The Theory of Insurance Reconsidered for Urban Analysis: An Expected Utility Approach

by

Robert H. Edelstein*

Working Paper No. 11-72 (Revised)

RODNEY L. WHITE CENTER
FOR FINANCIAL RESEARCH

The Wharton School

University of Pennsylvania

Philadelphia, Pa. 19104

This essay attempts to explore the demand for and supply of insurance through the use of simplified mathematical models. The theoretical analysis uses an expected utility-portfolio approach in order to derive implications for static equilibrium in the insurance market. In addition to the theoretical models, there are suggested applications of the theory for urban core insurance problems. Some of the major conclusions of the essay are (1) given the apparent risks involved in urban core areas, the government will probably be called upon to be the insurer or re-insurer for significant quantities of property insurance, (2) governmental price regulation in the insurance market is likely to be a partial cause for inadequate urban core coverage, and (3) Fair Access to Insurance Requirements plans (FAIR) and Excess Rates plans either currently in effect or proposed will be non-optimal and inadequate solutions to urban property insurance problems.

The essay is presented in two principal parts, The Theory of Insurance and The Applications of the Theory of Insurance to Urban Problems. These parts are divided in several main sub-sections:

III - Private Market Approaches to the Solution

IV - Existing Public Programs as Solutions

The Theory of Insurance

Economic theorists recently have lucubrated about the relationship between asset preferences and expectations regarding future events or future opportunities.[1] Expectations are multi-faceted; a world in which individuals have expectations about future states or conditions generally can not be described by a simple "parametric" index. We are not affected merely by what we suspect is the most probable outcome of a set of possible outcomes, or even by the average outcome (i.e., expected value); our behavior is determined by our expectations about less probable, or even the most improbable of possible outcomes, as witnessed by the public's demand for Irish sweepstakes tickets or certain Canadian uranium mining companies' stock. One of the most important aspects demonstrating the interconnectedness of expectations and asset preferences in the real world is the demand for insurance protection, and the resulting robustness of insurance companies' wealth. Unfortunately, the theory of insurance demand and supply are somewhat neglected in the literature.[2] This paper will explore, through the use of simplified models, the theoretical determinants of the demand for and the supply of insurance, and the implications for static equilibrium in the insurance market. The theory will be applied in subsequent sections to several problems associated with urban areas and insurance arrangements.[3, 4]

I. Optimal Property Insurance for an Individual

Consider an individual who owns real property as one of his assets. His total wealth, W, consists of the sum of the values of his piece of real property, H, and all his other assets, A. In our simplified world, we will assume this individual believes he faces the risk of, say, theft of H with a probability of π; and all other assets are theft free (i.e., riskless). We shall also assume that theft results in complete, non-recoverable loss of H. If insurance against theft is available at the premium rate of ρ per dollar of insured value, with the stipulation that insured value can not be greater than the actual value, what will be the optimal insurance coverage?

Assume that the individual is an expected utility[5] maximizer with wealth being the independent variable of the utility function:

$$u = u\,(W) \tag{1}$$

The situation depicted above without insurance yields

$$W = \begin{cases} A \text{ with probability of } \pi \\ A + H \text{ with probability of } 1 - \pi \end{cases} \tag{2}$$

An insurance policy can be contracted such that the insurance company will reimburse you for λH if H is stolen ($0 \leq \lambda \leq 1$) for the modest cost of $\rho\,\lambda\,H$. That is, for a premium payment of $\rho\,\lambda\,H$, the wealth distribution will be

$$W = \begin{cases} A + \lambda H(1 - \rho) \text{ with probability } \pi \\ A + H(1 - \lambda\rho) \text{ with probability } 1 - \pi \end{cases} \tag{3}$$

In other words, the quest for optimal insurance coverage against theft is equivalent to discovering the optimal proportion of self- or co-insurance an individual should undertake (i.e., $0 \leq 1 - \lambda \leq 1$ is the proportion of self-insurance he indulges in) given his utility function, his subjective expectations about loss, the values of the risk and riskless assets, and the premium rate structure.

a) Insurance and the Range of Wealth

As an example consider the item to be insured against theft is a necklace valued at \$10,000. If the probability of theft is .01, without insurance your expected wealth would be A + \$9,900. The range of wealth would be \$10,000 (i.e., A + \$10,000 in the case of no theft and less A in the case of theft). An insurance policy costing \$1 per \$100 of coverage (i.e., an actuarial fair policy, given your expectations), and covering \$5,000 of value in case of theft will have a total premium of \$50. Expected wealth will be exactly the same at A + \$9,900 but the range of wealth will have been altered to \$5,000 (i.e., A + \$9,950 less A + \$4,950). Increments in insurance do not reduce expected wealth, while they do reduce the range of wealth by lowering the value of the upper extreme event and raising the lower extreme event. In our simple theft-protection model, the range of wealth R is

$$R = (1 - \lambda)H \tag{4}$$

and the expected wealth is

$$E(W) = \pi \cdot (A + \lambda H(1 - \rho)) + (1 - \pi)\cdot(A + H(1 - \lambda\rho)). \tag{5}$$

Clearly, the relevant measure of risk in our simple example is the range of wealth. Also, we would expect the insurance premium rate to be greater than \$1 per \$100 of protected value.[6]

b) Full Coverage and Maximum Premium Rate

What would be the maximum premium rate one would be willing to pay for protection, and accept complete coverage against theft? The value of the maximum premium rate would occur when the expected utility of the non-insured risk situation was equal to the utility of the completely insured situation (i.e., $\lambda = 1$). Mathematically we can discover the appropriate $\rho = \rho^*$ by solving the following equation.

$$\pi\, u(A) + (1-\pi)\, u(A+H) = u(A+H(1-\rho^*)) \qquad (6)$$

If we assume that marginal utility of wealth is positive, and it is decreasing as wealth increases,[7] it is clear that ρ^* will be actuarially unfair in favor of the insurance company. Figure 1 demonstrates this graphically. An actuarially fair premium rate for full coverage would reduce the range to zero with a wealth level (for certain) of $\pi A + (1-\pi)(A+H) = C$. This yields an expected utility of E' versus E for the uninsured situation. The maximum an individual would be willing to pay for full coverage insurance (i.e., no range of wealth) has to yield the identical utility as the risk situation, E. At position C a premium rate $\rho = \pi$, or a total premium payment for full-coverage of $\pi H = \rho H$ represents the actuarially fair policy. At point B, the total insurance bill would be C-B greater than πH. Therefore,

$$\rho^* = \frac{C-B}{H} + \pi \qquad (7)$$

If there are no administrative costs, $\frac{C-B}{H}$ represents "pure" expected return above the expected costs per dollar of coverage.

