Chapter 2 The Supply of Insurance

1 Introduction

In modelling the market supply of goods in general, we proceed by first developing a theory of the firm, and then analysing its supply behaviour. The key underlying relationship is the production function, showing how the the output quantities that can feasibly be produced vary with the input quantities used. The general properties of this function are important because they determine the nature of the firm's costs, in particular how they vary with output. The production function and its properties are treated in a very general way. As economists we are not interested in the details of the engineering or technological relationships involved in producing some specific good, but only in their broad characteristics - the behaviour of marginal productivity as input quantities vary, the nature of the returns to scale - that allow us to put restrictions on the form of the firm's cost function.

In most of the economics literature on insurance markets, a much simpler approach is taken. It is just assumed that the market is "competitive", the "production costs" of insurance are zero, and as a result there is a perfectly elastic supply of insurance cover at a fair premium. This approach can be justified when the purpose of the model is to analyse specific issues that would only be unnecessarily complicated by a more complete specification of the supply side of the market. Later in this book for example we shall see this in the analysis of the implications of information asymmetries for the existence and optimality of insurance market equilibrium. It will not suffice however when we are concerned with the general analysis of insurance markets. Therefore in this chapter we develop a theory of the insurance firm and analyse its supply behaviour.

Our first concern will be with the "technology" of insurance. This has two aspects. On the one hand, there are the activities involved in physically "producing" insurance: drawing up and selling new insurance contracts, administering the stock of existing contracts, processing claims, estimating loss probabilities, calculating premiums, and administering the overall business. The costs involved in these activities are often referred to as "transactions costs", but since they clearly extend beyond what in the economics literature are normally referred to as transactions costs, we will call them **insurance costs**. We would generally expect them to increase with the "output" of insurance, the amount of cover sold, though there may well be a fixed overhead component independent of this.

The second aspect of insurance technology is conceptual rather than physical, and concerns the **pooling** and **spreading** of risk. When an insurer enters into insurance contracts with a number of distinct individuals, the probability distribution of the aggregate losses they suffer will in general differ from the loss distribution facing any one individual. We are interested in the nature of this aggregate loss distribution, or, more precisely, the distribution of claims on the insurer to which it gives rise. In particular, we want to establish its properties as the number of contracts sold becomes large.This is the issue of risk pooling. In addition, the insurer will typically not be a single individual, but rather a group of individuals. Each member of this group may face **unlimited liability**, in the sense that he will be liable to meet insurance losses to the full extent of his wealth; or **limited liability**, where his possible loss is limited to the extent of his shareholding. In the former case we refer to an insurance **syndicate**, in the latter to an insurance **company** or **firm**. In each case the insurance losses are being spread over a number of individuals, and we are interested in the question of how this affects the premium that would be set, given that the individuals may be risk averse.

Finally, a very important aspect of an insurer's operations are its investment activities. These arise in two ways. As we shall see, the insurer will have to hold reserves against the possibility that the aggregate value of loss claims will exceed its premium income. These will be invested in assets that yield a return. Secondly, since under every insurance contract premium revenue is collected in advance of the payout of any corresponding claim, this provides a flow of investible funds. For both these reasons large insurers are also major financial institutions. It is therefore of interest to examine how these two sides of the business, insurance and investment, interact.

In the next two sections we examine in a general way the economics of risk pooling and risk spreading in insurance markets, assuming that insurance costs are zero. We then go on to consider the implications of introducing insurance costs for an insurer in a competitive market, using a discrete version of a model first proposed by Artur Raviv.

In the following section we consider the implications of limited liabilty of insurers, and discuss the issue of choice of insurance reserves, which is closely bound up with the question of the regulation of insurance markets. In all this, we consider only pure insurance, and ignore the issues raised by the insurer's investment activities. In the concluding section of the chapter, we devekop a model that incorporates both insurance and investment activities.

2 Risk Pooling

We assume that the insurer enters into insurance contracts with n individuals, and we make the further assumption that the distribution of claims costs under each contract is identical, and independent across contracts. This assumption of **identically and independently distributed** (*i.i.d.*) risks is not essential for determining the aggregate claims distribution, but is very helpful in greatly simplifying the technicalities involved, while losing little of interest to the economist. Thus each contract is assumed to have the same probability distribution distribution of cover, and therefore of loss claims, $\tilde{C}_i$, with mean μ and variance σ^2, both finite, and with zero covariance between any pair of values C_i, C_j, $i, j = 1, .., n, i \neq j$. It follows from the standard properties of the sum of *i.i.d.* random variables that $\tilde{C}_n = \sum_{i=1}^{n} \tilde{C}_i$, is also a random variable with mean $n\mu$.We find its variance as

$$E[(\sum_{i=1}^{n} \tilde{C}_i - n\mu)^2] = E[\{\sum_{i=1}^{n}(\tilde{C}_i - \mu)\}^2] \tag{1}$$

and, since the covariances between the $\tilde{C}_i$ are all zero

$$E[\{\sum_{i=1}^{n}(\tilde{C}_i - \mu)\}^2] = \sum_{i=1}^{n} E[(\tilde{C}_i - \mu)^2] = n\sigma^2 \tag{2}$$

Note that the variance of the total claims cost increases linearly with n,while its standard deviation, $\sqrt{n}\sigma$, is strictly concave in n.

