Optimal Risk-Sharing and Deductables in Insurance.

Knut K. Aase
Norwegian School of Economics and Business Administration
5045 Sandviken - Bergen, Norway

Abstract

Risk-sharing in insurance is analyzed, with a view towards explaining the prevalence of deductibles.

First we introduce, in a modern setting, the main concepts of the theory of risk-sharing in a group of agents. This theory we apply to the risk-sharing problem between an insurer and an insurance customer. We motivate the development through simple examples, illustrating some of the subtle points of this theory.

In order to deduce deductibles endogenously, not explained in the neoclassical model, we separately introduce (i) the insurable asset as a decision variable, (ii) administrative costs, and (iii) moral hazard, and illustrate by examples.

KEYWORDS: Reinsurance Exchange, Equilibrium, Pareto Optimality, Representative Agent, Core Solution, Individual Rationality, Deductibles, Costs, Moral Hazard.

Introduction

Deductibles in some form or another is often part of real world insurance contracts, but most of the classical theory of risk-sharing is unable to explain this phenomenon. In this paper we discuss deductibles from the perspective of economic risk theory, and try to point out when optimal contracts contain deductibles.

The model we consider is a special case of the following: Consider a group of \( I \) agents, that could be consumers, investors, reinsurers (syndicated members), insurance buyers etc., having preferences over a suitable set of random
prospects. These preferences are assumed to be represented by expected utility, meaning that there is a set of utility functions \( u_i : R \rightarrow R \), such that \( X \) is preferred to \( Y \) if and only if \( E u_i(X) \geq E u_i(Y) \). Here the random variables \( X \) and \( Y \) may represent terminal wealth, or consumption in the final period. They could also represent insurance risks by a proper reformulation. We assume smooth utility functions with the properties that more is preferred to less, i.e., these functions are strictly increasing, and agents are risk averse, i.e., these functions are all convex.

Each agent is endowed with a random payoff \( X_i \) called his initial portfolio. Uncertainty is objective and external, and there is no informational asymmetry. All parties agree upon the space \((\Omega, \mathcal{F}, P)\) as the probabilistic description of the stochastic environment, the latter being unaffected by their actions. Here \( \Omega \) is the set of states of the world, \( \mathcal{F} \) is the set of events, formally a sigma-field generated by the given initial portfolios \( X_i \), and \( P \) is the common belief probability measure. It will be convenient to posit that both expected values and variances exist for all the initial portfolios, since this it true in real life applications of the theory.

We suppose the agents can negotiate any affordable contracts among themselves, resulting in a new set of random variables \( Y_i \), representing the possible final payout to the different members of the group, or final portfolios. The transactions are carried out right away at “market prices”, where \( \pi(Y) \) represents the market price for any non-negative random variable \( Y \), and signifies the group’s common valuation of \( Y \) relative to the other random variables having finite variance. The essential objective is then to determine:

(a) The market price \( \pi(Y) \) of any “risk” \( Y \) from the set of preferences of the agents and the joint probability distribution of the \( X_i \)’s.

(b) For each individual, to determine the final portfolio \( Y_i \) most preferred by him among those satisfying his budget constraint.

Finally, market clearing requires that the sum of the \( Y_i \)’s equals the sum of the \( X_i \)’s, denoted \( X_M \).

The outline of the paper is as follows: In Section 1 we define key concepts like Pareto optimality, risk tolerance (aversion), and give Borch’s characterization of Pareto optimality. In our subsequent treatment we more or less take for granted the concepts of a competitive equilibrium, Borch’s Theorem, as well as the representative agent construction that leads to the solution of (a) and (b) above. With this as a starting point, in Section 2 the above general formulation is applied to the problem of finding the optimal risk exchange between an insurer and a policy holder. In this section we also present a simple, motivating example. Deductibles in the optimal contracts are treated in Section 3. First we discuss deductibles stemming from no frictions, then from administrative costs and finally deductibles arizing from moral hazard.
Section 4 concludes.

A glaring omission in our treatment of deductibles is that of adverse selection. Rothschild and Stiglitz (1976) showed in their seminal paper that the presence of high risk customers caused the low risk customers to accept a deductible. Since the analysis of this problem deviates quite substantially from the treatment in the rest of the paper, we have omitted it. Adverse selection is, however, another prominent example of asymmetric information that causes deductibles and less than full insurance to be optimal.

Risk tolerance and aggregation

First we define what is meant by a Pareto optimum. The concept of Pareto optimality offers a minimal and uncontroversial test that any social optimal economic outcome should pass. In words, an economic outcome is Pareto optimal if it is impossible to make some individuals better off without making some other individuals worse off.