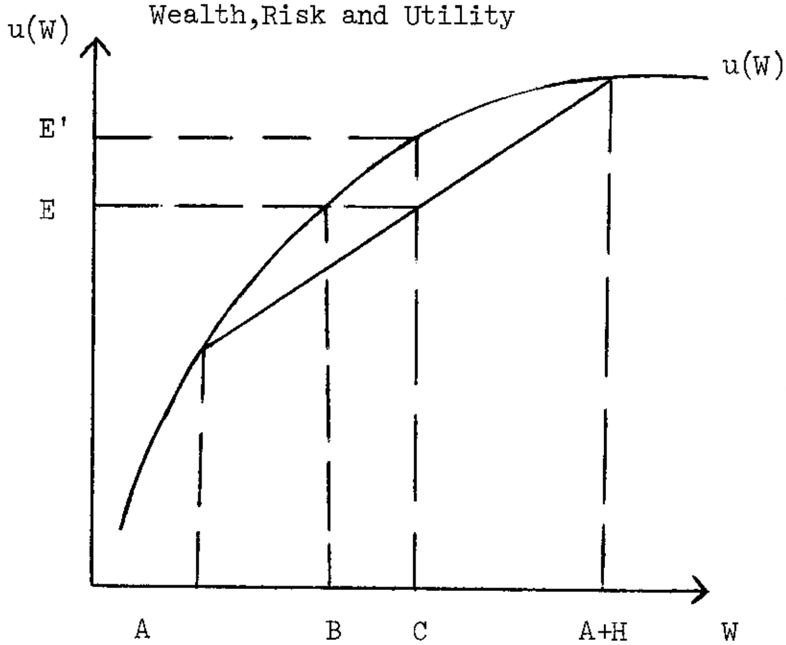
c) States of the World and Uncertainty

The individual attempting to maximize his expected utility is presumed to have an immutable subjective notion of π, the probability of theft. This assumption[8] is not essential for reaching our analytic conclusions, but is maintained in order to simplify our graphics and mathematics. For example, suppose the probability of theft during the future period was considered dependent upon who was in charge of the police force. That is, if the mayor next year were the "Law and Order" candidate rather than the "Rad-Symp" candidate, presumably he would reduce burglaries through more stringent law enforcement policies. The probability of election of a particular candidate in our example is called the "state" probability, and the sum of the state probabilities must be unity (i.e., some one will be elected mayor). If there are only the two aforementioned candidates, and we were to assess the probability of election of the Law and Order candidate as P, the probability of the Rad-Symp candidate becoming mayor is (1-P). In addition, if the Rad-Symp wins our conditional probability of theft is π_2, and conditional upon the Law and Order candidate's election, our theft probability is π_1. Presumably, with $0 \leq \pi_1 < \pi_2 \leq 1$, our problem of finding an appropriate insurance coverage solution can be conducted as before. The maximum premium rate ρ^* that an individual would be willing to pay for complete insurance coverage is dependent upon U, P, π_1, π_2, A and H, and can be discovered by the solution of

$$P[\pi_1 U(A) + (1-\pi_1) U(A+H)] + (1-P)[\pi_2 U(A)+(1-\pi_2) U(A+H)]=U(A+H(1-\rho^*)) \quad (8)$$

Equation (8) represents the balancing of total expected utility from the uninsured situations with the two states, each with different underlying theft probabilities, and the utility derived from the completely insured situation. The uninsured situation is equivalent to a double lottery: the probability of participating in lottery 1 is P, and the probability of participating in lottery 2 is 1-P. If lottery 1 obtains, we can discover its expected utility using π_1 (as we did for π in figure 1 above) yielding an expected utility level (such as E); lottery 2 can be

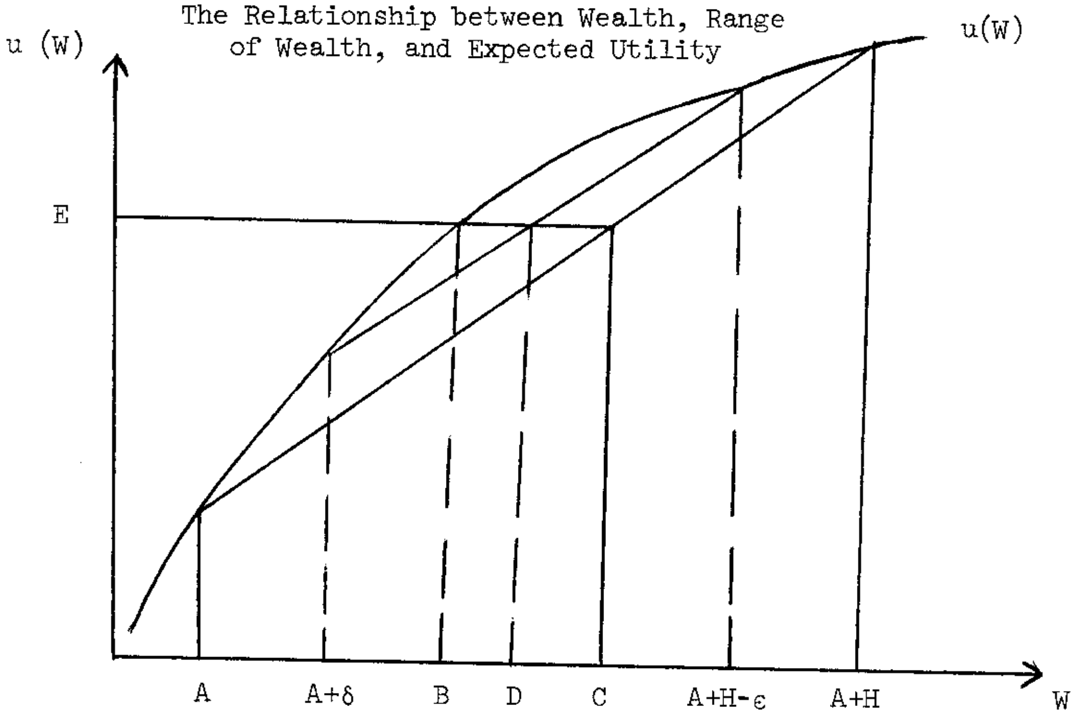
Figure 1
The Relationship between
Wealth, Risk and Utility



$B = A + H\ (1-\rho^*)$

$C = \pi A + (1-\pi)\ (A+H)$

Figure 2
The Relationship between Wealth, Range
of Wealth, and Expected Utility



$D = \pi(A+\delta) + (1-\pi) \cdot (A+H-\varepsilon)$

$C = \pi \cdot A + (1-\pi) \cdot (A+H)$

B = Minimum wealth for maintaining $E(U) = E$
(no range of wealth)

analyzed similarly using π_2. The expected utilities of these two lotteries are then treated as the utility extremes (i.e., such as U(A) and U(A+H) in figure 1.). The combined expected utility of the two lotteries is the sum of the expected utilities of each of the lotteries weighted according to its state probability.

In essence, uncertainty about the true value of π does not affect our analysis, if we assume that the individual has a subjective assessment of the probability distribution for the potential values of π. The uncertainty about π can be transformed into an expected utility analysis with the state probabilities weighting each expected utility combination for appropriate values of π.[9] Therefore, the analysis presented can be extended without additional theoretical concern to conditions of uncertainty.