One immediate implication of this is that if the insurer sets the premium on each contract equal to the expected value of cover or claims cost μ, and insurance costs are zero, it will just break even **in expected value**, since its total premium revenue $n\mu$ will equal the expected value of claims costs. This is the reason for calling μ the fair premium. However, it must be emphasised that any one realisation of $\tilde{C}_n$, that is, actual aggregate claims costs in any one period, may be larger or smaller than $n\mu$, no matter how large the number of contracts sold, since the variance $n\sigma^2$ is always positive and increases with n. If the insurer is to avoid **insolvency**, *i.e.* the situation in which claims costs exceed the funds available to meet them, it will have to carry what are called technical or insurance reserves. Now, it is reasonable to assume that each contract has a maximum cover $C_{\max}$, and so there is a maximum aggregate claims cost $nC_{\max}$. Thus, in principle, if the insurer sets a premium amount P per contract and also carries reserves (ignoring for the moment investment income and associated risk) $R_{\max} = n(C_{\max} - P)$, it will have a zero probability of insolvency. In practice, however, the probability that actual claims costs will be in the region of $nC_{\max}$ is typically extremely small, while, for a large insurer, attempting to raise a capital of $R_{\max}$ could be extremely costly. Consequently, insurers proceed by choosing a so-called

ruin probability, which we denote by ρ, and, given the distribution of aggregate claims costs, they then choose a level of reserves $R(\rho) = C_\rho - nP$, where C_ρ satisfies

$$\Pr[\tilde{C}_n > C_\rho] = \rho \tag{3}$$

That is, reserves are set at a level such that the probability is ρ that actual claims costs will exceed premium revenue plus reserves (again ignoring investment income) and the insurer will be insolvent.

Figure 1 illustrates, for the case in which the insurer sets the fair premium, $P = \mu$. The aggregate loss claims distribution is bounded below by zero and above by $nC_{\max}$, and C_ρ is the value of aggregate claims such that with probability ρ the insurer will be insolvent. For a given value of ρ, the value C_ρ will increase with the number of contracts n.Since μ is independent of n, this means that the value of the required reserves $R(\rho)$ must also increase with n.

Figure 1 about here

It is clearly of interest to ask how the ruin probability ρ is determined. It will result from a solution to the problem of the optimal trade-off between the costs of insolvency and the cost of holding reserves. We shall explore this problem in more detail in section X below. First, we consider the implications of the *Law of Large Numbers* for the value of the loss and insurance reserves *per contract*.

Consider a particular realisation $C_1, C_2, ..., C_n$ of the claims under the n individual contracts. We can regard this as a random sample from a distribution with mean μ and variance σ^2,both finite. Let $\bar{C}_n$ denote the sample mean, or average loss per contract, *i.e.* $\bar{C}_n = \frac{1}{n}\sum_{i=1}^{n} C_i$. Then the version of the Law of Large Numbers (there is more than one) relevant for present purposes says that for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr[\left|\bar{C}_n - \mu\right| < \varepsilon] = 1 \tag{4}$$

In words, as n becomes increasingly large, this sample mean, the average loss claim per contract, will be arbitrarily close to the value μ with probability approaching 1. Put loosely, this says that for a sufficiently large number of insurance contracts, it is virtually certain that the loss per contract is just about equal to μ, the mean of the individual loss distribution. As the number of contracts increases, so the probability that the loss per contract lies outside an arbitrarily small interval around μ goes to zero.

It is also useful to look at the variance of $\bar{C}_n$. This is given by

$$E[(\frac{1}{n}\sum_{i=1}^{n} C_i - \mu)^2] = E[\frac{1}{n^2}(\sum_{i=1}^{n} C_i - n\mu)^2] = \frac{1}{n^2}E[(\sum_{i=1}^{n} C_i - n\mu)^2] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \tag{5}$$

Thus the variance of the realised loss per contract about the mean of the individual loss distribution goes to zero as n goes to infinity.