Karl Borch’s characterization of a Pareto optimum \( Y = (Y_1, Y_2, \cdots, Y_I) \) simply says that there exist positive "agent" weights \( \lambda_i \) such that the marginal utilities at \( Y \) of all the agents are equal modulo these constants, i.e.,

\[
\lambda_1 u'_1(Y_1) = \lambda_2 u'_1(Y_2) = \cdots = \lambda_I u'_1(Y_I) = u'_\lambda(X_M), \quad \text{a.s.}
\]

The function \( u'_\lambda(\cdot) \) is interpreted as the marginal utility function of the representative agent. The risk tolerance function of an agent \( \rho(x) : R \to R_+ \), is defined by the reciprocal of the absolute risk aversion function \( R(x) = \frac{-u''(x)}{u'(x)} \), where the function \( u \) is the utility function in the expected utility framework. Consider the following nonlinear differential equation:

\[
Y'_i(x) = \frac{R_\lambda(x)}{R_i(Y_i(X_M))}, \tag{1}
\]

where \( R_\lambda(x) = \frac{u''(x)}{u'(x)} \) is the absolute risk aversion function of the representative agent, and \( R_i(Y_i(X_M)) = \frac{-u''(Y_i(x))}{u'(Y_i(x))} \) is the absolute risk aversion of agent \( i \) at the Pareto optimal allocation function \( Y_i(x) \), \( i \in I \). There is a neat result connecting the risk tolerances of all the agents in the market to the risk tolerance of the representative agent in a Pareto optimum. It goes as follows:

(a) The risk tolerance of "the market", or the representative agent, \( \rho_\lambda(X_M) \) equals the sum of the risk tolerances of the individual agents in a Pareto optimum, or

\[
\rho_\lambda(X_M) = \sum_{i \in I} \rho_i(Y_i(X_M)) \quad \text{a.s.} \tag{2}
\]
(b) The real, Pareto optimal allocation functions $Y_i(x): R \rightarrow R$, $i \in I$ satisfy the nonlinear differential equations (1).

The result in (a) first appeared in Wilson (1968), and can also be found by Borch (1985). It is in fact a direct consequence of Borch’s Theorem; see also Bühlmann (1980) for the special case of exponential utility functions.

The result has several important consequences, among which we mention one: Suppose there is a risk neutral agent. In a Pareto optimum all the risk will be carried by this agent.

The risk exchange between an insurer and a policy holder

Some of the issues that may arise in finding optimal insurance contracts are, perhaps, best illustrated by an example:

Example 1: A potential insurance buyer can pay a premium $\alpha p$ and thus receive insurance compensation $I(x) := \alpha x$ if the loss is equal to $x$. He then obtains the expected utility

$$U(\alpha) = E\{u(w - X - \alpha p + I(X))\},$$

where $0 \leq \alpha \leq 1$ is a constant of proportionality. It is easy to show that if $p > EX$, it will not be optimal to buy full insurance, i.e., $\alpha^* < 1$ (see e.g., Mossin (1968)).

On the other hand, the following approach was suggested by Karl Borch (e.g., Borch (1990)). The situation is the same as in the above, but the premium functional is $p = (\alpha EX + c)$ instead, where $c$ is a non-negative constant. It follows that here $\alpha^* = 1$, i.e., full insurance is indeed optimal:

$$Eu(w - X + I(X) - EI(X) - c) \leq u(w - EX - c),$$

which follows by Jensen’s inequality, in fact for any compensation function $I(\cdot)$. Equality above is obtained when $I(x) = x$, i.e., full insurance is optimal (if $c$ is small enough such that insurance is preferred to no insurance.) In other words, we have just demonstrated that the optimal value of $\alpha$ equals one. □

The seeming inconsistency between the solutions to these two problems has caused some confusion in the insurance literature. In the two situations of Example 1 we only considered the insured’s problem. Let us instead take the full step and include both the insurer and the insurance customer, in which case “optimality” mean Pareto optimality.
Consider a policy holder having initial capital \( w_1 \), a positive real number, and facing a risk \( X \), a non-negative random variable. The insured has utility function \( u_1 \), where monotonicity and strict risk aversion prevails, \( u_1' > 0, u_1'' < 0 \). The insurer has utility function \( u_2, u_2' > 0, u_2'' \leq 0 \), so that risk neutrality is allowed, and initial reserves \( w_2 \), also a positive real number. These parties can negotiate an insurance contract, stating that the indemnity \( I(x) \) is to be paid by the insurer to the insured if claims amount to \( x \geq 0 \). It seems reasonable to require that \( 0 \leq I(x) \leq x \) for any \( x \geq 0 \). Notice that this implies that no payments should be made if there are no claims, i.e., \( I(0) = 0 \). The premium \( p \) for this contract is payable when the contract is initialized.