There is another important case in which π may not be constant. The probability of theft may be altered by behavior of the insured, such as the installation of more effective locks or placing valuables in safety vaults. Put differently, the insured may expend resources in order to reduce the probability of the undesired event occuring. We will defer discussion about the relationship between actions which alter probabilities and the optimal insurance purchasing policy to a subsequent section.

d) The Maximum Premium Rate and Parametric Changes

It is interesting to note the effects of a change in one of the given parameters on the value of ρ^*. If the probability π increases, say, because of an alteration in expectations, the maximum ρ^* will increase. This follows because when π increases, the expected loss from the uninsured situation (πH) will increase causing expected wealth and expected utility to decrease. Therefore, the insurer could react by increasing ρ^* in order to cover at least part of the greater expected loss. The insurer can conduct this new policy as an individual's expected utility is decreased by the altered probability distribution of π and therefore decreased expected wealth because the insured is willing to pay a larger proportion of his maximum wealth (A+H) in order to reduce the range of wealth to zero. This could be seen in figure 1: if π were to increase, points B and C would shift to the left causing E to decline, and implying ρ^* will increase.[10]

It is also clear that as H, the value of the risky asset increases, the total an individual is willing to pay for full-insurance will increase.[11] Of more interest is the effect of change in the value of A upon the maximum premium rate. If A constitutes the bulk of the wealth of the individual, we will be tracing the "wealth" effect upon his willingness to purchase full-insurance coverage.

If we totally differentiate (6) and solve:

$$\frac{d\rho^*}{dA} = \frac{-(\pi\ u'(A) + (1-\pi)\ u'(A+H) - u'(A+H(1-\rho^*))}{H \cdot u'(A+H(1-\rho^*))} \tag{9}$$

The denominator is always positive. Therefore, the sign of equation (9) depends upon the sign of the numerator, which can be shown to be negative if we assume decreasing risk aversion.[12] Therefore, an individual's maximum total payment and premium rate for full-insurance coverage against the theft of H will decrease as his total wealth (H constant) increases.

II. Expected Utility and Variable Coverage

a) Expected Utility as a Function of Range and Wealth.

Expected utility has been shown to be a convex linear combination of the utility levels of the extremes of wealth. Also, if we reduce the range of wealth the individual can remain as well-off by lowering his expected wealth level. In other words, an individual who has the assumed utility function will pay something to eliminate the vagaries of wealth associated with risk assets through an insurance program. Using Figure 2, we can demonstrate this in the Wealth-Range plane. In this problem an individual will pay to reduce the range, which is the relevant measure of risk of total wealth.

$$E(U) = \pi\, U(A+\delta) + (1-\pi)\, U(A+H-\epsilon) \tag{10}$$

where for each $0 \leq \delta$, we can solve for an appropriate $\epsilon \geq 0$.[13] For any value of expected utility given the utility function on wealth, A, H, and π, there exists a unique ϵ for each δ, such that:

$$\pi(A+\delta) + (1-\pi)(A+H-\epsilon) \leq \pi \cdot A+(1-\pi)(A+H).$$[14]

That is, the expected wealth which maintains the same level of satisfaction will decline as the range of wealth declines. Finally, expected wealth will be a minimum for the given utility level when the range is identically zero (i.e., B in figure 2). In figure 3, this information has been transformed into the W - R plane. Additional indifference curves could be derived in the same manner.[15] An example is illustrated on figure 3 as B'E'. Note that B'E' is clearly preferred to BE because at each level of expected wealth it has a smaller range, and at each level of range, it has a higher expected level of wealth.

b) Opportunities to Buy Insurance

For the moment, assume that insurance protection against theft can be bought according to the insurer's book rate of $\$\rho$ per dollar of coverage. Expected wealth will be a function of the proportion of H you choose to insure. In general:

$$E(W) = \pi(A+\lambda H(1-\rho)) + (1-\pi)(A+H(1-\rho\lambda)) \tag{11}$$

The range of wealth varies according to your coverage:

$$R = H(1-\lambda) \tag{12}$$

The opportunity line for purchasing insurance can be found by solving (12) for λ and substituting into (11):

$$E(W) = A+H(1-\rho) + H(1-\lambda)(\rho-\pi)$$

or (13)

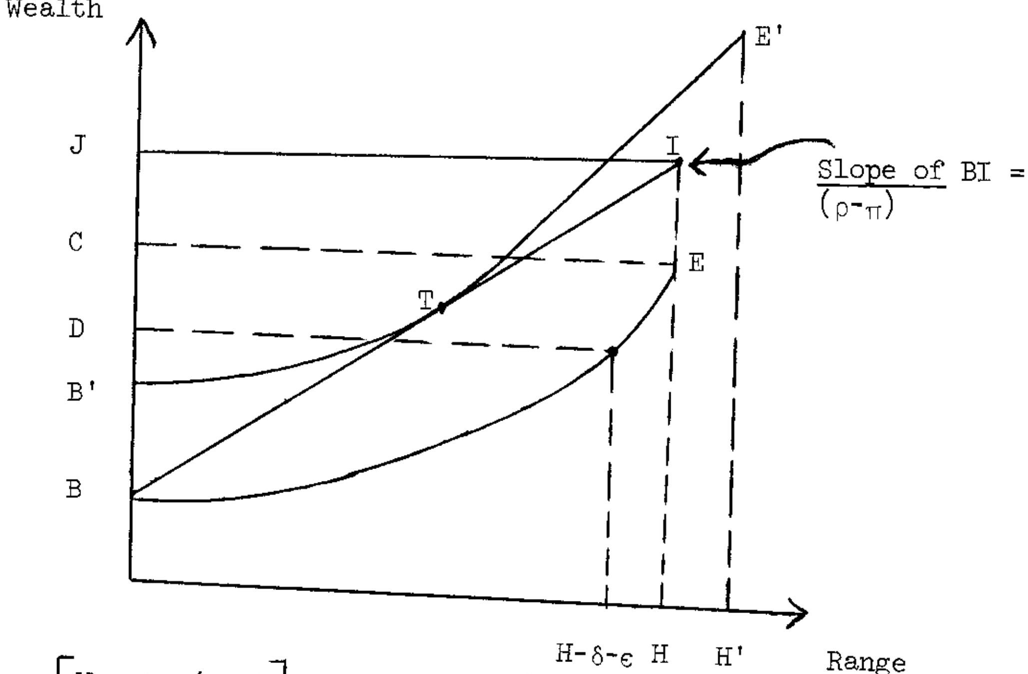
$$E(W) = (A+H(1-\rho)) + R(\rho-\pi)$$

Figure 3 depicts equation (13).[16] Point B represents full-insurance (i.e., the range is reduced to zero) and point I shows the totaly uninsured individual. Note if the insurer charges a premium rate exactly equal to the actuarial cost π, IJ will be the opportunity line. (In equation (13), the range term becomes identically zero when $\rho = \pi$.) In general, we would expect insurance companies to charge more than the acturarial costs of insuring, therefore $\rho > \pi$, and BI will be a straight line with a positive slope.