Now, in the Law of Large Numbers statement in (.) set $\varepsilon = \sigma^2/n$, so that we have

$$\lim_{n\to\infty} \Pr[\left|\bar{C}_n - \mu\right| < \frac{\sigma^2}{n}] = 1 \tag{6}$$

This suggests that as the number of individual insurance contracts sold by an insurer becomes very large, the risk that the claims cost per contract will exceed the fair premium becomes vanishingly small. We can interpret this as a type of **economy of scale**: although the variance of aggregate claims increases with n, so the insurance reserves will have to increase in *absolute* amount, the required reserve *per contract* tends toward zero: required reserves increase less than proportionately with size of the insurer, measured in terms of the number of individual insurance contracts.

3 Risk Spreading

Suppose now that the insurer is either a syndicate with N members or a company with N shareholders. It will simplify the analysis without losing much of economic interest if we assume that these individuals are all identical and share the net income from the insurance business equally, so that each receives a share $s = 1/N$ of this net income. The main difference between these two types of insurer in the present context is that if it is a syndicate, the total wealth of the members will have to be at least equal to the insurance reserve implied by the chosen ruin probability, while for a company, the equity capital would have to be at least this amount. That said, we will ignore the distinction for the time being, by assuming it is costless to hold reserves. We also assume that the individuals are risk averse. The question of interest is: what, if any, are the implications of increasing the number of individuals N in the syndicate or company, *i.e* of spreading the risky income over a larger number of individuals? The intuition would be that this should in some sense reduce the riskiness of the individual incomes and therefore reduce the risk premium that they would demand as a condition of taking a share in the insurance, thus reducing the insurance premium. Again we would have a type of economy of scale. Support for this intuition can be gained by applying the discussion of the Arrow-Pratt measure of absolute risk aversion given in the Appendix to the previous chapter. There we saw that we can derive the approximation

$$r \approx -\frac{1}{2}\frac{u''(y_0)}{u'(y_0)}\sigma_z^2 \approx \frac{1}{2}A(y_0)\sigma_z^2 \tag{7}$$

where r is the individual's risk premium, or amount he would have to be paid to compensate him for accepting a small risky income z with zero mean and variance σ_z^2, y_0 is his income in the absence of this risk, and $A(y_0)$ is his Arrow-Pratt index of risk aversion. Now let this risky income z consist of the share $s = 1/N$ in a given aggregate risky income $\tilde{Z}$, which has mean zero and variance σ_Z^2. Then we have

$$\sigma_z^2 = E[(\frac{\tilde{Z}}{N})^2] = \frac{\sigma_Z^2}{N^2} \tag{8}$$

Then clearly $\lim_{N\to\infty} r = 0$. For a sufficiently large number of individuals, each becomes essentially risk neutral. More to the point, if we consider Nr, the *sum* of the risk premia

$$Nr \approx \frac{N}{2}A(y_0)\sigma_z^2 = \frac{1}{2}A(y_0)\frac{\sigma_Z^2}{N} \tag{9}$$

it is clear that the total cost of risk bearing to the syndicate as a whole, Nr, goes to zero as the syndicate size grows, so that for sufficiently large N the risk aversion of the individuals can be ignored, and the insurer can be treated as risk neutral. Intuitively, the individual risk premia fall at a rate determined by N^2, while the sum of risk premia grows at a rate determined by N, and so overall this sum goes to zero as N grows. We now consider a more rigorous and general formalisation of this intuition.

3.1 The Arrow-Lind Theorem

This theorem has many applications over and beyond insurance markets, but is also of central importance here. It confirms the intuitive idea that the larger the number of syndicate members who share in a given distribution of income from a risky insurance business, the smaller the cost of the risk associated with that business, even though the individual syndicate members are risk averse. More importantly, it makes clear a necessary condition for this result, namely that *the covariance between the member's income from the insurance business, and his marginal utility of income if he does not share in this business, be zero.* Thus let $\tilde{Z}$ be the aggregate income from the insurance business, and $E[\tilde{Z}]$ its expected value. There are N members of the syndicate or shareholders of the company, and each receives a share $s = 1/N$ of the random income $\tilde{Z}$. Assume each member has an identical risk averse utility function $u(.)$ and non-insurance income y,which may be a random variable. The key assumption is that $Cov[\tilde{Z}, u^{'}(y)] = 0$ (which of course certainly holds if y is certain).

Now, define the certain amount of income r to satisfy

$$E[u(y + s\tilde{Z} + r) = E[u(y)] \tag{10}$$

Note that this is an identity in r, and implicitly defines r as a function of s. We could think of r as the amount the individual would require to be paid to induce him to participate in the insurance business. If this is negative, it is the amount he would pay for a share in the business. It is obvious that as $N \to \infty$, *i.e.* as $s \to 0$, we have $r \to 0$. For example, a risk averse decision taker with a certain income would always be indifferent about accepting a fair bet if it is small enough - to the first order expected utility would be unchanged. What the theorem shows, however, is the somewhat less obvious fact that, on the given assumptions, the *sum* $Nr(s) = r(s)/s \to -E[\tilde{Z}]$ as $N \to \infty$. We can interpret this as saying that for sufficiently large N, the aggregate market value of the insurance business can be taken as the expected value of its net income - we can treat the insurer as risk neutral.We now show this.

