European Economic Review 11 (1978) 321–341. © North-Holland Publishing Company

THE EFFICIENCY OF INSURANCE MARKETS*

Hans-Werner SINN

University of Mannheim, 68 Mannheim, Federal Republic of Germany

Received January 1978, final version received June 1978

The paper provides a systematic analysis of various types of behaviour changes induced by insurance, which all tend to increase the insured losses. The welfare implications of these changes as well as the means to improve the working of insurance markets are studied.

1. Introduction

The task of this paper is to examine whether insurance markets can be expected to work efficiently and, if necessary, how they could be improved. Our analysis concentrates on the role of the insured parties who are assumed to choose among insurance contracts offered to them by the company and, simultaneously, among activities which enable them to change the loss distributions which the company would bear, at least to some extent. It can often be observed that these activities are selected in such a way that after buying insurance, losses tend to be higher than before. We shall study systematically four important mechanisms for this phenomenon (sections 2.2.3, 2.2.4, 3 and 4) and attempt to evaluate them from the welfare point of view. Surprisingly enough, the analysis will show that it is not always true that the insurance induced increase in losses indicates misallocation, as one could suppose at first glance and as it is suggested by the term 'moral hazard' which is sometimes used in this connection.

2. The welfare gain from insurance under alternative rating systems

It is an old argument of insurance theory whether the companies should follow an equivalence or community rating scheme. In this section we shall

*The paper has been developed from an earlier version published by the author in German: Sinn (1977). Permission of the editor is acknowledged.

point out which allocation will result on the insurance market under both alternatives.[1]

2.1. The risk theoretic model

Since we concentrate on the part of the insured, we refer to a very simple model of the insurer: The company is risk neutral and receives a premium which is just enough to cover the estimated mathematical expectation of claims and proportional costs, since there is competition or a government supervision of the market. This assures that the company is indifferent towards activities of an insured person, such that we can state an increase in welfare whenever the latter can be made better off.

Each insured is globally risk averse and trys to maximize expected utility of period income Y. For an opportunity set of probability distributions which all belong to a linear class,[2] this means that his goal can also be expressed as[3,4]

$$\max U(\mu,\sigma),$$
$$\frac{\partial U}{\partial \mu}>0, \qquad \frac{\partial U}{\partial \sigma}<0,$$
$$\left.\frac{\mathrm{d}\mu}{\mathrm{d}\sigma}\right|_U \begin{cases} =0 & \text{if} \quad \sigma=0, \\ >0 & \text{if} \quad \sigma>0, \end{cases}$$
$$\left.\frac{\mathrm{d}^2\mu}{\mathrm{d}\sigma^2}\right|_U >0, \tag{1}$$

with

$$\mu \equiv \mu(Y) \quad \text{and} \quad \sigma \equiv \sigma(Y).$$

Without insurance the decision maker, i.e., the potential customer, is faced with a subjective probability distribution of the type

$$Y=b-C, \qquad C \geqq 0, \tag{2}$$

[1] We do however not discuss the problem of adverse selection which arises under community rating if the companies require different tariffs. For this problem, see e.g. Gürtler (1936), Pauly (1974) and Rothschild–Stiglitz (1976).

[2] A linear distribution class is defined such that all its members have the same standard distribution $Z=[Y-\mu(Y)]/\sigma(Y)$ from which they can, figuratively speaking, be derived if one shifts it by the amount $\mu(Y)$ and stretches it by the amount $\sigma(Y)$ (cf. fn. 3).

[3] $\mu(Y)=\sum_{i=1}^{n} p_i y_i$ and $\sigma(Y)=\left\{\sum_{i=1}^{n} p_i[y_i-\mu(Y)]^2\right\}^{1/2}$, where $y_1,\dots,y_n$ are the variates of the random variable Y, and $p_1,\dots,p_n$ the corresponding probabilities. Note that $\mu(a+bY)=a+b\mu(Y)$ and $\sigma(a+bY)=|b|\sigma(Y)$.

[4] See Tobin (1958) and e.g. Schneeweiß (1967, pp. 126–128).

where b denotes his *normal income* and C his stochastic loss distribution. The normal income is the income net of all loss prevention costs if the decision maker is lucky enough to avoid any loss. The losses are in principle monetary, but we can also include non-monetary losses if we define C, such that it denotes the amounts of money which, if provided unconditionally, are sufficient to compensate the decision maker for the damages in question. We exclude the possibility of a default risk, such that without insurance the decision maker has to bear all losses himself.[5]

In general we cannot assume that the decision maker is faced with a given distribution of Y. Instead he will normally be able to manipulate both, b and C. Even if, in this section, we disregard the possibility of wilful damage, there remain such possibilities as to install burglar alarms and sprinklers, to built-in fire- and theftproof safes, to utilize fireproof building materials, to substitute dangerous production techniques by others with a higher degree of safety, to diminish danger by better controls, etc. Common to all these activities is the fact that they reduce the possible losses, C, but also bring about an increase of prevention costs in a broader sense, i.e., diminish b.

We consider two types of offers the insurance company makes to his customer: First, an 'all or nothing' offer, where he can either buy full coverage insurance for a given premium or stay uninsured.[6] Second, an offer where he can also choose arbitrary degrees of coverage (θ) between these extremes.[7] Then, in general, the decision maker's probability distribution is given by

$$Y=b-C+\theta\{C-\pi\mu^*[\mu(C)]\}, \qquad \pi \geqq 1, \quad \begin{matrix} \theta=0,1,\text{or} \\ 0 \leqq \theta \leqq 1, \end{matrix} \tag{3}$$

where π denotes the premium loading required and μ^* the expected loss as estimated by the company. With $\mu^*=\mu^*[\mu(C)]$ we allow for the possibility that there is a functional relationship between the expectation of the company and that of the insured. For the time being it is assumed that the monetary compensation from the company, θC, is provided regardless for what purpose the insured utilizes it. This assumption will be abandoned in section 4.

In order to assure that all obtainable distributions of Y belong to the same linear class, we assume that $C=\lambda Z_C$ where Z_C denotes an arbitrary basic

[5] According to this assumption we can only allow for liability risks small enough to prevent the insured's wealth from becoming negative.

[6] Insurance under this kind of offer was first studied by Barrois (1834) on the basis of the expected utility criterion. See Rothschild–Stiglitz (1976) for a theoretical explanation of the great practical relevance of this offer.