However, it was stated earlier that π was a subjective evaluation of the probability of theft. Moreover, it is possible for individuals to systematically misestimate π. We assume that insurance companies with their wider range of insuring experience can predict accurately the true loss probabilities, and charge higher rates (ρ) than the true loss probabilities in order to cover administrative costs as well as to clear a profit and accumulate cash reserves. Therefore, it is possible for BI to be downward sloping because π is substantially overestimated

Figure 3

Wealth-Range Indifference Map, Opportunity-Insurance Curve, and Optimal Insurance Purchasing

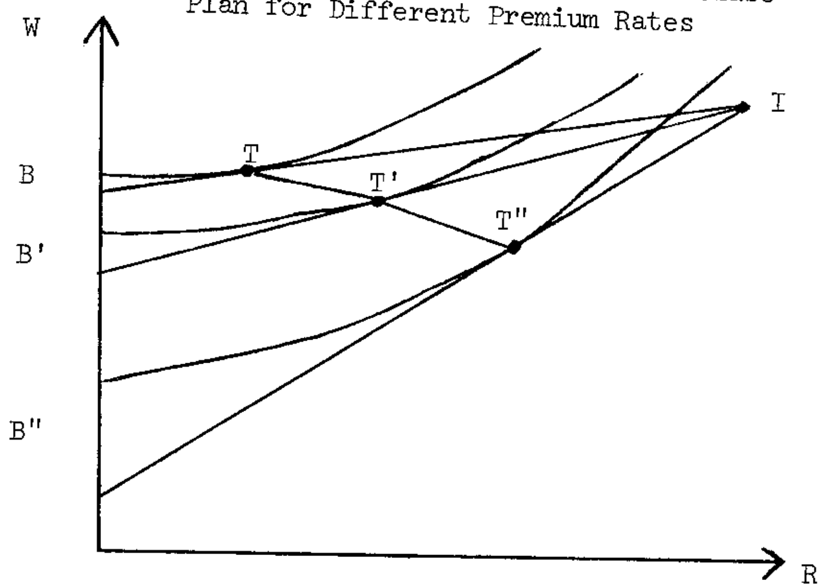


Point B $\begin{Bmatrix} W = A+H(1-\rho) \\ R = 0 \end{Bmatrix}$

Point I $\begin{Bmatrix} W = A+H(1-\pi) \\ R = H \end{Bmatrix}$

Figure 5

An Individual's Optimal Insurance Purchase Plan for Different Premium Rates



(i.e., $\pi > \rho$) by individuals seeking coverage, even after insurance company loadings are considered.

c) Choosing Your Optimal Coverage

A very strong distaste for risk would imply that the individual's indifference curves would be steeper than the opportunity line at all points. Therefore, this individual might engage in a fully insured program where risk is reduced to zero. Clearly, an individual might choose to insure a fraction of H rather than all or nothing. An individual who finds "risky" situations disquieting is likely to insure a larger proportion of the risky asset's value than one who possesses more of a gambler's nature. That is, if he has distaste for risk, it is possible that he will engage in some partial coverage, such as point T in figure 3. Point T has the same expected utility as any other point on B'E', which mathematically is

$$E(U) = \text{Constant} = \pi U(A+\lambda H(1-\rho)) + (1-\pi)U(A+H(1-\rho\lambda))^{17} \quad (14)$$

For small changes around the optimal λ, the proportion of H that would be insured would obey the following "marginal" rule:

$$\pi(1-\rho)U'(A+\lambda H(1-\rho)) - \rho(1-\pi)U'(A+H(1-\lambda\rho)) = 0 \quad (15)$$

The first term of (15) indicates the gain in expected marginal utility that accrues to the insured for the last dollar of coverage at cost ρ, leaving $(1 - \rho)$, if the theft occurs with the probability π. When the theft does not occur (i.e., the probability of this event is $1 - \pi$), the last dollar of insurance coverage, ρ, represents a loss of wealth and utility. Therefore, the second term represents the expected loss of utility at the margin caused by the purchase of insurance when no theft occurs. The first term is a gain in expected utility, and in equilibrium, it will be balanced by the second term, the loss in expected utility. That is, the costs and benefits weighted by the appropriate probabilities and marginal utilities will be precisely equal in equilibrium.

Graphically, the optimum at point T is tangent to the opportunity line BI, with the slope of each being $(\rho - \pi)$. Note that point T has a certainty equivalence of B', which is greater than B, the point representing full-insurance, and is simultaneously on a higher utility-indifference curve than point I, which represents no insurance.

Mathematically, the optimal insurance purchasing plan by the individual is discovered by differentiating (14) with respect to λ, yielding (16) which is equated with zero if we assume a tangency solution and second order conditions are satisfied.[18]

$$\frac{dE(U)}{d\lambda} = \pi H(1-\rho)U'(A+\lambda H(1-\rho))-(1-\pi)\, H\rho U'(A+H(1-\lambda\rho)) = 0 \qquad (16)$$

Note that the tangency solution (16) yields our marginal rule (15) for individual behavior. Of particular interest are the conditions for full coverage insurance and no coverage. For complete coverage,

$$\frac{dE(U)}{d\lambda} = (\pi - \rho)\ HU'(A+H(1 - \rho)) \geq 0 \text{ for } \lambda = 1$$

That is, full coverage will occur if $\rho \leq \pi$, which implies that the premium <u>is no worse than</u> actuarially fair, given the subjective evaluation of theft by the insured. A corollary of this is that if the premium rate is <u>not at least</u> actuarially fair, given the subjective evaluation of theft by the insured, less than full-insurance coverage is optimal. Also, given any $0 < \pi \leq 1$, there exists some $\rho > 0$, which will engender complete insurance coverage by the individual.

Similarly, we can state that no insurance coverage is optimal (i.e.,

$\lambda = 0$) if

$$\frac{dE(U)}{d\lambda} = \pi(1-\rho)HU'(A)-\rho(1-\pi)U'(A+H) \leq 0.$$

Clearly this condition is impossible if $\pi = 1$. In other words, if the subjective belief of theft is certain some insurance is optimal. (If $\rho < 1$, our case above for full insurance is met.) However, if $1 > \pi > 0$, one cannot state a priori if being uninsured is optimal. As a corollary of this, given $1 > \pi > 0$, it is always possible to set some rate $\rho > 0$ for which no insurance coverage will be optimal. For $\pi = 0$, the case when one believes theft is impossible, no insurance coverage is optimal.

d) Variable Insurance Coverage and Wealth

We will assume that an individual has purchased some insurance coverage (i.e., $0 < \lambda < 1$). That is, given his total wealth, asset composition between risky and riskless assets, utility function, subjective belief about theft probabilities, and the premium rate structure, he has chosen his optimal coverage such that $0 < \lambda H < H$. It is interesting to analyze the effects of a change in wealth, A, upon the individual's optimal insurance coverage. This can be accomplished by differentiating (16), the first order optimization condition, with respect to A. This yields

$$\frac{d\lambda}{dA} = \frac{\rho(1-\pi)U''(A+H(1-\lambda\rho)) - \pi(1-\delta)U''(A+\lambda H(1-\rho))}{H(\rho^2(1-\pi)U''(A+H(1-\lambda\rho))+\lambda(1-\rho)U''(A+\lambda H(1-\rho)))} \tag{17}$$

The denominator of (17) is clearly negative because $U''(W) < 0$. Hence, the total effect of (17) depends upon the sign of the numerator. This can be discovered by substituting from (18) into the numerator, $G(\lambda)$, for π and $(1-\pi)$:

$G(\lambda) =$

$$\frac{\rho(1-\rho)U'(A+\lambda H(1-\rho))U'(A+H(1-\lambda\rho))}{(1-\rho)U'(A+\lambda H(1-\rho))+\rho U'(A+H(1-\lambda\rho))}\left[\frac{-U''(A+\lambda H(1-\rho))}{U'(A+\lambda H(1-\rho))}-\frac{-U''(A+H(1-\lambda\rho))}{U'(A+H(1-\lambda\rho))}\right] \quad (18)$$

It is clear that if we assume decreasing risk aversion, (18) will be positive because the bracketed term is positive since

$$A+\lambda H(1-\rho) < A+H(1-\lambda\rho) \text{ for } \lambda \in (0,1).$$

Hence, (17) will be negative, signifying that an increase in total wealth (H constant) will lead to a decrease in the optimal coverage λ. In other words, ceteris paribus richer individuals optimally co-insure more than poorer ones. In micro-economic terminology, the insurance-consumption/wealth function is a declining function of wealth for risk averters.

