset under the alternative contract and the right side is expected utility when there is a loss in either subset under the original contract.
13. When $\hat{\ell} < x$, the bracketed expression in (12) is increasing in $\hat{\ell}$. And when $\hat{\ell} < x - c$, the third term is negative, so there will exist a critical value of $\hat{\ell}$ for which the first two terms together are positive and (12) equals zero (unless the expression is negative throughout the interval $[0, x - c]$, in which case it would not be desirable to demonstrate any low losses).
14. The demonstration of all ℓ could be accomplished either by inducing all individuals to demonstrate ℓ, as by setting $x \leq 0$ and $x(\ell) = \ell + c$, or by having the contract specify that ℓ be demonstrated in all cases (if the insurance company can independently spend c to demonstrate ℓ). Observe that the insurance company is unable to act strategically in light of whether individuals choose to demonstrate ℓ. If an individual demonstrates ℓ there is nothing for the insurance company to do. If an individual does not demonstrate ℓ, the insurance company can do nothing or can demonstrate ℓ itself. The former corresponds to the case in which the insurance company cannot act ex post to demonstrate ℓ and the latter amounts to always demonstrating ℓ.
15. This result does not depend on the fact that, in the model, individuals rather than the insurance company pay c. If the insurance company paid c, the incentive to demonstrate all ℓ would exist when the mean of the $x(\ell)$ over $\mathcal{L}$ was less than x by at least c, but the values of the $x(\ell)$ would be less than otherwise by c (see note 7).
16. Consider the following case (which does not precisely fit the assumptions of the model concerning the form of $u(\cdot)$ and that ℓ is distributed as a continuum, but clearly indicates the existence of more complex examples that would):

$$w = 100$$

$$u'(y) = \begin{cases} \infty, & \text{for } y \in [0, 90) \\ 1 - .00001y, & \text{for } y \geq 90 \end{cases}$$

$$p = .01$$

$$\ell = \begin{cases} 0, & \text{with probability } .5 \\ 100, & \text{with probability } .5 \end{cases}$$

$$c = 10$$

It is apparent that the optimal scheme if the insurance company is not permitted to require demonstration of ℓ involves some x in the interval [90,100) and the loss never being demonstrated. (If the loss is demonstrated, $x(0) = 10$ and $x(100) = 110$.) The insurance company, however, would have an incentive to require that ℓ be demonstrated, for the expected insurance payment when ℓ is demonstrated is 60, while the payment when ℓ is not demonstrated is at least 90. (If the insurance company must pay the cost of 10, the $x(\ell)$ would each be lower by 10—see note 7—and the incentive would be the same.)

17. Increasing the magnitude of the losses, keeping their dispersion unchanged, need not affect the relative desirability of demonstrating high losses because the amount paid to those who do not demonstrate their losses increases by the amount of any upward shift in the amount of the losses. Of course, if losses are greater, it is more likely that a contract that allows excessive demonstration of losses will be preferred to no insurance at all.
18. Optimal coverage for demonstrated losses equates the marginal utility of wealth in those states with the average marginal utility of wealth for all states. Unlike in the original model, the average marginal utility of wealth for undemonstrated losses exceeds the average marginal utility of wealth for states with no losses because it is not possible to give coverage of x when losses are incurred but not demonstrated without also giving x to individuals who suffer no losses.
19. The analysis here proves that incentives are excessive even when, ex post, the claimant bears all the costs of demonstrating claims; if some of the costs are borne by the insurer, incentives are even more excessive. There is, however, a factor that in some contexts may make these features inefficient: the optimal contract described herein may encourage insureds to increase their losses in order to qualify for full coverage (see Huberman, Mayers, and Smith [1983]).

References

ARROW, Kenneth J. [1963]: "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, 53 (December), 941-973.

BLAZENKO, George [1985]: "Optimal Indemnity Contracts," *Insurance: Mathematics and Economics*, 4, 267–278.

GOLLIER, Christian [1987]: "Pareto-Optimal Risk Sharing with Fixed Costs per Claim," *Scandinavian Actuarial Journal*, 13, 62–73.

HUBERMAN, Gur, MAYERS, David, and SMITH, Clifford W. [1983]: "Optimal Insurance Policy Indemnity Schedules," *Bell Journal of Economics*, 14 (Autumn), 415–426.

KAPLOW, Louis [1994]: "The Value of Accuracy in Adjudication: An Economic Analysis," *Journal of Legal Studies*, 23 (January), 307-401.

MOOKHERJEE, Dilip and PNG, Ivan [1989]: "Optimal Auditing, Insurance, and Redistribution," *Quarterly Journal of Economics*, 104 (May), 399–415.

MYERSON, Roger [1979]: "Incentive Compatibility and the Bargaining Problem," *Econometrica*, 47, 61–74.

RAVIV, Artur [1979]: "The Design of an Optimal Insurance Policy," *American Economic Review*, 69 (March), 84–96.

REINGANUM, Jennifer F. and WILDE, Louis L. [1985]: "Income Tax Compliance in a Principal-Agent Framework," *Journal of Public Economics*, 26, 1–18.

SHAVELL, Steven [1979]: "On Moral Hazard and Insurance," *Quarterly Journal of Economics*, 93 (February), 541-562.

SHAVELL, Steven [1978]: "Theoretical Issues in Medical Malpractice," in *The Economics of Medical Malpractice*, Simon Rottenberg (ed.), American Enterprise Institute, Washington, D.C., 35–64.
SOCIAL SECURITY ADMINISTRATION [1991]: Annual Statistical Supplement, *Social Security Bulletin.*
SPIER, Kathryn E. [1992]: "Incomplete Contracts and Signalling," *Rand Journal of Economics*, 23 (Autumn), 432–443.
TOWNSEND, Robert M. [1979]: "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, 21, 265–293.