[7] For the case of a given risk this kind of offer was first analysed by Mossin (1968).

distribution and λ is the parameter under control.[8] This assumption allows us to refer to the μ, σ criterion introduced above. If we characterize the loss prevention policy of the decision maker by the parameters

$$\mu_M \equiv \mu(b-C) = b - \mu(C), \tag{4}$$

and

$$\sigma_M \equiv \sigma(b-C) = \sigma(C), \tag{5}$$

and if we note further that

$$\mu(C) = k\sigma_M, \tag{6}$$

where

$$k \equiv \mu(C)/\sigma(C) = \mu(Z_C)/\sigma(Z_C) = \text{const.},$$

then the mean of the probability distribution characterized in (3) is

$$\mu = \mu_M + \theta[k\sigma_M - \pi\mu^*(k\sigma_M)], \tag{7}$$

and the standard deviation is

$$\sigma = (1-\theta)\sigma_M. \tag{8}$$

The decision maker's task is to manipulate these parameters via μ_M, σ_M and θ such that $U(\mu, \sigma)$ is maximal. Fortunately this work can be simplified somewhat, as [cf. (1)] he will for any given σ_M and θ trivially choose the highest possible value for μ_M. Since this value, call it $\bar{\mu}_M$, is a function of σ_M alone, we can substitute

$$\mu_M = \bar{\mu}_M(\sigma_M) \tag{9}$$

as a functional description of the efficiency frontier within (7), and the parameters under control reduce to θ and σ_M. For simplicity's sake we assume $\bar{\mu}_M(\sigma_M)$ to be twice differentiable in the relevant range with $\bar{\mu}''_M(\sigma_M) < 0$.

For a graphical illustration of the decision problem, see fig. 1. There the shaded area represents the *original opportunity set* (M) of loss prevention

[8] If all probability distributions Y belong to the same linear class, $Z = [Y - \mu(Y)]/\sigma(Y)$ should be independent of b, λ, and θ. This is indeed the case since inserting (3) gives $Z = -[Z_C - \mu(Z_C)]/\sigma(Z_C)$.

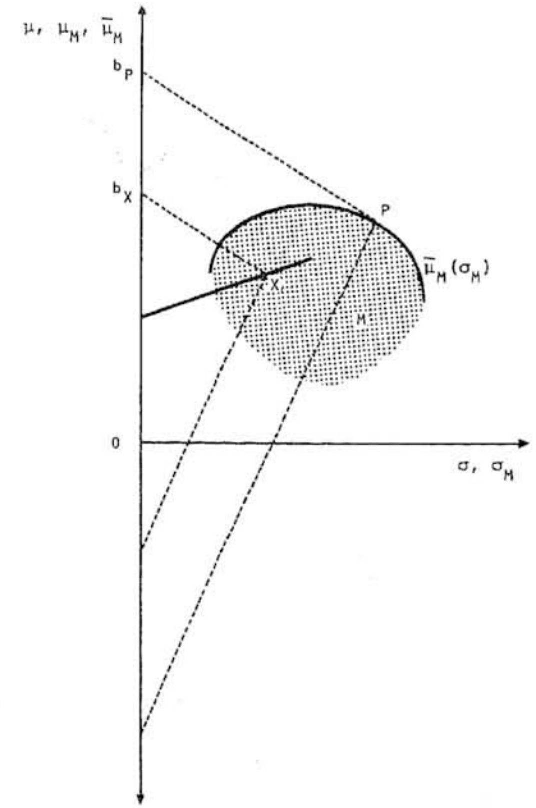

Fig. 1. The opportunity locus of the insurance customer.

policies, which are each characterized by particular values of μ_M and σ_M. For each point of the opportunity set, there exists what we call an *insurance line* connecting this point with the ordinate. The insurance line is the locus of points representing those probability distributions of income which can be obtained by alternative degrees of insurance coverage, given the loss prevention policy. However, whereas the whole length of a particular insurance line is relevant if the company allows for $0 \leqq \theta \leqq 1$, the decision maker can only reach its two ends if the company offers $\theta = 0, 1$. From (7) and (8) the explicit equation for the insurance line is

$$\mu = \mu_M + k\sigma_M - \pi\mu^*(k\sigma_M) + \sigma k\left(\pi\frac{\mu^*(k\sigma_M)}{k\sigma_M} - 1\right). \tag{10}$$

The line will normally have a positive slope since $k > 0$, $\pi \geqq 1$, and $\mu^* = k\sigma_M = \mu(C)$ if the company estimates the same expected loss as the insured does.

Including the insurance company's offer, the whole opportunity set available to the decision maker consists of the shaded area and the admissible parts of all insurance lines belonging to the points within this area.

It will be of great help if we try to obtain a plastic impression of the shape of the probability distributions within the opportunity set. Note that for a linear distribution class the distance between the mean and any characteristic distribution parameters like e.g. the mode, the percentiles, or the boundaries are proportional to the standard deviation. Thus the shape of the probability distribution represented by a point in the μ, σ diagram can be derived, figuratively speaking, by projecting the common standard distribution rectangularly towards the ordinate with the point in question as the projection centre. This is illustrated for points P and X in fig. 1. The upper rays represent the upper boundaries or normal incomes, b_P and b_X, of the corresponding probability distributions. Since it follows from (4) and (6) that

$$\mu_M = b - k\sigma_M, \tag{11}$$

these rays have the slope $-k$. Analogously the lower rays characterize the lower boundaries of the probability distributions in question.

According to (9) we know already that only the upper boundary of the original opportunity set (shaded area) is efficient.[9] Let us compare all the probability distributions which are represented by the points on this efficiency frontier. Obviously the points to the left of point P, which is characterized by $\bar{\mu}'(\sigma_M) = -k$, such that an upper ray is there tangent to the opportunity set, show those probability distributions to which we referred in the above examples:[10] A reduction of $\sigma_M = \sigma(C)$, which, according to (6), implies a proportional reduction of $\mu(C)$, has to be 'paid' by a reduction of normal income, b, due to the costs of loss protection. However, at point P normal income is maximal indicating that prevention costs are zero. To the right of P it would even be necessary for the decision maker to bear costs, i.e., to accept a lower b, if he tried to enlarge the loss distribution above its 'natural shape'.

2.2. *The results*

Having illustrated the nature of the decision problem so far, we are now prepared to study the decision maker's behaviour under alternative assum-

[9]This can be verified by aid of the insurance line eq. (10), since for a given σ_M it is obviously optimal to choose the highest possible line, i.e., the line where μ_M is maximal.