e) Variable Insurance Coverage and the Price of Insurance

The quantity of insurance purchased is related to price per dollar of coverage, ρ, by the demand function. Changes in the price ρ, ceteris paribus, will generally be accompanied by an adjustment in the quantity of insurance protection, λH, which is considered optimal. Assume $0 < \lambda < 1$ is the current proportion of H which is insured. One can examine the static equilibrium adjustment by an individual in terms of λ by differentiating (16) with respect to ρ:[19]

$\frac{d\lambda}{d\rho} =$

$$\frac{\pi U'(A+\lambda H(1-\rho))-(1-\pi)U'(A+H(1-\lambda\rho))+\lambda H(\pi(1-\rho)U''(A+\lambda H(1-\rho))-\rho(1-\pi)U''(A+H(1-\lambda\rho)))}{H(1-\rho)^2\pi U''(A+\lambda H(1-\rho))+H\rho^2(1-\pi)U''(A+H(1-\lambda\rho))} \quad (19)$$

If $U''(W) < 0$ the denominator is always negative for $0 < \lambda < 1$. Hence, the net result of a change in ρ upon λ will depend upon the numerator's sign. Define $T(\lambda)$ as the numerator of (19). Then substitute for π and

$(1 - \pi)$ from (16), the original static equilibrium solution, into $T(\lambda)$ yielding:

$T(\lambda) =$

$$\frac{\lambda H\rho(1-\rho)U'(A+\lambda H(1-\rho))U'(A+H(1-\lambda\rho))}{(1-\rho)U'(A+\lambda H(1-\rho))+\rho U'(A+H(1-\lambda\rho))}\left[\frac{1}{H\lambda} + \left(\frac{-U''(A+H(1-\lambda\rho))}{U'(A+H(1-\lambda\rho))} - \frac{-U''(A+\lambda H(1-\rho))}{U'(A+\lambda H(1-\rho))}\right)\right] \quad (20)$$

for $\lambda \in (0,1)$. The sign of $T(\lambda)$ is not determinate. The term outside the brackets is constant and positive. The first term inside the brackets is positive. The sum of the two terms in the parenthesis inside the brackets, assuming decreasing risk aversion, is negative because

$$A + \lambda H(1-\rho) < A + H(1-\lambda\rho) \text{ for } 0 < \lambda < 1.$$

Therefore, a change in ρ can either increase or decrease λ in (19). Only if the terms in the brackets are a positive sum will the insurance coverage decrease with an increased ρ. It should be noted therefore that $U''(W) < 0$ is not a necessary or sufficient condition for a normal demand curve. Taking the term outside of the bracket in $T(\lambda)$ and dividing it by the denominator yields a function $F(\lambda) < 0$. Therefore (19) can be represented as

$$\frac{d\lambda}{d\rho} = \frac{F(\lambda)}{H\lambda} + F(\lambda)(R(A+H(1-\lambda\rho)) - R(A+\lambda H(1-\rho))) \quad (21)$$

where $R(W)$ is the risk aversion function of wealth. The first term on the right hand side is negative, and is similar to the substitution effect. For example, an increase in ρ for a person who has constant risk aversion (i.e., bracketed term in (21) is identically zero), will cause a decrease in his purchase plan for insurance. The wealth-security effect (the second term on the RHS of 21) for risk averters is always positive. This occurs because an increase in ρ effectively reduces wealth at each λ; that is, it reduces the opportunity to have both expected wealth and the certainty of wealth. Moreover, a risk-averter with this apparent decrease in wealth will place a higher subjective weight upon security. The obverse holds true, too. A decrease in ρ is equivalent to an opportunity to buy increased security and increased expected wealth. However, increased wealth with decreasing risk aversion leads the individual to reduce his insurance coverage because security in the form of insurance protection of H becomes subjectively less attractive. In addition, the wealth-security effect will diminish as H becomes smaller in absolute value as well as a proportion of total wealth. On the other hand, the smaller H is the larger will be the substitution effect.

If the wealth-security effect is weak relative to the substitution effect (assuming decreasing risk aversion), an increase in ρ will result in a decrease in λ. In the parlance of micro-exonomics, insurance will be a normal good. This case is demonstrated in Figure 4. The initial insurance opportunity line BI is tangent to the utility map at X. A decrease in ρ, alters the slope of the insurance opportunity to NI (see equation (13) above). The new tangency solution for insurance occurs at Z, and represents increased insurance protection (i.e., the range, $H(1-\lambda)$, is reduced). This equilibrium adjustment from X to Z may be decomposed into a wealth-security effect of XY where W'W' is constructed parallel to BI and tangent to B"E", the level of satisfaction equal to point Z, and a price-substitution effect of Y to Z. As previously noted, the substitution effect is negative (i.e., a decrease in ρ leads to an increase in λ), and the wealth-security effect is positive (i.e., decrease in ρ leads to a decrease in λ). Moreover, the substitution effect overwhelms the wealth-security effect, leading to a net increase in insurance coverage with a decrease in the premium rate.

f) Market Demand for Insurance

A typical individual's demand for insurance can be analyzed using Figure 5. We will suppose that an increase in the premium rate, his subjective expectations being constant, will result in a steepening of the opportunity line BI through successive positions B'I and B"I. The equilibrium points trace out a locus T, T', and T", indicating that the higher premium rate results in a smaller proportion of the risky asset being insured. The total market demand for insurance can be determined by summing the demand for all individuals at each ρ.