First, since the RHS of (10) is independent of s, we can apply the Implicit Function Theorem to obtain

$$\frac{dr}{ds} = -\frac{E[u^{'}(y + s\tilde{Z} + r)\tilde{Z}]}{E[u^{'}(y + s\tilde{Z} + r)]} \tag{11}$$

Now consider

$$\lim_{s \to 0} \frac{r(s)}{s} \tag{12}$$

Since both numerator and denominator go to zero, we apply l'Hôpital's Rule

$$\begin{aligned}
\lim_{s \to 0} \frac{r(s)}{s} &= \lim_{s \to 0} \frac{dr(s)/ds}{ds/ds} && (13)\\
&= \lim_{s \to 0} -\frac{E[u^{'}(y + s\tilde{Z} + r)\tilde{Z}]}{E[u^{'}(y + s\tilde{Z} + r)]} && (14)\\
&= -\frac{E[u^{'}(y)\tilde{Z}]}{E[u^{'}(y)]} && (15)\\
&= \frac{-E[u^{'}(y)]E[\tilde{Z}] - Cov(u^{'}(y), \tilde{Z})}{E[u^{'}(y)]} && (16)
\end{aligned}$$

Given the assumption $Cov(u^{'}(y), \tilde{Z}) = 0$ we have

$$\lim_{s \to 0} \frac{r(s)}{s} = -E[\tilde{Z}] \tag{17}$$

Thus, the aggregate value of the insurance business to the participants is equal to its expected value, with no adjustment for risk, as long as the uncertain net income has a zero covariance with the individuals' marginal utility of income from outside the business. Note that if this covariance were positive, implying, since $u'' < 0$, a negative covariance between y and $\tilde{Z}$, the aggregate value of the insurance business to its shareholders would exceed its expected value, and conversely if the covariance were negative. In the former case, the insurance business offers the shareholders a way of diversifying their asset portfolio.

.

4 Insurance Costs and the Raviv model

The previous two sections have discussed conditions under which an insurer may act as if it were risk averse in setting a fair premium for cover: if the number of insurance contracts with *i.i.d* risks is sufficiently large, then the cost per contract of holding reserves is close enough to zero to be ignored; and if the number of syndicate members or company shareholders is sufficiently large, and the zero covariance condition is met, then their individual risk aversion can be ignored and the insurer treated as risk neutral. All this however ignored insurance costs. We now analyse the implications of introducing these costs for an insurer supplying insurance in a perfectly competitive market. To focus the analysis on the effect of insurance costs, we assume

- insurance buyers are identical and their relevant characteristics - utility function, income, loss distribution - are fully known to the insurer

- the insurer can be treated as risk neutral

- we can ignore the cost of reserves in pricing individual insurance contracts.

The loss distribution is $\{0, L_1, L_2, ..., L_S\}$, with corresponding probabilities $\{\pi_0, \pi_1, \pi_2, ..., \pi_S\}$, all positive, and with $0 < L_1 < L_2 < ... < L_{S.}$ The premium (amount) is P, and so the insurance buyer's incomes are

$$y_0 = y - P \quad (18)$$

$$y_s = y - P - L_s + C_s \quad s = 1, .., S \quad (19)$$

where $C_s \geq 0$ is cover in state s. The insurer's profits on any one contract in the respective states are

$$x_0 = P - F \quad (20)$$

$$x_s = P - C_s - K(C_s) \quad s = 1, .., S \quad (21)$$

where the **insurance cost function** $K(C_s)$ has $K^{'}(.) \geq 0$, $K^{''}(.) \geq 0$, and $K(0) = F \geq 0$, a fixed cost. Note that this cost function relates to one individual insurance contract. In effect the model is assuming that the cost of having n contracts is just n times the cost of one contract.