[10]We should not be surprised that points right to P and points below the efficiency frontier cannot be observed in reality, for, as will be shown below, all these possibilities are excluded if people behave in accordance with our model.

ptions concerning his opportunity set. To summarize the previous discussion, his task is

$$\begin{aligned} &\max_{(\theta, \sigma_M)} U(\mu, \sigma), \\ &\mu = \bar{\mu}(\sigma_M) + \theta[k\sigma_M - \pi\mu^*(k\sigma_M)], \\ &\sigma = (1-\theta)\sigma_M. \end{aligned} \tag{12}$$

Bearing in mind eq. (6), we obtain for the case of interior maxima[11] from $\partial U / \partial \sigma_M = 0$,

$$\bar{\mu}'_M(\sigma_M) - \theta k\left(\pi \frac{\partial \mu^*}{\partial \mu(C)} - 1\right) = \left.\frac{d\mu}{d\sigma}\right|_U (1-\theta), \tag{13}$$

and from $\partial U / \partial \theta = 0$,

$$k\left(\pi \frac{\mu^*}{\mu(C)} - 1\right) = \left.\frac{d\mu}{d\sigma}\right|_U. \tag{14}$$

The behavioural implications of these conditions will now be discussed.

2.2.1. *Loss prevention without insurance*

If $\theta = 0$ then (14) is irrelevant and according to (13), the decision maker chooses a loss prevention policy such that

$$\bar{\mu}'_M(\sigma_M) = \left.\frac{d\mu}{d\sigma}\right|_U. \tag{15}$$

This is illustrated in fig. 2, where an indifference curve as characterized by (1) is tangent to the original opportunity locus M at point S. Since $d\mu/d\sigma|_U > 0$ for $\sigma > 0$ the solution indicates that the decision maker will always exhibit some loss prevention activities. Moreover, he will even extend these activities up to a point, where the expected income is less than maximal.

2.2.2. *The welfare gain from insurance under given risks*

Suppose now that the decision maker has the possibility to buy insurance. If he did not change his loss prevention policy, then in fig. 2 the insurance

[11]Note that $\partial^2 U / \partial \theta^2 < 0$ is always true, and $\partial^2 U / \partial \sigma_M^2 < 0$ is true if $\theta k^2 \partial^2 \mu^* / \partial \mu^2(C) > \overline{\mu''} - (1-\theta) \cdot d^2\mu / d\sigma^2|_U$. We assume the latter, which is e.g. the case if $\mu^*[\mu(C)]$ is linear.

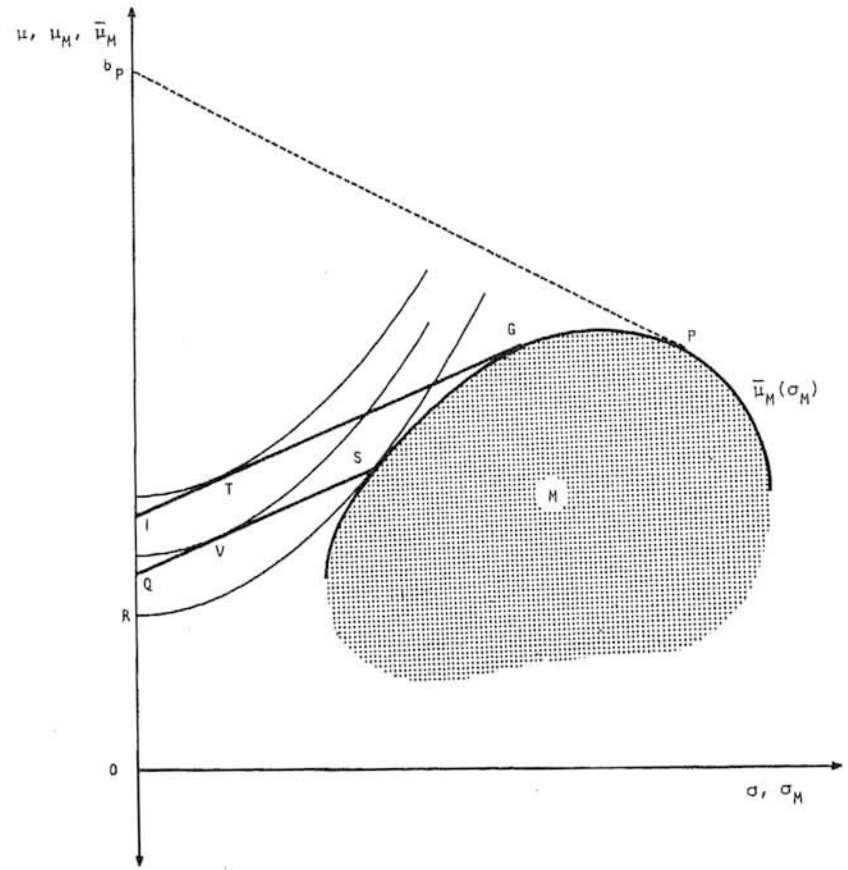

Fig. 2. The welfare gain under ideal insurance.

line $\overline{SQ}$ would be relevant for him. Thus, in the case of an 'all or nothing' offer he would buy full coverage insurance if Q is above R, where the indifference curve passing through S originates from the ordinate. Since the slope of the indifference curves is, except at the ordinate, everywhere positive this is always the case if the company requires a unitary premium loading factor π, such that according to (10) the insurance line is horizontal. However, even if there are managing costs of the company, such that $\pi > 1$ assures a positive slope of the insurance line, the decision maker would still buy insurance if the costs were not too high.

If variable degrees of coverage are allowed and consequently the whole range of the insurance line is feasible for the decision maker, then of course he cannot be worse off. In the case $\pi = 1$ he cannot improve his situation, since then (14) requires $\mathrm{d}\mu/\mathrm{d}\sigma|_U = 0$ which implies, according to (1) and (8), that $\theta = 0$. But, however, if $\pi > 1$ he can clearly become better off by choosing a degree of coverage between zero and unity as illustrated by point V in fig. 2.

Thus, remembering that by assumption the company is indifferent towards all that the insured person does, there is no doubt that insurance will bring about a welfare gain in the sense of the Pareto criterion, if the insured maintains his loss prevention policy without change. However, a quick look at (13) informs us that he will never behave in this way.