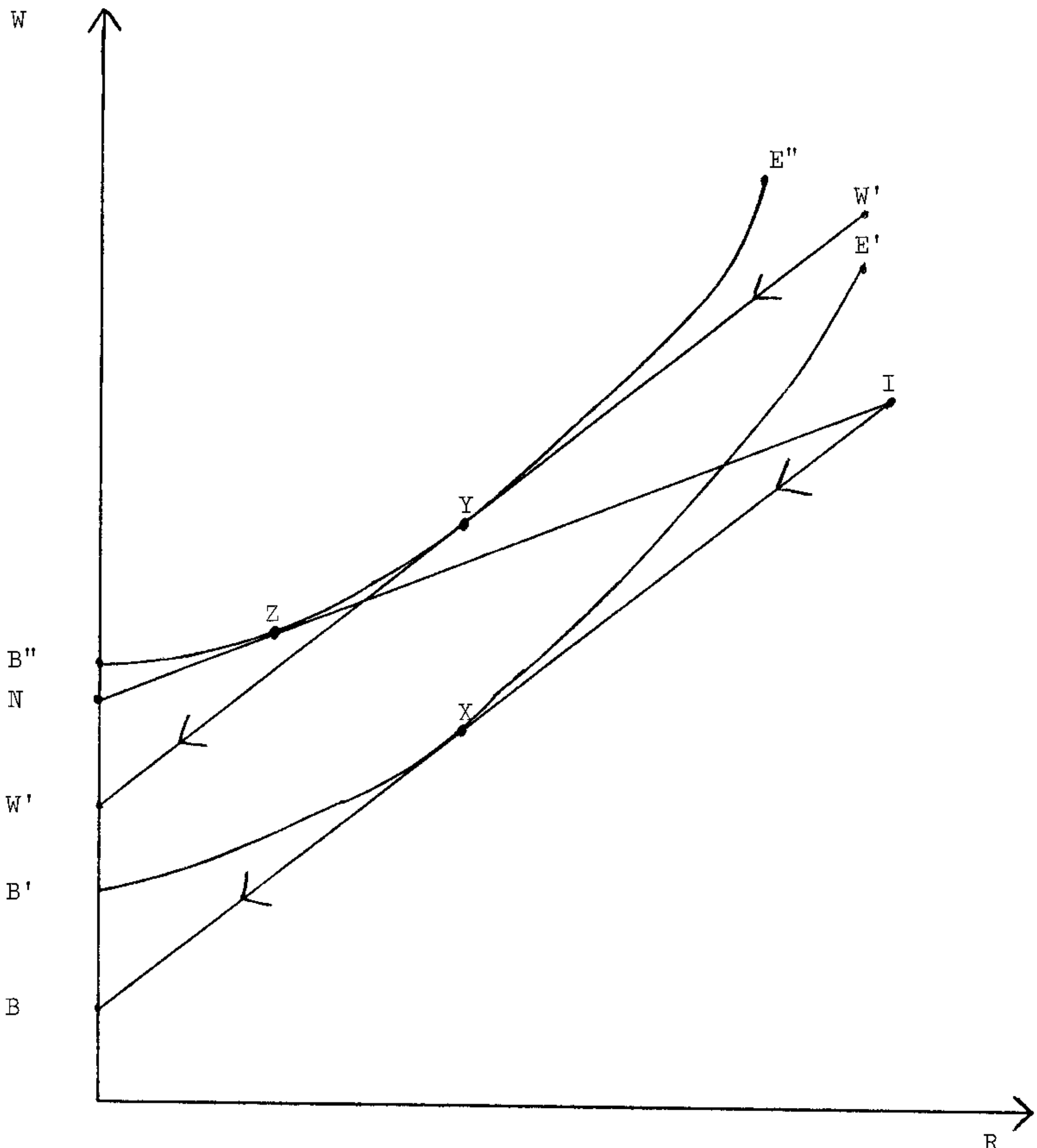
Casual empirical observation indicates that our model of insurance demand has several apparent drawbacks:

(1) Insurance companies do charge actuarially "unfair" rates, which, according to our analysis should lead to less than full coverage. This is not the way people appear to behave in the "real-world." There are at least two plausible explanations for this. First, people do not "balance at the margin." In other words, they are not rational economic decision-makers. Second, claims are not limited to the value of the object or the exact value that will be paid by the insurance upon submission of a claim is also not always clear. (Also, we assume negligible transaction costs.)

(2) Insurance purchasing does change behavior. Many an auto fender-dent has been repaired because the individual was "covered." The medical insurance industry has been confronted with a similar type of problem: It takes the form sometimes of a patient going to a doctor, when he would

Figure 4

Insurance Coverage and Price Changes



XY = Wealth-Security Effect

YZ = Price Substitution Effect

W'W' is parallel to BI

BI is initial opportunity line

NI is new opportunity line

not have if insurance did not pay for the visit; perhaps a doctor will administer more medical services to a patient who is covered by insurance than he might have otherwise. All of these can be classified as problems of "Moral Hazard."[20] Our analysis has by-passed this issue, but it is not difficult to understand that moral hazard is the resultant from inadequate economic or — philosophically speaking — moral incentives. In our theft problem, thieves might be more likely to steal from insured people than from uninsured people because "insurance companies are large and can absorb the loss." To correct this, an insurance company might insist upon, say, an alarm system against theft as a precondition for a policy contract.

We will discuss this type of situation in section (i) below from a slightly different but related vantage point. That is, we shall examine how an individual's optimal insurance purchasing program will be affected by the possibilities of actions which make the undesired event either less likely to occur or, if it occurs, less costly to the individual. It will be shown that our analysis can readily include such circumstances without difficulty; in general we omit it from our analysis in order to simplify our presentation.

(3) The analysis does not take into account either variable loss (i.e., thefts resulting in less than full loss) and deductibility clauses. Our next two sections are directed specifically to these issues. It will be shown that both characteristics can be included without vitiating our general conclusions.

g) Variable Random Loss and Optimal Insurance Coverage

The magnitude of the loss caused by, say, theft might itself be a random event. When a theft occurs, it is possible that a portion of your possessions will be stolen, while the remainder is left untouched. This type of situation is relevant to most automobile accident insurance, medical illness coverage, and fire insurance, where partial damage is possible. Fortunately, variable random loss can readily be incorporated into our analysis. It has a marked similarity to the discussion about uncertainty and probability states above in section I-c.

Each possible random loss, L_i, where $0 \leq L_i \leq H$, is associated with a state probability $\pi_i \geq 0$. The expected utility of the complete self-insurance (i.e., no insurance purchased) is

$$E(U) = \Sigma\, \pi_i\; U(A+H-L_i)$$

with the side condition $\Sigma\, \pi_i = 1$. An insurance policy can be introduced into the expected utility analysis as before

$$E(U) = \sum_i \pi_i\; U(A+H(1-\rho\lambda)) + \sum_h \pi_h\; U(A+H(1+\lambda(1-\rho))-L_h) \qquad (22)$$

where $\Sigma\, \pi_i + \Sigma\, \pi_h = 1$; $\pi_i, \pi_h \geq 0$ and $i \in I$ s.t. $0 \leq L_i \leq \lambda H$ and $h \in I$ s.t. $\lambda H < L_h \leq H$. The probabilities π are partitioned into two subsets i and h according to the proportion of H that is insured. Subset i represents loss events (including $L = 0$), which are completely insured; subset h represents loss events which are not entirely covered through insurance. If an event h occurs, the loss is greater than λH resulting in a net loss to the insured of $L - \lambda H$.

The optimal insurance policy can be found by differentiating (22) with respect to λ:

$$\frac{dE(U)}{d\lambda} = -\,\rho H\, \Sigma\pi_i U'(A+H(1-\lambda\rho)) + H(1-\rho)\Sigma\pi_h U'(A+H(1+\lambda(1-\rho))-L_h) \qquad (23)$$

The terms containing π_i represent, up to the margin, the loss in expected utility by having paid for insurance which is more than adequate to cover losses. The terms containing π_h represent the gain of expected utility caused by disbursement by the insurance company to the individual for the last marginal dollar spent on coverage, conditioned on a theft larger than λH occurring. This is completely analogous to the binary loss model optimum, equation (16). In equilibrium, (23) will be zero for the optimal λ.[22]

The condition for complete self-insurance is also similar to our earlier results. Evaluating (23), at $\lambda = 0$ yields:

$$- \rho H \Sigma\pi_i U'(A+H) + H(1-\rho)\Sigma\pi_h U'(A+H-L_h)$$

which must be non-positive[23] to be a static equilibrium solution. Note that $\Sigma\pi_i$ is the probability that no theft will occur, and clearly if that sum is unity our condition is met for self-insurance. If $\Sigma\pi_h > 0$, *a priori* one can not claim that complete self-insurance is optimal. On the other hand, if $\Sigma\pi_h = 1$, it is always optimal to purchase some insurance for $0 \leq \rho < 1$.