We use the assumption that the insurance market is perfectly competitive in formulating the problem of the insurer's choice of an optimal contract. Perfect competition implies two things:

- the insurer will make zero profits in expected value, $\sum_{s=0}^{S} \pi_s x_s = 0$, since otherwise entry or exit of risk neutral insurers will take place. In other words, the contract we derive is a long run equilibrium contract

- this equilibrium contract must also maximise the expected utility $\bar{u} = \sum_{s=0}^{S} \pi_s u(y_s)$ of the insurance buyer since, if not, a competitior could offer a superior contract and compete away the business

We state and discuss the main results before deriving them formally. From the analysis of insurance demand in the previous chapter we know that a buyer will choose full cover if offered a fair premium, while otherwise he prefers a contract with a deductible over all other types of contract with the same expected cost to the insurer. The existence of insurance costs suggests that a fair premium will not be feasible, and thus, given the competitive market assumption, contracts with a deductible are likely to make an appearance. This is what we find. With positive marginal costs $K^{'}(.) > 0$, we find that the optimal contract will give partial cover, and is likely to involve a deductible (if the loss distribution were continuous on $[0, L_S]$ it would be certain to). We also obtain two further interesting results. First, if marginal costs are increasing, $K^{''}(.) > 0$, then the optimal contract involves *coinsurance above a deductible,* or an increasing gap between loss and cover, whereas if marginal costs are constant, $K^{''}(.) = 0$, there is simply a deductible. Secondly, if marginal costs are zero, so that the insurance costs take the form of a fixed cost per contract, $K(.) = F > 0$, there is full cover, implying that the insurer offers a premium $P = F + \sum_{s=1}^{S} \pi_s C_s$. Thus *at the margin* the premium is fair, inducing the buyer to choose full cover, and the insurer covers its costs by making a lump sum charge in addition (this is known as a two-part tariff). Figure 2 gives a summary illustration of these results.

Figure 2 about here

The assumption that the insurance market is perfectly competitive implies that the optimal insurance contract is given by the solution to the following problem

$$\max_{P,C_s} \bar{u} = \sum_{s=0}^{S} \pi_s u(y_s) \tag{22}$$

$$s.t. \quad P = \pi_0 F + \sum_{s=1}^{S} \pi_s [C_s + K(C_s)] \tag{23}$$

$$C_s \geq 0 \quad s = 1, 2, ..., S \tag{24}$$

The Lagrange function is

$$L = \sum_{s=0}^{S} \pi_s u(y_s) + \lambda[P - \pi_0 F - \sum_{s=1}^{S} \pi_s (C_s + K(C_s)] \tag{25}$$

The first order conditions are

$$\frac{\partial L}{\partial C_s} = \pi_s u^{'}(y_s^*) - \lambda^* \pi_s [1 + K^{'}(q_s^*)] \leq 0 \;\; C_s^* \geq 0 \;\; C_s^* \frac{\partial L}{\partial C_s} = 0 \tag{26}$$

$$\frac{\partial L}{\partial P} = -[\pi_0 u^{'}(y_0^*) + \sum_{s=1}^{S} \pi_s u^{'}(y_s^*)] + \lambda^* = 0 \tag{27}$$

$$\frac{\partial L}{\partial \lambda} = P^* - \pi_0 F - \sum_{s=1}^{S} \pi_s [C_s^* + K(C_s^*)] = 0 \tag{28}$$

Note, we have assumed $P^* > 0$, which in turn requires that at least one $C_s^* > 0$, otherwise the problem is trivial. We now establish the main results.

It is first useful to prove that if cover is positive in state s, it must be positive in all higher loss states $s + 1, ..., S$

Proposition 1: *If* $C_s^* > 0$, *then* $C_{s+t}^* > 0$, *for all* $t = 1, 2, ..., S - s$

Proof: Suppose to the contrary that $C_s^* > 0$, and $C_{s+1}^* = 0$. Then the first order conditions give

$$u'(y_s^*) - \lambda^*[1 + K'(C_s^*)] = 0 \tag{29}$$

$$u'(y_{s+1}^*) - \lambda^*[1 + K'(0)] \leq 0 \tag{30}$$

This implies (since $K''(.) \geq 0$)

$$u'(y_s^*) = \lambda^*[1 + K'(C_s^*)] \geq \lambda^*[1 + K'(0)] \geq u'(y_{s+1}^*) \tag{31}$$

But this implies (given $u'' < 0$)

$$y_s^* = y - P^* - L_s + C_s^* \leq y - P^* - L_{s+1} = y_{s+1}^* \tag{32}$$

But this cannot be true if $L_s < L_{s+1}$ and $C_s^* > 0$. Thus we have a contradiction. The argument can then be repeated for $t = 2, .., S - s$.

We now show that zero marginal costs imply full cover.