2.2.3. The loss prevention policy under equivalence rating

How the decision maker changes his loss prevention policy if he buys insurance depends on how the premium he has to pay varies with a change in his behaviour, i.e., in other words on how the expected loss, as estimated by the company, depends on the insured person's expectation. In this paragraph we assume

$$\mu^* = \mu(C). \tag{16}$$

This implies that both parties have the same information about the loss probability distribution and that moreover the company follows an equivalence rating scheme. The latter seems to be the case for many private insurance companies, but public insurance organizations often follow the solidarity principle, which is a striking contrast to equivalence rating.

The condition for an optimal loss prevention policy is (13). If the decision maker buys full coverage insurance either because the company makes an 'all or nothing' offer or because the premium is low enough such that the customer comes to this decision deliberately, then (13) becomes

$$\bar{\mu}'_M(\sigma_M) = k(\pi - 1). \tag{17}$$

If the decision maker chooses partial coverage, then not only (13) but also (14) must be satisfied. However, since $\mu^*/\mu(C) = \partial\mu^*/\partial\mu(C) = 1$ substitution of (13) into (14) gives again eq. (17). Thus the optimal loss prevention policy is independent of the particular kind of insurance contract.[12]

The condition (17) can easily be understood if we note that according to (10) and (6) all feasible insurance lines have, independently of σ_M, the slope $k(\pi - 1)$. So it is obvious that the decision maker chooses the highest possible insurance line, which is tangent to the original opportunity locus where (17) is satisfied. In fig. 2 the point of tangency is G.

[12]Moreover it is interesting to note that the loss prevention policy is also independent of the particular shape of the decision maker's indifference curves. This result reminds of Tobin's (1958) separation theorem for portfolio theory. Cf. also Ehrlich–Becker (1972, pp. 636 f.) who reached a similar conclusion for the insurance market in a different risk theoretic framework.

It is important to note that G is always right of S, representing the pre-insurance loss prevention activity, if insurance is purchased. The reason for this is that whenever the slope of the insurance line is equal to or greater than the slope of the efficiency frontier at point S, the decision maker's position deteriorates in comparison to S if he buys insurance. Thus under equivalence rating insurance induces a reduction of the loss prevention efforts. This clearly renders the decision maker better off, which is illustrated in fig. 2 where points I and T are on higher indifference curves than Q or V, respectively.

The insurance induced reduction of the customer's loss prevention activities is sometimes associated with the term 'moral hazard'. However this term is misleading in the case at hand. Not because the behaviour is only rational from the standpoint of the insured,[13] but it is even desirable from the welfare aspect: there is indeed no allocation which could further improve the decision maker's situation, provided that the company is merely compensated for its costs and that the type of insurance contract is given.[14] The result is very plausible, for it doesn't make any sense at all to bear excessive costs of loss prevention activities if there is ultimately a cheaper way of cancelling out risks by consolidation in insurance companies.[15]

We can easily find many examples for the prosperous allocation effects of insurance. One very striking example could already be seen in the development of insurance in ancient Venice: A Venetian merchant who sent a ship to foreign harbours was engaged in a risky business, since it often occurred that he lost the ship and its goods. Thus for a long time the profession did not attract many, and the journeys went no further than neighbouring coasts. But at some point it proved to be advantageous to shift the burden of risk to the shoulders of speculators, to whom it was possible to consolidate risks because of their multiple engagements. This increased the Venetian society's risk bearing capacity to such an extent that business men ventured even to the most distant coasts and the Venetian merchant fleet soon dominated the Mediterranean.

[13]Cf. the discussion between Pauly (1968) and Arrow (1970, pp. 220–222) about moral hazard and individual rationality.

[14]If we allow for other types of contracts so that e.g. the insured can choose an optimal deductable, then the allocation could possibly be further improved. Cf. Arrow (1970, pp. 212–219, and 1974) who unfortunately only refers to the case of given risks.

[15]Although this result does not reflect the popular judgement on the allocation effect of insurance, it was already stated by Mahr (1951, pp. 88 f., 91 f.) and Arrow (1970, pp. 137 f.). For the special case of zero managing costs similar results were obtained in the time state preference approaches of Ehrlich–Becker (1972, pp. 636 f.) and Helpman–Laffont (1975, p. 21). Cf. also Marshall (1976) who reaches a different conclusion for a two-person game situation where the insurer is risk averse. A view popular among active insurance brokers, according to which insurance companies increase welfare by helping the insured persons to prevent losses, is expressed by Slanec (1972).

2.2.4. *The loss prevention policy under community rating*

If assumption (16) were a correct description of reality for all kinds of insurance, then we should expect enormous utility gains from an insurance of practically all economic risks. In addition to this we should require the total welfare state which guarantees an absolutely safe income for everyone.

Unfortunately, however, for many branches of insurance it is typical that the insurance company has much less information about the loss distribution than the insured person himself, since for the latter there are always some possibilities to manipulate risks without the company's knowledge. In order to express this pointedly, we assume for the time being that the insurance company is completely unaware which of the distributions of the original opportunity set the decision maker chooses.[16] In this case the company has no other choice than to practice community rating, i.e., to calculate μ^* according to the average loss of all insureds as observed in the past. For the calculation of the individual this means that

$$\mu^* = \text{const.}, \tag{18}$$

such that

$$\partial\mu^*/\partial\mu(C) = 0.$$

Under this assumption for the case of full coverage insurance we get from (13)

$$\bar{\mu}'(\sigma_M) = -k. \tag{19}$$

This formula holds, regardless of whether full coverage insurance was chosen since the company had offered an 'all or nothing' contract or whether it was chosen because $\pi\mu^*/\mu(C) \leqq 1$, although the whole range $0 \leqq \theta \leqq 1$ was feasible. As it can be verified by aid of figs. 1 and 2 formula (19) implies a drastic change in the loss prevention policy in comparison to the situation without insurance, as well as to the situation under equivalence rating: The decision maker chooses point P, which is characterized by a complete lack of loss prevention activities.

If instead partial coverage is chosen, then a substitution of condition (14) into (13), which both must be simultaneously satisfied, gives[17]

[16]Other theoretical approaches to this important question are: Spence–Zeckhauser (1971), Seidl (1972), Ehrlich–Becker (1972), Pauly (1974), Helpman–Laffont (1975), Marshall (1976) and Eisen (1976). A practically oriented analysis for fire insurance can be found in Menges (1970).