Full-insurance coverage requires that (23), evaluated at $\lambda = 1$, be non-negative. That is,

$$HU'(A+H(1-\rho)) [\Sigma\pi_h - \rho] \geq 0.$$

For this to occur, the probability of any type of loss must be no less than the premium rate. However, this is to say that the premium is *at worst* actuarially fair, which is analogous to our results in the binary model. In like fashion, the other results derived from the binary-theft-no-theft model apply to the random-variable loss model.[24]

h) Deductibility Clauses and Insurance Coverage

Another common feature of insurance plans for medical, fire, theft or automobile coverage is a deductible clause, where the insured pays the first d dollars of a loss, and the insurer pays a prespecified proportion of the loss above d. Returning to our binary-theft-no-theft model,[25] this can be accounted for in our expected utility analysis:

$$E(U) = \pi U(A+\lambda H(1-\rho(d))-d) + (1-\pi)U(A+H(1-\lambda\rho(d))) \qquad (24)$$

where $\rho(d)$ is the premium rate charged per dollar of coverage above the deductible d, and, as before, λ is the proportion of H, the risk asset, which is insured (above the deductible). $\rho'(d) < 0$, denoting that the larger deductible is associated with a less expensive premium rate.

Professor Arrow has demonstrated, if insurance is purchased, the optimal insurance policy for the risk-averting individual is a full-coverage program above some deductible, depending only that all actuarially equivalent insurance risks have equal total premium bills.[25] (Of course, this does not preclude that the deductible be identically zero, and,therefore,the insured has full coverage against loss.) Hence, our analysis will fix $\lambda = 1$ and discover the appropriate deductible, d, by differentiating (24):

$$\frac{\partial E(U)}{\partial d} = -\pi U'(A+H(1-\rho(d))-d)[H\rho'(d)+1]-H\rho'(d)(1-\pi)U'(A+H(1-\rho(d)))\geq 0 \qquad (25)$$

Let $X = A+H(1-\rho(d))-d \leq A+H(1-\rho) = Y$, the following holds upon rearrangement of (25), where clearly

$$U'(X) \geq U'(Y) \text{ and } U'(X)(1-\delta) = U'(Y)$$

for some $\delta \geq 0$ and $0 \leq d \leq H$.

$$[-\rho'(d)H(1-\delta \cdot (1-\pi))-\pi]U'(X) \geq 0 \qquad (26)$$

In equilibrium, an individual will have full-coverage above the deductible, and the premium rate need not be actuarially fair. That is, if $\pi-\rho(d)<0$, the bracketed term may be positive. In other words, the premium rate has to be not too unfair in order that equation (26) be positive. This conclusion appears to be contrary to our earlier results, which claimed that full-coverage (without a deductible) could be optimal only when $\pi \geq \rho$. It is not because, if $d = 0$, $U'(X) = U'(Y)$, which implies $\delta = 0$, and is consistent with the previous conclusion for full-coverage insurance, that is, $\pi \geq \rho$.

i) Optimal Insurance Coverage and Moral Hazard[27]

To this point we have neglected that there may be alternatives to the insurance market for protecting your risky assets. A protection alternative might reduce either the probability of the unfavorable outcome or the magnitude of the loss if the unfavorable outcome occurs. (In fact, a protection alternative might do both simultaneously.) For example, placing steel bars on your windows (hopefully) reduces the probability of burgulary, but not the expected loss if one occurs. Purchasing a hand fire extinguisher for home use is not expected to alter the likelihood of a fire, but is expected to reduce its effects. Building a fire resistant structure reduces the likelihood of a fire and the size of the loss from a fire.

Conceptually related to the possibility of asset protection in a form other than the insurance market is the problem of "moral hazard." If individuals are rational expected utility maximizers, and if there exist techniques for altering the probabilities of events or their expected outcomes, the insurer must take these into consideration when determining his premium rates. Put differently, the insurer must be aware of the opportunities confronting the insured in order to create appropriate pecuniary incentives between insurance and alternative protection purchases.

Suppose we change our binary model such that an individual could reduce the proportion, ℓ, of his hitherto risky asset, H, which will be subject to loss by expending S: presumably, $\ell(S=0) = 1$, $0 \leq \ell \leq 1$, and $\ell'(S) \leq 0$. The expected utility, under these circumstances, will be

$$E(U) = \pi \cdot U(A+H(1+\ell(\lambda(1-\rho)-1))-S)+(1-\pi)\cdot U(A+H(1-\rho\lambda\ell)-S) \quad (27)$$

The individual may purchase insurance and/or loss reducing protection. (We are assuming for the moment that the probability of loss, π, is unchangable.) If second order conditions are satisfied, the marginal conditions for expected utility maximization are

$$\frac{\partial E(U)}{\partial \lambda} = \pi(1-\rho)U'(A+H(1+\ell(\lambda(1-\rho)-1))-S)+\rho(1-\pi)U'(A+H(1-\rho\lambda\ell)-S)=0 \quad (28)$$

$$\frac{\partial E(U)}{\partial S} = \pi(H(\lambda(1-\rho)-1)\ell'(S)-1)U'(A+H(1+\ell(\lambda(1-\rho)-1))-S) - (1-\pi)(H\rho\lambda\ell'(S)+1)U'(A+H(1-\rho\lambda\ell)-S) = 0 \quad (29)$$

The first marginal condition is similar for equation (15) above. That is, if we knew the optimal amount of resources S* to spend on loss reduction, the solution of (28) would yield the optimal λ^* which would balance the gains and losses of expected marginal utilities created by insurance coverage between the unfavorable and favorable events. Equation (29) can be interpreted similarly; if we knew λ^*, the solution of this equation would yield S*, which would balance the loss and gain of expected marginal utility from the last dollar spent on loss reduction activities. The simultaneous solution of these equations will maximize the expected utility of equation (27). Also, we can use the two marginal conditions to produce the general condition (30):

$$\rho = \frac{-1}{\ell'(S)H} \tag{30}$$

This implies that the value of loss reduction expenditures will, at the margin, be the same as the cost of insurance. For example, if H = \$1000 and the cost of insurance were 10¢ per dollar of coverage, then for S^*, $-\ell'(S^*) = .01$. In other words, the last dollar spent on insurance "protects" the individual as much as the last dollar spent on loss reduction.

In a similar fashion, we could assume that the probability, $\pi(S)$, is alterable through appropriate expenditures. We will not elaborate on this line of thought now. However, the relationship among insurance coverage, probability reducing expenditures, and "moral hazard" has been explored elsewhere by Ehrlich and Becker.[28] Their basic conclusion is that alternatives to the insurance market for protecting ones assets may be a prime cause of "moral hazard" faced by insurers.

For our purposes, we shall assume that the insurer, where necessary, takes into account the alternative responses of individuals. In fact, insurers at times issue insurance contracts contingent upon stipulated behavior by the insured. For example, automobile theft insurance is in force only if the vehicle was locked when the theft occured, or a precondition for fire insurance is an inspection of the premises, and so forth. Additionally, there are events that are basically beyond the control of the individual in terms of altering probabilities or losses. For instance, an individual living in the urban core is unlikely to affect either the likelihood of or his losses sustained from a riot.

III. Insurance Company Behavior[29]

In this section, we will focus upon the relationship between insurance issuance, profitability and claims' risk. Clearly, other aspects of an insurance company's operation (such as management costs, dividend policy, and investment of capital) interact in the determination of total risk and profitability. However, in order to simplify our discussion, we will take as given all variables other than those related to insurance policy issuance and claim risk structure.