We recognize that we may employ our established theory for generating Pareto optimal contracts. Doing this, Moffet (1979) was the first to show the following:

The Pareto optimal, real indemnity function \( I: \mathbb{R}^+ \rightarrow \mathbb{R}^+ \), satisfies the following nonlinear, differential equation

\[
\frac{\partial I(x)}{\partial x} = \frac{R_1(w_1 - p - x + I(x))}{R_1(w_1 - p - x + I(x)) + R_2(w_2 + p - I(x))},
\]

where the functions \( R_1 = -\frac{u_1''}{u_1'} \), and \( R_2 = -\frac{u_2''}{u_2'} \) are the absolute risk aversion functions of the insured and the insurer, respectively.

In our setting this result is an immediate consequence of the result (b) of the previous section, provided the premium \( p \) is taken as a given constant.

From this result we see the following: If \( u_2'' < 0 \), it follows that \( 0 < I'(x) < 1 \) for all \( x \), and, together with the boundary condition \( I(0) = 0 \), by the mean value theorem that

\[
0 < I(x) < x, \quad \text{for all} \quad x > 0,
\]

stating that full insurance is not Pareto optimal when both parties are strictly risk averse. Notice that the natural restriction \( 0 \leq I(x) \leq x \) is not binding at the optimum for any \( x > 0 \), once the initial condition \( I(0) = 0 \) is employed.

Also observe that contracts with a deductible \( d \) can not be Pareto optimal when both parties are strictly risk averse, since such a contract means that \( I_d(x) = x - d \) for \( x \geq d \), and \( I_d(x) = 0 \) for \( x \leq d \) for \( d > 0 \) a positive real number. Thus either \( I'_d = 1 \) or \( I'_d = 0 \), contradicting \( 0 < I'(x) < 1 \) for all \( x \).

However, when \( u_2'' = 0 \) we notice that \( I(x) = x \) for all \( x \geq 0 \): When the insurer is risk neutral, full insurance is optimal and the risk neutral part, the insurer, takes all the risk. Clearly, when \( R_2 \) is uniformly much smaller than \( R_1 \), this will approximately be true even if \( R_2 > 0 \).

This gives a neat resolution of the "puzzle" in Example 1. We see that the premium \( p \) does not really enter the discussion in any crucial manner.
when it comes to the actual form of the risk sharing rule $I(x)$, although this function naturally depends on the parameter $p$.

**Deductibles**

**No frictions**

The usual explanation of deductibles in insurance goes via the introduction of some frictions of the neoclassical model. There are a few notable exceptions, of which we mention one: The common model of insurance takes the insurable asset as given, and then finds the optimal amount of insurance given some behavioral assumptions. In order to explain the wealth effect in insurance, Aase (2007) uses a model where the amount in the insurable asset is also a decision variable. In this situation solutions can arise where the insurance customer would like to short the insurance contract, i.e., $I^*(x) < 0$ for some losses $x > 0$, for the optimal solution $I^*$. Since this violates $I^* \in [0, x]$, a deductible naturally arises, since no insurance then becomes optimal at strictly positive losses.

In fact this seems to be the most common reason why deductibles arise also under frictions; when some constraint becomes binding at the optimum, usually caused by a friction, a deductible may typically arise.

The first case we discuss in some detail is that of administrative costs.

**Administrative costs**

The most common explanation of deductibles in insurance is, perhaps, provided by introducing costs in the model. Intuitively, when there are costs incurred from settling claim payments, costs that depend on the compensation and are to be shared between the two parties, the claim size ought to be beyond a certain minimum in order for it to be Pareto optimal to compensate such a claim.

First we mention that Arrow (1970) has a result where deductibles seem to be a consequence of risk-neutrality of the insurer.

Following Raviv (1979), let the costs $c(I(x))$ associated with the contract $I$ and claim size $x$ satisfy $c(0) = a \geq 0$, $c'(I) \geq 0$ and $c''(I) \geq 0$ for all $I \geq 0$; in other words, the costs are assumed increasing and convex in the indemnity payment $I$, and are ex post.

Raviv then shows that if the cost of insurance depends on the coverage, then a nontrivial deductible is obtained. Thus Arrow’s (1970, Theorem 1) deductible result was not a consequence of the risk-neutrality assumption.
(see also Arrow (1974)). Rather it was obtained because of the assumption that insurance cost is proportional to coverage. His theorem is then a direct consequence of above cited result. A consequence of this is the following:

If \( c(I) = kI \) for some positive constant \( k \), and the insurer is risk neutral, the Pareto optimal policy is given by

\[
I(x) = \begin{cases} 
0, & \text{if } x \leq d; \\
x - d, & \text{if } x > d.
\end{cases}
\]

where \( d > 0 \) if and only if \( k > 0 \).