[17]Since $\mu(C) = k\sigma_M$ this formula implies a function $\sigma_M(\theta)$. Similarly (14) implies a function $\theta(\sigma_M)$ since $d\mu/d\sigma|_\theta$ depends on θ for a given σ_M. Thus, given the preference structure and given the function $\bar{\mu}_M(\sigma_M)$, θ and σ_M can be calculated explicitly.

$$\bar{\mu}'_M(\sigma_M) = -k + (1-\theta)k\pi \frac{\mu^*}{\mu(C)}. \tag{20}$$

According to this formula the decision maker chooses a point to the left of P on the efficiency frontier of the original opportunity set, if $0<\theta<1$. Contrary to the full coverage case, he will thus not completely cancel his loss prevention efforts. The formula shows moreover that the lower the degree of coverage, the higher the remaining efforts. This raises the question whether perhaps for sufficiently low degrees of coverage, the allocation could even be the same as under equivalence rating. Unfortunately, however, we cannot hope that this will be the case.

Let us assume that there is a great community of insured persons who all have the same time invariant preferences and the same, but stochastically independent, time invariant opportunity sets. Let us assume further that each insured knows his objective loss distribution. Since under these assumptions all insurance purchasers choose the same loss prevention policy, the average loss the company observes ex post is (except for small stochastic deviations) a correct estimator of each expected individual loss in the previous period. If the premium calculation of the following period is based on this estimator, then it is obviously impossible that an equilibrium of the insurance market, defined such that μ^* and $\mu(C)$ are constant in time, could exist without $\mu^* = \mu(C)$.[18] For if $\mu^* \neq \mu(C)$ then the true expected loss of a period deviates from its value in the previous period such that there is no equilibrium.[19] In respect of this result (20) simplifies to

$$\bar{\mu}'_M(\sigma_M) = k(\pi-1) - \theta k\pi. \tag{21}$$

In comparison to (17) this condition shows that we cannot hope that, for a sufficiently small degree of coverage, an insurance market equilibrium is possible where the allocation is the same as under a direct loss prevention control. Only at the limit $\theta \to 0$ such an equilibrium could be obtained, but then the insurance market would not exist any longer.

[18] Because of our simplifying assumptions, in equilibrium, community rating looks quite strange. Firstly, it resembles equivalence rating, since for each insured the premium is π times his individual expected loss, and secondly, all contracts are homogenous, since all insured persons choose the same loss prevention policy (although they do not have to do so). However, for just this reason now that property which is responsible for the striking difference to the allocation under equivalence rating becomes clear: The premium constancy for an individual change in expected loss. In the following it will be shown that this is an important source of welfare losses. This result should not be interpreted as if we deny the possibility of welfare losses under community rating, because the company requires an indiscriminate premium for heterogenous risks, as was shown by Pauly (1970). (The essence of Pauly's result is that even for given risks, under community rating, the high risk individual buys too much and the low risk individual too little insurance.)

[19] We only discuss equilibrium conditions, but exclude the stability question. For this problem, Mahr (1977) and Forster (1977).

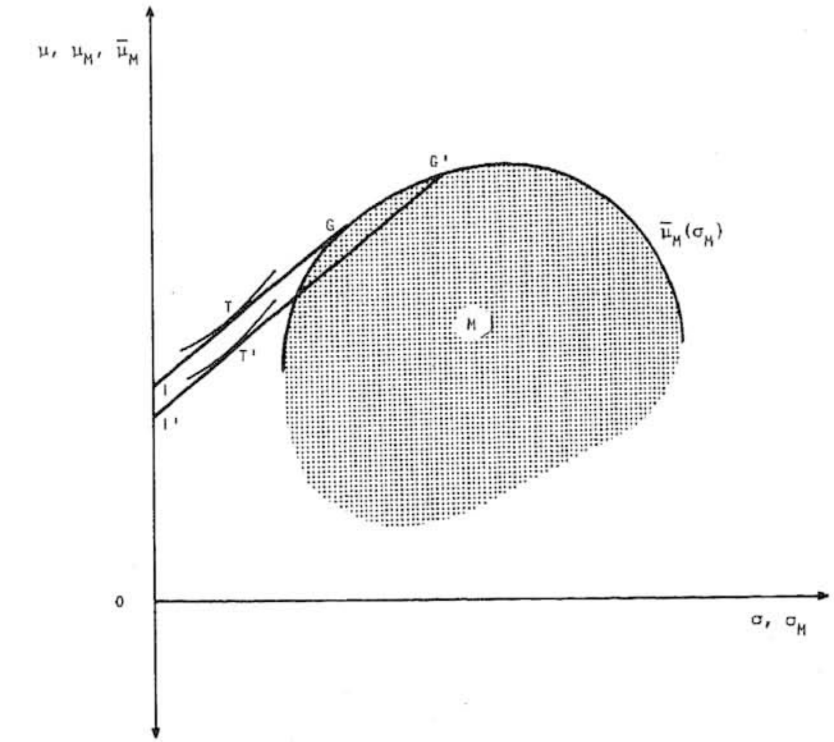

Fig. 3. The welfare loss of community rating.

As we have already found that the allocation described by (17) is Pareto-optimal the different allocation as described by (19) or (21), respectively, must imply a comparative welfare loss. For the case of partial coverage this loss is illustrated in fig. 3. There are two insurance lines. The upper one ($\overline{GI}$), which is tangent to the original opportunity set, is relevant if the insurance company can directly observe the insured's loss prevention activity; the lower one ($\overline{G'I'}$) is relevant if such a control is missing. As it can be seen by aid of (6) and (10), both lines have the same slope since for the reasons indicated above the insurer estimates the expected loss correctly, despite his lack of a direct control. Since in any case the insurance company is merely compensated for all its costs, the insured's decrease in utility, which is brought about by changing from T to T', indicates a clear deterioration in the sense of the Pareto criterion.

Of course this deterioration is measured against point G and not against the situation without insurance (point S of fig. 2). Whether there is even a deterioration against this initial situation is another question: The advantage of insurance as such (as discussed in section 2.2.2) might exceed the disadvantage arising from the fact that the insured reduces his loss prevention efforts too much. Fortunately in reality it is very easy to decide which is the case, for whenever insurance is bought on a liberal market it must obviously be worth it for the participants.