Proposition 2. *There is full cover in every loss state if* $K'(.) = 0$

Proof: Set $K'(.) = 0$ in the above first order conditions. We first show that cover is positive in all loss states. Suppose not. In the light of Proposition 1, this must mean that cover is zero for loss states $1, .., t - 1$, and positive for states $t, ...S$, for any t between 1 and S. Note that we must have, since $u'' < 0$,

$$u'(y_0^*) < u'(y_1^*) < ... < u'(y_{t-1}^*) \leq \lambda^* \tag{33}$$

while in the states where cover is positive

$$u^{'}(y_s^*) = \lambda^* \quad s = t, .., S \tag{34}$$

This implies that incomes are equal across these loss states

$$y - P^* - L_t + C_t^* = ... = y - P^* - L_S + C_S^* \tag{35}$$

Denote this common income by y^*. Substituting in the second condition gives

$$u^{'}(y^*)[1 - \sum_{s=t}^{S} \pi_s] = \pi_0 u^{'}(y_0^*) + \sum_{s=1}^{t-1} \pi_s u^{'}(y_s^*) \tag{36}$$

But since

$$[1 - \sum_{s=t}^{S} \pi_s] = \pi_0 + \sum_{s=1}^{t-1} \pi_s \tag{37}$$

we can write this equation as

$$\pi_0[u^{'}(y_0^*) - u^{'}(y^*)] + \sum_{s=1}^{t-1} \pi_s[u^{'}(y_s^*) - u^{'}(y^*)] = 0 \tag{38}$$

which in the light of (33) and (34) cannot be true. Thus we have a contradiction and cover must be positive in all loss states, implying that $t = 1$.Using this in (36) then gives

$$(1 - \sum_{s=1}^{S} \pi_s) u^{'}(y^*) = \pi_0 u^{'}(y_0^*) \tag{39}$$

But since $(1 - \sum_{s=1}^{S} \pi_s) = \pi_0$, this implies $y_0^* = y - P^* = y - P^* - L_s + C_s^* = y^*$. But this can only hold if $C_s^* = L_s$. Thus we have full cover.

Note that if marginal costs are zero but there is a fixed cost F, the first order condition (28) gives the form of the premium. We now prove a central result:

Proposition 3: *If marginal costs are positive and constant, optimal insurance takes the form of full cover above a deductible, $D > 0$, or more precisely*

$$C_s^* = max(0, L_s - D) \tag{40}$$

Proof: Let cover be positive in states $s = t, ..., S$. Denoting the constant marginal costs by $K^{'}$, condition (26) becomes

$$u^{'}(y_s^*) = \lambda^*[1 + K^{'}] \quad s = t, ..., S \tag{41}$$

This implies that

$$y - P^* - L_t + C_t^* = y - P^* - L_{t+1} + C_{t+2}^* = ... = y - P^* - L_S + C_S^* \tag{42}$$

or

$$L_t - C_t^* = L_{t+1} - C_{t+1}^* = ... = L_S - C_S^* \equiv D \tag{43}$$

We call this common difference the deductible D. Thus we have $C_s^* = L_s - D$. It is easy to show $D > 0$, and this is left as an exercise.

However, cover need not be positive in all states, and it is interesting to see why. Thus suppose $C_s^* = 0$, while $C_{s+1}^* > 0, .., C_S^* > 0$. The conditions then become

$$u^{'}(y_s^*) \leq \lambda^*[1 + C^{'}] \tag{44}$$

$$u^{'}(y_t^*) = \lambda^*[1 + C^{'}] \quad t = s + 1, ..., S \tag{45}$$

Just as before, we can show that $L_t - C_t^* = ... = L_S - C_S^*$, and we again call this difference the deductible. These conditions now imply

$$u^{'}(y_s^*) \leq u^{'}(y_{s+1}^*) \tag{46}$$

and therefore

$$y - P^* - L_s \geq y - P^* - L_{s+1} + C^*_{s+1} \tag{47}$$

or

$$L_s \leq L_{s+1} - C^*_{s+1} = D \tag{48}$$

Thus there is no cover in loss state s (and therefore in all lower loss states) because the loss is no bigger than the optimal deductible.

Proposition 4: if marginal costs are positive and increasing there is coinsurance above a deductible.

Proof: We now have that $K^{''}(.) > 0$. Let cover be positive in loss states $s = t, ..., S$. Then the conditions become

$$u^{'}(y^*_s) = \lambda^*[1 + K^{'}(C^*_s)] \quad s = t, ..., S \tag{49}$$

Consider now the logical possibilities:

(i) y^*_s stays constant as L_s increases through $s = t, .., S$, as would be the case with a deductible. In that case $u^{'}(y^*_s)$ would stay constant while $K^{'}(C^*_s)$ increases, because C^*_s must increase. Then the condition cannot be satisfied, and so we rule this case out.

(ii) y^*_s increases as L_s increases. In that case $u^{'}(y^*_s)$ would decrease while $K^{'}(C^*_s)$ increases, because C^*_s must increase, thus again we can rule this case out

(iii) y^*_s falls as L_s increases. In that case $u^{'}(y^*_s)$ would increase. Provided C^*_s also increases, $K^{'}(C^*_s)$ will increase, so this case, and only this case, is consistent with the conditions.