Here we obtain full insurance above the deductible. If the insurer is strictly risk averse, a nontrivial deductible would still result if \( c'(I) > 0 \) for some \( I \), but now there would also be coinsurance (further risk sharing) for losses above the deductible.

Risk aversion, however, is not the only explanation for coinsurance. Even if the insurer is risk neutral, coinsurance might be observed, provided the cost function is a strictly convex function of the coverage \( I \). The intuitive reason for this result is that cost function nonlinearity substitutes for utility function nonlinearity.

A, perhaps, more natural cost function than the one considered above, would be a zero cost if no claim, and a jump at zero, to \( a > 0 \) say, if a small claim is incurred. Thus the overall cost function is \( a\chi_{[I > 0]} + c(I) \), where \( c(0) = 0 \), \( \chi_B \) being the indicator function of \( B \). Then, so long as \( I < a \) it is not optimal for the insured to get a compensation, since the cost, through the premium, outweighs the benefits, and a deductible will naturally occur, even if \( c' \equiv 0 \).

To conclude this section, a strictly positive deductible occurred if the insurance cost depended on the insurance payment, and also if a cost at zero is initiated by a strictly positive claim. Coinsurance above the deductible \( d \geq 0 \) results from either insurer risk aversion or cost function nonlinearity.

For further results on costs and optimal insurance, see e.g., Spaeter and Roger (1995).

**Moral Hazard**

A situation involving moral hazard is characterized by a decision taken by one of the parties involved (the insured), that only he can observe. The other party (the insurer) understands what decision the insured will take, but can not force the insured to take any particular decision by a contract design, since he can not monitor the insured.
The concept of moral hazard has its origin in marine insurance. The old standard marine insurance policy of Lloyd’s - known as S.G. (ship and goods) policy - covered “physical hazard”, more picturesquely described as “the perils of the sea”. The “moral hazard” was supposed to be excluded, but it seemed difficult to give a precise definition of this concept. Several early writers on marine insurance (e.g., Dover (1957), Dinsdale (1949), Winter (1952)) indicate that situations of moral hazard were met with underwriters imposing an extra premium.

The following idea was initiated by Holmström (1979). In order to explain this, let, as in the above, \( u_2(x) \) and \( w_2 \) denote the insurer’s utility function and initial wealth, \( u_1(x) \), \( w_1 \) are similarly the corresponding utility function and initial wealth of the insured. In the latter case the function \( v(a) \) denotes disutility of level of care (or effort for short) \( a \), effort designated to minimize or avoid the loss. Only the insured can observe \( a \). The loss facing the insured is denoted by \( X \), having probability density function \( f(x, a) \). Observe that we are here thinking of part of the level of care that is unobservable by the insurer. Many insurance contracts are contingent upon certain actions taken by the insured, actions that can be verified or observed by the insurer. This is not what we have in mind at the present, we only consider non-observable level of care.

Notice that here we deviate from the neoclassical assumption that uncertainty is exogenous, since the stochastic environment is now effected by the insured’s actions.

The problem may be formulated as follows:

\[
\max_{I(x), a, p} E u_2(w_2 - I(X) + p)
\]

where \( I(x) \in [0, x] \), \( p \geq 0 \), subject to

\[
E u_1(w_1 - X + I(X) - p) - v(a) \geq \bar{h},
\]

and

\[
a \in \arg\max_{a'} \{ E u_1(w_1 - X + I(X) - p) - v(a') \}.
\]

The first constraint is called the participation constraint (individual rationality), and the last one is called the incentive compatibility constraint. This latter is the moral hazard constraint, and says that the customer will use the effort level \( a \) that suits him best, and this is also feasible, since the insurer does not observe \( a \). Technically speaking, the natural requirement \( 0 \leq I(x) \leq x \) will be binding at the optimum, and this causes a strictly positive deductible to occur. We illustrate by an example:
Example 2: Consider the case with \( u_2(x) = x \), \( u_1(x) = \sqrt{x} \), \( v(a) = a^2 \) and the probability density of claims \( f(x,a) = ae^{-ax} \) is exponential with parameter \( a \) (effort). Notice that \( P(X > x) = e^{-ax} \) decreases as effort \( a \) increases: An increase in effort decreases the likelihood of a loss \( X \) larger than any given level \( x \).