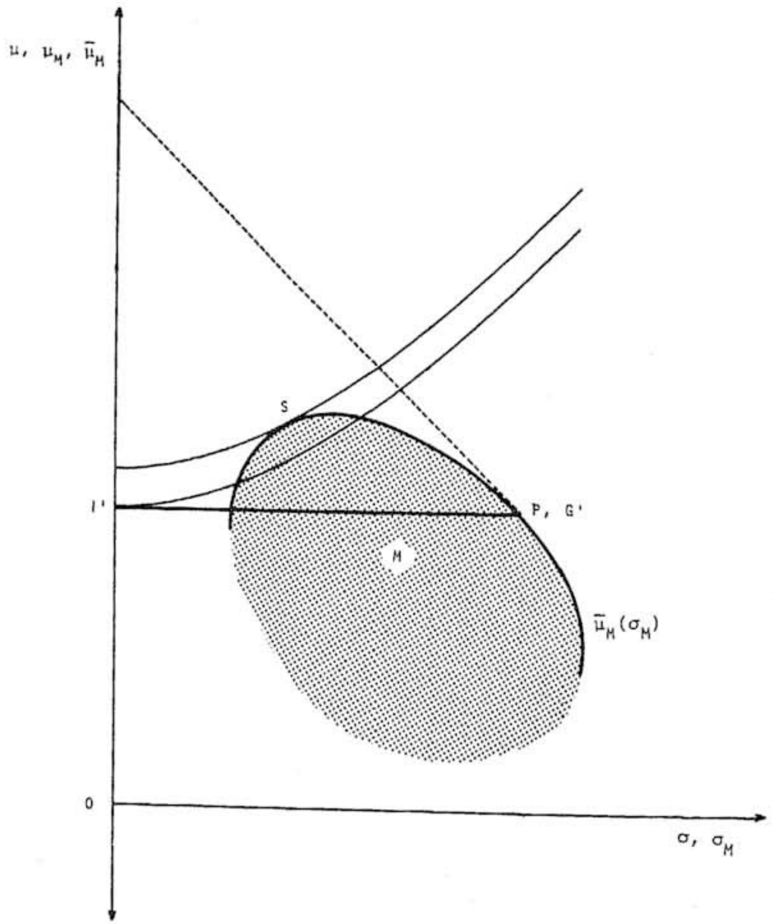

Fig. 4. The dilemma of the welfare state.

A problem can only arise if there is compulsory insurance. A not necessarily unrealistic situation which could be denoted the *dilemma of the welfare state* is illustrated in fig. 4. There it was assumed that the members of a community of identical citizens are obliged to buy full coverage insurance

and that the government requires a premium which is just enough to cover the observed average loss. Under these assumptions all citizens choose the original distribution P which is characterized by a complete lack of loss prevention efforts. So the individual has done the best from his personal point of view. But since all losses eventually have to be borne by the community, premiums (or taxes) must be so high that each citizen's income is $\overline{OI'}$, which he clearly dislikes more than the probability distribution S, which he would have chosen without compulsory insurance.[20] Of course it would be best for the community if each of its members were to choose a loss prevention policy which makes his original expected income maximal, but unfortunately people seem to lack collective rationality which would be necessary to behave in this way.[21]

2.3. *Practical implications for insurance rating*

At the private insurance market neither equivalence rating nor community rating is the rule: On the one hand the insurance company usually refers to a considerable number of objective, observable classification criteria for a definition of tariff categories; on the other hand, for the insured there remains in most cases a certain amount of scope for an unobservable manipulation of risk within such a category. By exploiting this scope, the insured causes misallocation, but in choosing the tariff category itself, he comes to the allocatively correct decision.

Thus it is obvious that, from an allocative point of view, a great variety of classification criteria is desirable, such that there is only a small scope of unobserved, but for the loss distribution relevant behaviour patterns of the insured. How small this scope should exactly be is, of course, an economic problem, since the advantage of a better control of the insurance purchaser, explained above, will be counterbalanced by its costs at any point. Unfortunately, we cannot say in general whether insurance companies have reached this point already or not. However, community rating, which is not practiced because of increasing classification costs, but, for example, in order to satisfy a 'solidarity principle', can definitely be rejected. It brings about clear welfare losses. The goal of a systematic redistribution of incomes between low and high risk individuals, which public compulsory insurance often tries to reach by following the solidarity principle, should rather be approached by a tax–subsidy mechanism parallel to insurance.

[20]This causes considerable doubts concerning a proposition of Arrow (1970, pp. 184 and 202) according to which those risks which are not yet covered by private insurance companies should be insured by the government in order to increase welfare.

[21]It must be admitted that collective rationality could in principle be achieved by the formulation of a strict moral code. But unfortunately communities which tried this in the past have not been very successful.

3. Insurance fraud

Among the insurance induced behaviour changes insurance fraud is certainly a very apparent one. Thus we shall only briefly review its allocative implications and means by which it can be avoided. There seem to be three types of fraudulent delicts.

The first type is *simulation of loss* in order to enjoy the insurance compensation. This type is allocatively neutral, since a destruction or suboptimal utilization of resources does not take place. The only thing that happens is an income redistribution from the mass of premium payers to the swindler, the judgment of which is a matter of our sense of justice. Unfortunately there does not seem to be a simple means to avoid the fraud.

The second type is induced by *overinsurance* ($\theta > 1$), where the insured can profit from a wilful destruction of the insured object or by increasing a damage that has already occurred, since the payment from the company overcompensates the loss. It seems to have some practical importance. For example we hear already from Haynes (1895, p. 445) that at the end of the last century in USA 35–50% of fire insurance damages were due to arson. But also today it seems that the increase of damages caused by fire during recessions must be explained in this way. Of course this effect indicates a clear misallocation: If all insured persons behaved in the described manner each one would obtain a compensation which could not exceed his premium, but the insured object would be destroyed. Fortunately, however, there is a very simple means to avoid this misallocation: The company only has to take care that no loss is covered by more than 100%.[22]

The third type of fraud is induced by *insurance of other's risk for one's own gain*. Here the insured person can make a profit of the difference between the insured value and the premium, if he causes the damage to the third party. Fisher (1906, pp. 294–295) for example reports the so-called 'graveyard insurance' which was possible in the States. Basically it was to take out an insurance contract on the lives of other people. It is not difficult to imagine in which horrible way insurance fraud was enacted with this kind of insurance. Also the judgment of its allocative value is of course obvious. Fortunately this type of insurance has no practical importance today, since it is forbidden by law in most countries.