Thus we have

$$y - P^* - L_t + C_t^* > y - P^* - L_{t+1} + C_{t+1}^* > ... > y - P^* - L_S + C_S^* \qquad (50)$$

implying

$$L_1 - C_t^* < L_{t+1} - C_{t+1}^* < ... < L_S - C_S^* \qquad (51)$$

but also

$$C_t^* < C_{t+1}^* < ... < C_S^* \qquad (52)$$

This then is the case of an "increasing deductible", or, if we define $D \equiv L_t - C_t^*$, as coinsurance above a deductible.

5 Limited Liability and Insurance Reserves

Under limited liability, a shareholder is liable for the debts of a company only up to the value of his shareholding. As we shall now see in a simple example, this may create an incentive for an insurer to provide insufficient reserves to cover loss claims, which in turn can be used as an argument for the regulation of insurance markets by a public agency. The example shows that given limited liability and an extreme form of asymmetric information, it could be in an insurer's interest to run a higher risk of insolvency than is desirable from the policyholder's point of view.

Consider an individual who faces a 10% chance of a loss of £1,000. The expected value of loss is £100, but because she is risk-averse, we assume she is prepared to pay a premium of £150 in return for full compensation in the event of loss. A risk-neutral insurer will certainly accept this. If the premium of £150 is paid at the beginning of the period, while the compensation would have to be paid at the end, and the interest rate is 10%, then, to be sure he is able to cover the loss, the insurer will need to put up a starting capital of £760, so that the initial investment of £910, capital plus premium income, will produce £1,000 at the end of the period. Note that putting up this capital does not involve any direct cost to the insurer, since he can invest it as "insurance company capital" at exactly the same rate, 10%, as if he invested it privately on the market. His end-of-period expected wealth is £900, his final capital less expected claims cost of £100. The expected present value of profit from the insurance business is £50, given by £150, the premium income, *minus* £100, the expected value of loss.

Suppose now that, *unknown to the policy-holder*, the insurer puts up no insurance capital, but instead invests his £760 privately. In the event of loss, he simply pays out £165 at the end of the period and declares the insurance company bankrupt. Then his end-of-period expected wealth is £984.50 (£836 for sure from his investment of the capital, plus $0.9 \times £165$), and the expected present value of profit from the insurance business is the net expected wealth gain of £135. By not putting up any capital the insurer simply *truncates the loss distribution* he faces, thus reducing the expected value of his claims liabilities. In that case his expected wealth gain is £85 higher than if he puts in enough reserves to ensure solvency. We will soon see that the basic point of this simple example can be shown to hold in much more general cases.

Two objections can be raised to this example. It may pay the insurer not to put any capital into the insurance business in this one-off case, but what if in fact he is in business "for the long term", *i.e.* the number of periods can be increased indefinitely? If he becomes insolvent, he loses the right to continue in the insurance business in the future, and the loss of future profits may be enough to induce him to put up capital to avoid insolvency today. In this example however this argument does not hold. If in every period the insurer puts the requisite capital into the company, his expected present value of profit over an infinite horizon with a 10% interest rate is £550. If he puts up no capital, and allows for the fact that in each period he runs a 10% chance of going out of business, his expected present value of profit is about £660. It is possible to construct realistic examples where the insurer would find it profitable to put up the required capital to avoid insolvency. Nevertheless, it remains the case that under quite plausible circumstances it pays the insurer to put up none of his own capital.

A more fundamental point concerns the buyer's information about the insurer's capital. In the above example, it was assumed that the insurance buyer believed that the insurer would meet her claim, otherwise she would not have bought insurance in the first place - she could have obtained exactly the same degree of coverage in the default case by herself investing the premium. Clearly, if the insurance buyer is fully informed about the default risk, it always pays the insurer to put up the capital, since otherwise he would not be able to sell insurance and would lose even the expected profit of £50. This point can be generalised: if the insurance buyers are fully informed about the risk of insolvency, so that *this is reflected in their willingness to pay* for insurance, then it always pays the insurer to put up enough capital to ensure losses can be met. The intuition is straightforward, and can be given most simply for the case of a risk neutral insurer (the insured is always risk averse). If there is an insolvency risk, the risk averse policy holder would always be prepared to pay more than the fair premium (expected value of loss) to buy insurance against this, and the insurer would always find it profitable to sell it to her. He can only do this however if he puts up enough capital to cover the loss.