We consider a numerical example where the initial certain wealth of the insurer \( w_1 = 100 \), and his alternative expected utility \( h = 19.60 \). This number equals his expected utility without any insurance. In this case the optimal effort level is \( a^* = 0.3719 \).

In the case of no moral hazard, we solve the problem without the incentive compatibility constraint, and obtain what is called the first best solution. As expected, since the insurer is risk neutral, full insurance is optimal: \( I(x) = x \), and the first best level of effort \( a^{FB} = .3701 \), smaller than without insurance, which is quite intuitive.

The expected utility of the representative agent we may denote the welfare function. Here it is:

\[
Eu_2(w_2 + p - I(X)) + \lambda(Eu_1(w_1 - p - X + I(X)) - v(a)) \\
= w_2 + 195.83,
\]

since \( \lambda = \lambda^{FB} = 9.8631 \). Here \( p^{FB} = 2.72 \). Moving to the situation with moral hazard, full insurance is no longer optimal. We get a contract with a deductible \( d \), and less than full insurance above the deductible:

\[
I(x) = \begin{cases} 
0, & \text{if } 0 \leq x \leq d; \\
x + p - w_1 + (\lambda + \frac{x}{a} - \mu x)^2, & \text{if } d < x \leq \frac{\lambda}{a} + \frac{1}{a}; \\
x + p - w_1, & \text{if } x > \frac{\lambda}{a} + \frac{1}{a}.
\end{cases}
\]

Here the deductible \( d = 11.24 \), the second best level of effort \( a = a^{SB} = .3681 \). Notice that this is lower that \( a^{FB} \). The Lagrangian multipliers of the two constraints are: \( \lambda^{SB} = 9.7214 \) and \( \mu = 0.0353 \). The second best premium in this case is \( p^{SB} = 0.0147 \).

Due to the presence of moral hazard there is now a welfare loss: The expected utility of the representative agent has decreased:

\[
Eu_2(w_2 + p - I(X)) + \lambda^{SB}(Eu_1(w_1 - p - X + I(X)) - v(a^{SB})) \\
= w_2 + 187.73,
\]

implying a welfare loss of \( 8.10 \) compared to the first best solution. Notice that the premium \( p^{SB} \) is here lower than for the full insurance case of the first
best solution, due to a smaller liability, by the endogenous contract design, for the insurer in the situation with moral hazard. □

In the above example we have used the first order approach, justified in Jewitt (1988).

Again we see that deductibles may result, and also coinsurance above the deductible, when the classical model predicts full insurance and no deductible.

A slightly different point of view is the following: If the insured can gain by breaking the insurance contract, moral hazard is present. In such cases the insurance company will often check that the terms of the insurance contract are being followed. Note, in the above model the insurance company could not observe the action $a$ of the insured. This was precisely the cause of the problem in Example 2. Here, checking is possible, but will cost money, so it follows that the mere existence of moral hazard will lead to costs. This is a different type of costs from those of Example 2, but has the same origin - moral hazard. The present situation clearly invites analysis as a two-person game played between the insurance company and its customer.

Borch (1980) takes up this challenge, and finds a Nash equilibrium in mixed strategies. In this model the insured pays the full cost of moral hazard through the premium.

It seems to have been Karl Borch’s position that moral hazard ought to be met by imposing an extra premium.

From the above example we note that increased premiums may not really be the point in dealing with this problem: Here we see that the second best premium $p^{SB}$ is actually smaller than the first best $p^{FB}$. The important issue is that the contract creates incentives for the insured to protect his belongings, since he carries some of the risk himself under the second best risk sharing rule. This is brought out very clearly in the above example if the first best solution is implemented when moral hazard is present. Then the insured will set his level of effort $a = 0$, i.e., to its smallest possible level, since he has no incentive to avoid the loss, resulting in a very large loss with high probability (a singular situation with a Dirac distribution at infinity). Since the second best solution is the best when moral hazard is present, naturally this leads to a low welfare, in particular for the insurer.

Conclusions

Optimal risk sharing was considered from the perspective of Pareto optimality. Relying on the concepts of a competitive equilibrium, Pareto optimality, Borch’s Theorem, and the representative agent construction, we discussed
some key results regarding risk tolerance and aggregation, which we applied to the problem of optimal risk-sharing between an insurer and a customer. The development was motivated by a simple example, illustrating some of the finer points of this theory. It turned out that the concept of a Pareto optimal risk exchange is exactly the right notion of optimality in the present situation.

In order to explain deductibles, which do not readily follow from the neoclassical theory, we separately introduced (i) the insurable asset as a decision variable, (ii) administrative costs, and (iii) moral hazard, the latter also illustrated by an example.

References