4. The excess burden of the cost recompensation principle

In this section we study an insurance induced behaviour change which is, contrary to insurance fraud, of great practical importance. Its essence is that there is an exaggerated repair demand (in a broader sense) if a damage has occurred. The examples of the insured car owner who has his whole car

[22]Cf. Haynes (1895, pp. 445–446), Fisher (1906, pp. 294f.) and Arrow (1970, pp. 142 and 148).

painted at the cost of the company, although there is only a small scratch, and the housewife never missing her monthly cozy chat at the doctor's, thereby accumulating a mountain of medicine, are certainly familiar. The reason for the excessive demand is that the company does not pay an unconditional money compensation as we had assumed in the previous sections, but makes the compensation dependent on the repair costs or even pays the entire costs.[23]

The decision problem of the insured suffering from a damage can be illustrated by aid of fig. 5, which shows an indifference curve system for the goods x and y, where x measures the number of repair units received and y the money income of the person in question, which we regard as a quantity index of all the other goods he consumes. The indifference curve system is conditional on a particular insurance damage having occurred already. The more repair units the decision maker has already bought, the less is he willing to pay for an additional unit; thus the indifference curves are convex. Before buying any repair and before receiving a compensation from the company, the decision maker is at point A.[24] If he buys repair without cost recompensation, then he can move along the budget line $\overline{BA}$, the position of which is determined by the constant competitive price of a repair unit,[25]

$$P_r = \overline{OA}/\overline{OB}, \tag{22}$$

and income $\overline{OA}$. Things are different if the company pays back the share θ of the documented costs. In this case the cost recompensation reduces the net price for the insured (P_{rn}) to

$$P_{rn} = (1-\theta)P_r = \overline{OA}/\overline{OC} \tag{23}$$

with

$$\theta = \overline{BC}/\overline{OC},$$

[23]This was first pointed out by Pauly (1968). Cf. also Zeckhauser (1970), Grubel (1971), Feldstein (1973) and Rosett–Huang (1973). As far as these authors are concerned with the welfare loss of the cost recompensation principle, their approaches are objectionable since they calculate the loss in terms of consumer rents. The indifference curve analysis utilized here avoids the strong assumptions necessary for such a calculation.

[24]$\overline{OA}$ is a particular variate of $Y = b - C$ as defined in eq. (2), where C is that amount of money which compensates the insured if provided unconditionally.

[25]This assumption requires that, in the relevant range, the transformation curve between the repair commodity and the bundle of other goods is linear. (This is e.g. the case if there are constant-returns-to-scale technologies and the factor price ratio is constant despite the change in demand. The reason for the constancy of the factor price ratio may be that the sector in question is comparatively small or that it employs the production factors, given their prices, at the same ratio as the other sectors do.)

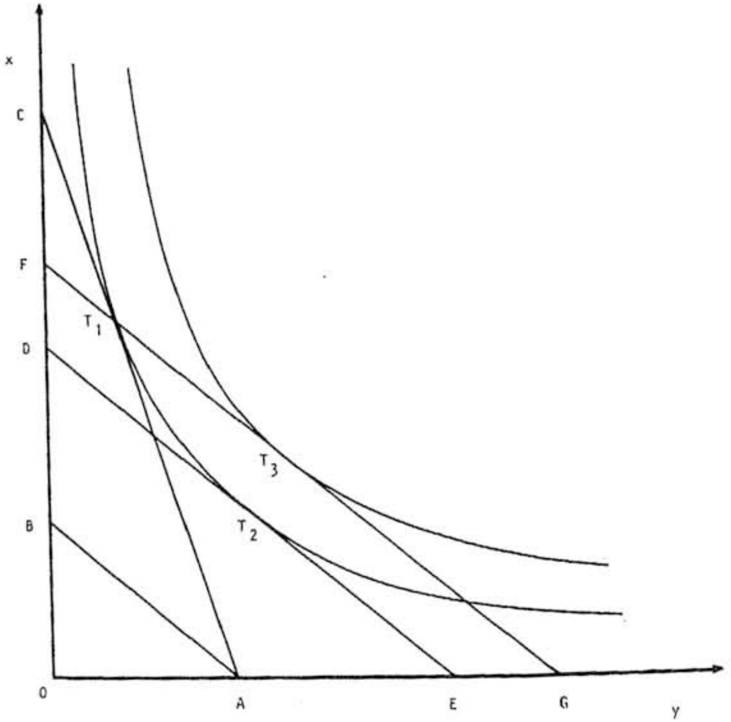

Fig. 5. The welfare loss of the cost recompensation principle.

such that the budget line turns to the new position $\overline{AC}$. Now, on this line the insured chooses the commodity bundle T_1, which implies that the company pays an amount of money up to $\overline{AG}$.[26]

Unfortunately the choice of the bundle T_1 which is optimal from the viewpoint of the insured indicates a clear misallocation: If the company had paid the amount $\overline{AG}$ unconditionally, then the insured would have chosen the bundle T_3 which represents a higher utility than T_1. Conversely, if the company had not linked the recompensation to the repair costs, then it could have saved the amount $\overline{EG}$ without making the insured worse off. The reason is that the insured would then have chosen the bundle T_2 which he likes as much as the bundle T_1.

The useless excess costs $\overline{EG}$ are undoubtedly a burden on the insurance market. In section 2, we found that a risk averse insurance purchaser should be willing to pay a premium beyond the expected monetary compensation

[26] Note that $\overline{OG}$, where G is the intersection with the abscissa of a budget line through T_1, parallel to $\overline{BA}$, measures the monetary value of the bundle represented by T_1, but that after the damage the insured's income was only $\overline{OA}$.

payments from the company, provided that these payments are unconditional. In order to make this theory applicable if the insured is indemnified according to the cost recompensation principle, we must transform the probability distribution of the company's payments into an equivalent distribution of unconditional payments. Then of course the insured would be willing to pay a premium in excess of the mean of the latter distribution. However it is by no means assured that he would also be willing to pay more than the mean of the former, as required by the company.

The allocative evaluation of the cost recompensation principle is obvious, since the excess costs $\overline{EG}$ induced by this principle are nothing but a useless wasting of resources. Whether insurance, despite this excess burden, causes a net welfare gain in comparison to the situation without insurance is again a question which can only be answered by testing as to whether a liberal market would exist or not. (This is similar to the case of community rating discussed above.) Anyhow, it is very dangerous if public compulsory insurances work according to the cost recompensation principle, since then there is no such test.