We now generalise this example. We take an infinite time horizon, with a sequence of discrete time periods (say years). At the beginning of each period, the insurer must decide on a level of capital K for the insurance business, in the light of a given distribution of claims costs C, described by the distribution function $F(C)$ with (differentiable) density $f(C)$, defined over the interval $[0, C_{\max}]$. For the moment, we take it that premium income P is also exogenous, and in particular independent of the level of capital chosen. This assumes not only that insurance buyers are uninformed, but also that they do not perceive a relationship between the insurer's capital and the likelihood that their claim will be met. The premium income P is collected at the beginning of the period and invested along with the capital. The only capital market asset is a riskless security with gross return $r > 1$. If at the end of the period assets $A \equiv (P + K)r$ are at least enough to meet claims costs C, then the insurer remains in business and receives a continuation value V, that is the expected present value of being in the insurance business at the end of the first period. If claims costs turn out to be greater than assets, the insurer pays out his assets and defaults on the remaining claims, losing the right to the continuation value V. Because of limited liability he does not have to pay out to claimants more than A.

The insurer can always choose to guarantee solvency by putting in enough capital, since we have assumed that the upper limit $C_{\max}$ on possible claims is finite. The question of interest is: under what circumstances would the insurer choose to stay solvent, thus making regulation unnecessary?

We assume the insurer is risk neutral and the only cost of capital put into the insurance business is r, the riskless rate of return on the capital market. It follows that he maximises the expected present value of net wealth from the insurance business

$$V_0(K) = \int_0^A (\frac{V}{r} + K + P - \frac{C}{r})dF - K \;\; s.t. K \in [0, K_{\max}] \tag{53}$$

where $K_{\max} = (C_{\max}/r) - P$ is the capital required to ensure no default. Now since at the beginning of each period the future is identical, we have $V = V_0(K)$, and so using this in () and rearranging gives

$$V_0(K) = [\int_0^A (K + P - \frac{C}{r})dF - K]/(1 - \frac{F(A)}{r}) \tag{54}$$

So far nothing beyond differentiability has been assumed for the claims distribution F. Empirically however insurance claims distributions typically belong to the class of "increasing failure rate" distributions, with the property that

$$\frac{d}{dC}[\frac{1-F(C)}{f(C)}] < 0 \tag{55}$$

An important implication of this property is then that *only corner solutions to the insurer's wealth maximisation problem are possible*: either he chooses $K = 0$, or $K = K_{\max}$. We show this in

Proposition 1 *Proposition1: Given the property of the claims distribution in (55), any solution to the insurer's wealth maximisation problem is a corner solution.*

Proof. Suppose not, *i.e.* there exists a value $K^* \in (0, K_{\max})$ such that $V(K^*)$ is a maximum. Then $V_0'(K^*) = 0, V_0''(K^*) \leq 0$. Using (54) to evaluate these derivatives gives

$$V_0'(K^*) = [V_0(K^*)f - (1-F)]/(1-\frac{F}{r})] = 0 \tag{56}$$

$$V_0''(K^*) = r[V_0(K^*)f' + f)]/(1-\frac{F}{r})] \leq 0 \tag{57}$$

Then (56) implies

$$V_0(K^*) = (1-F)/f \tag{58}$$

while (57) implies

$$f^2 + f'(1-F) > 0 \tag{59}$$

and so substituting for $V_0(K^*)$ from (58) into (59) yields a contradiction.

Note that a solution to the problem does exist, since the objective function is continuous on the compact interval $[0, K_{\max}]$. Which endpoint is optimal is given by the straightforward comparison of the values

$$V_0(0) = F(rP)(P - \frac{\bar{C}_0}{r})r/[r - F(rP)] \tag{60}$$

$$V_0(K_{\max}) = (P - \frac{\bar{C}}{r})r/[r - 1] \tag{61}$$

where $\bar{C}$ is the mean of the claims distribution and $\bar{C}_0 = [F(rP)]^{-1} \int_0^{rP} C dF < \bar{C}$ is the mean of the truncated distribution. As these expressions clearly show, the advantage to not putting up any capital is that the expected present value of claims falls. The disadvantage is that there is a risk of going out of business, $F(rP) < 1$. It does not seem possible to say that one of these endpoints is always better than the other. Figure 3 illustrates the possibilities.

Figure 3 about here

There are two major limitations of this model of the insurance firm which could make any policy conclusions derived from it of limited relevance. The first is that the only assets available on the market are safe assets. An interesting question in relation to real insurance companies concerns the interaction between the risks associated with their asset portfolios and those associated with their insurance activities. The second limitation is the exogeneity of the premium income. This is not simply a matter of allowing the firm to choose the premium or volume of insurance sold by maximising profit with respect to a given demand function. More specifically it implies the assumption that *the demand for insurance is* independent of the seller's insolvency risk, which is clearly a strong and ultimately unacceptable assumption.