Where it is practiced, the cost recompensation principle should be abandoned, except that the repair is a merit good or induces positive external effects, as might be the case in health insurance. One must however take care that, in this case, the decision maker is not better off after the damage than before, since this would induce insurance fraud as discussed above. So, for example, for full coverage insurance a change in the recompensation principle has to be accompanied by a reduction of the loss payments by the company, which, of course, makes it possible to reduce premiums, too.

5. Conclusion

The insurance of previously uncovered risks leads clearly to a reduction of the efforts on the part of the insured to prevent damages. How far this reduction goes and how it is to be judged from the allocative standpoint depends upon the characteristics of the insurance contracts the company offers.

If the insurer indemnifies the insured by unconditional monetary payments, if he can observe the damage prevention policy, and if he demands strict equivalence premiums, then a Pareto-optimal allocation (under the given conditions) will be guaranteed through the decision of the insured: A welfare gain will arise with the substitution of costly loss prevention measures through less expensive insurance protection. On the other hand, if the insurance company practices community rating because, e.g., it follows the solidarity principle, then there will be a greater than optimal incentive to reduce loss prevention. The greater the degree of coverage, the stronger this will be. In the limiting case of full coverage the stimulus will be so strong that no attempts to reduce losses will be undertaken at all.

Besides two fraudulent behaviour changes induced by insurance, which cause misallocation in the cases of overinsuring and the insurance of other's risks, there is another type which deserves particular attention. It has its cause in the cost recompensation principle and leads to an exaggerated repair demand in the damage case. Clear welfare losses can also be ascribed to this type of behaviour change. They can be avoided by conversion to a lump sum compensation.

The welfare losses caused by the cost recompensation principle and by community rating can reduce substantially the utility gain for risk averse persons brought about through insurance as such. Of course it is very difficult to estimate welfare losses quantitatively. However, one thing can still be said: Whenever there is a free insurance market, there must be a net welfare gain in comparison to the situation without insurance. In the case of insurance which comes into existence by law one must be careful, though: Here the possibility of a net welfare loss cannot be excluded. From our theoretical explanation of the functioning of the insurance market it is possible to derive some concrete criteria for its allocatively correct construction. The three most important, expressed somewhat too pointedly, are:

- equivalence instead of community rating,
- unconditional damage compensation instead of cost recompensation,
- liberal insurance markets (for non-liability risks[27]).

[27]Cf. fn. 5.

References

Arrow, K.J., 1970, Essays in the theory of risk-bearing (Amsterdam and London).

Arrow, K.J., 1974, Optimal insurance and generalized deductibles, Scandinavian Actuarial Journal, 1–42.

Barrois, M.T., 1834, Essay sur l'application du calcul des probabilitié, Aux assurance contre l'incendie, Mémoires de la Société Royale des Sciences, de l'Agriculture et des Arts, Lille, 85–282.

Ehrlich, I. and G.S. Becker, 1972, Market insurance, self-insurance, and self-protection, Journal of Political Economy 80, 623–648.

Eisen, R., 1976, Unsicherheit und Information, Unkontrollierbares Verhalten und das Problem des moralischen Risikos, Jahrbücher für Nationalökonomie und Statistik 191, 193–211.

Feldstein, M.S., 1973, The welfare loss of excess wealth insurance, Journal of Political Economy 81, 251–280.

Fisher, I., 1906, The nature of capital and income (London).

Forster, E., 1977, Äquivalenz und Disvalenz in Versicherung und Staat, Zeitschrift für die Gesamte Versicherungswissenschaft 66, 241–253.

Grubel, H.G., 1971, Risk, uncertainty, and moral hazard, The Journal of Risk and Insurance 38, 99–106.

Gürtler, M., 1936, Die Kalkulation der Versicherungsbetriebe (Berlin).

Haynes, J., 1895, Risk as an economic factor, The Quarterly Journal of Economics 9, 409–449.

Helpman, E. and J.-J. Laffont, 1975, On moral hazard in general equilibrium theory, Journal of Economic Theory 10, 8–23.

Mahr, W., 1951, Einführung in die Versicherungswirtschaft, Allgemeine Versicherungslehre (Berlin).

Mahr, W., 1977, Das moralische Risiko in Individual- und Sozialversicherung, Zeitschrift für die Gesamte Versicherungswissenschaft 66, 205–240.

Marshall, J.M., 1976, Moral hazard, The American Economic Review 66, 880–890.

Menges, P., 1970, Die Prämienpolitik der Feuerversicherer, Dissertation (University of Hamburg).

Mossin, J., 1968, Aspects of rational insurance purchasing, Journal of Political Economy 76, 553–568.

Pauly, M.V., 1968, The economics of moral hazard: Comment, The American Economic Review 58, 531–537.

Pauly, M.V., 1970, The welfare economics of community rating, The Journal of Risk and Insurance 37, 407–418.

Pauly, M.V., 1974, Overinsurance and public provision of insurance: The roles of moral hazard and adverse selection, The Quarterly Journal of Economics 88, 44–62.

Rosett, R.N. and L. Huang, 1973, The effect of health insurance on the demand for medical care, Journal of Political Economy 81, 281–305.

Rothschild, M. and J. Stiglitz, 1976, Equilibrium in competitive insurance markets: An essay on the economics of imperfect information, The Quarterly Journal of Economics 90, 629–649.

Schneeweiß, H., 1967, Entscheidungskriterien bei Risiko (Berlin, Heidelberg, New York).

Seidl, Ch., 1972, Moral Hazard: Die individuelle Maximierung des Erwartungsnutzens als Quelle von Wohlfahrtsminderungen, Zeitschrift für Nationalökonomie 32, 425–448.

Sinn, H.-W., 1977, Die Allokationswirkungen der Versicherung, Zeitschrift für die gesamte Versicherungswissenschaft 66, 507–538.

Slanec, E., 1972, Versicherung – Produkt und Faktor in der Wirtschaft, Österreichische Hochschulzeitung 24, 15–16.

Spence, M. and R. Zeckhauser, 1971, Insurance, information, and individual action, The American Economic Review 61, 380–391.

Tobin, J., 1958, Liquidity preference as behavior towards risk, Review of Economic Studies 25, 65–86.

Zeckhauser, R., 1970, Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives, The Journal of Economic Theory 2, 10–26.