Using the binary-theft-no-theft model, we will assume that there are n identical people in the world who own similar risky assets H and riskless assets A. Each person has the same subjective probability evaluation of theft, π. If the insurance company assesses the probabilities of theft as identically π and independent for each person, what is the optimal premium rate it should charge? The answer to this question is our first goal of this section.

a) The Insurance Company and Individual Policy Risks

The decision to insure a particular individual is part of the company's decision to insure a certain class of individuals. That is, the decision is one of a large number of similar, and independent decisions, and when major importance is attached to the average or over-all result of the decision, we have an insurable situation involving actuarial risk. Note the asymmetry between the rational behavior of the insurer and the insured. The insured reduces the expected value of his wealth when he takes out insurance, and may be behaving rationally when he does so, but the rational policy for the insurer, who has the advantage of pooling numerous risks, is to maximize his expected net wealth.

It is easy to show that in this model the insurance company in offering theft-protection coverage should maximize its expected wealth without regard to any other characteristics of the probability distribution of wealth (e.g., the standard deviation). In other words, the "risk factor" is irrelevant to the firm's decision, subject to the qualifications of sub-sections c) and d) below. The expected value of wealth for the insurance company will be (assuming no administrative costs):

$$E(W) = n\rho\lambda H - nE(C) \qquad E(C) = \pi\lambda H \tag{31}$$

where n is the number of clients, λ is the proportion of H insured by each, and ρ is the premium rate. The variance of the expected costs to revenue approaches zero (i.e., expected total costs approach a fixed proportion of total revenue, A) if we assume the outcomes for each policyholder are statistically independent, as n becomes large.[30]

$$E(A) = \frac{\sum_{i=1}^{n} E(C_i)}{\sum_{i=1}^{n} \rho\lambda H} \; ; \; \sigma_A^2 = \frac{\sum_{i=1}^{n} \sigma_{C_i}^2}{(\sum_{i=1}^{n} \rho\lambda H)^2} \tag{32}$$

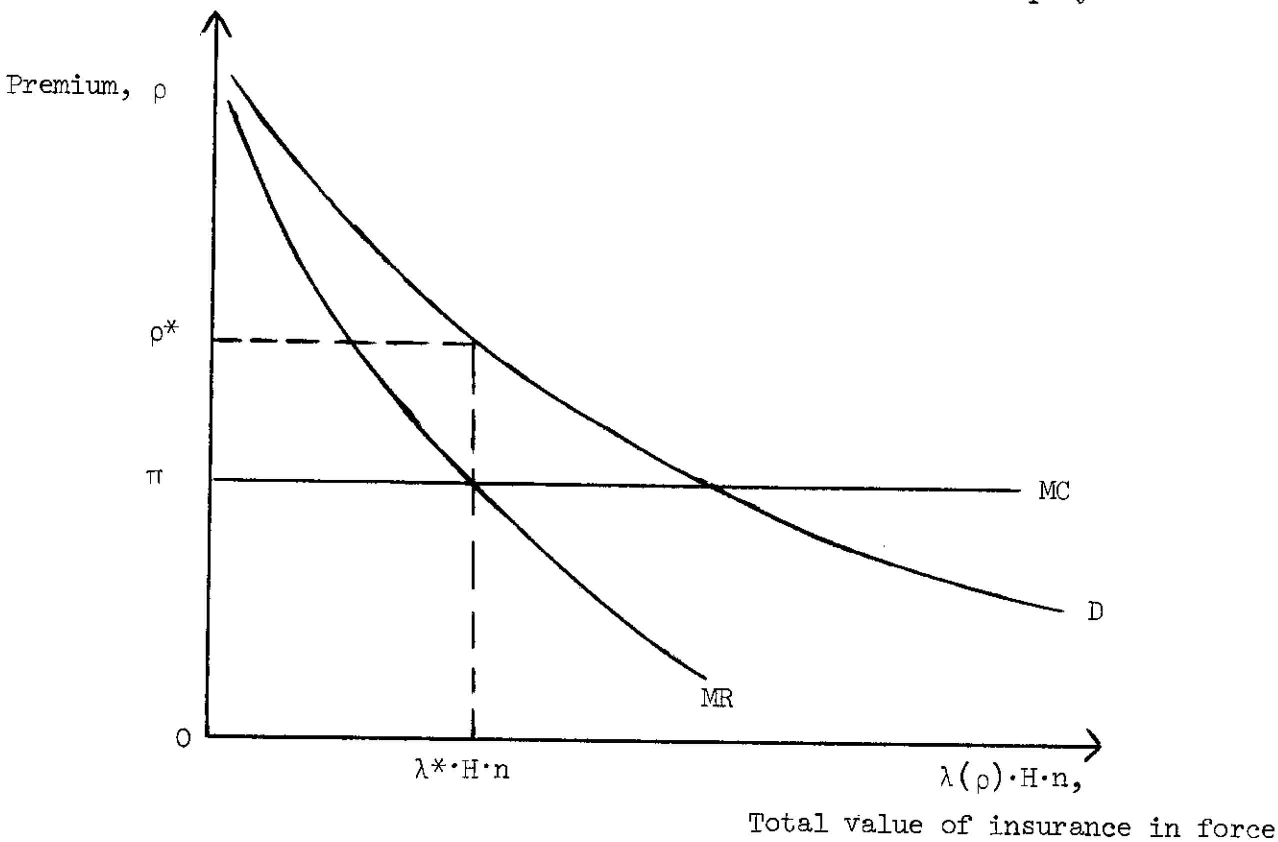
where $\sigma_{C_i}^2$ is the variance of the cost associated with the i^{th} individual. It can be seen that the variance of A is small if n is large by writing $\bar{\sigma}^2$ for the average variance, averaged over all individuals. Replacing the sum in the formula, recalling that $\rho\lambda H$ is the expected average revenue from each policyholder:

$$\sigma_A^2 = \frac{n\bar{\sigma}^2}{(n\rho\lambda H)^2} = \frac{\bar{\sigma}^2}{n(\rho\lambda H)^2} \tag{33}$$

This clearly decreases as n increases, provided $\bar{\sigma}$ does not grow in proportion to n. Therefore, an insurance company attempts to maximize wealth, disregarding any measure of risk.

Figure 6

Expected Wealth Maximization by the Insurance Company



MR = Marginal Revenue Function

MC = Marginal Cost Function

D = Demand or Average Revenue Function

b) Wealth Maximization[31]

The choice of the optimal premium can be determined by differentiating equation (31) with respect to ρ:

$$\frac{\partial E(W)}{\partial \rho} = n\lambda H + n\rho H \frac{\partial \lambda}{\partial \rho} - n\pi H \frac{\partial \lambda}{\partial \rho} = 0 \quad (34)$$

The first two terms on the right hand side of (34) are the expected marginal revenue, and it must be equal to the last term, the expected marginal cost. Graphically this is illustrated in figure 6. (Note that the total demand curve, faced by the company, is derived from the aggregation over all individuals of the information in figure 5.) ρ^* represents the wealth maximizing insurance premium rate and $\lambda^* H \cdot n$ is the insurance coverage provided. The total revenue accrued to the insurance company is $\rho^* \lambda^* \cdot H \cdot n$; the total expected payment on claims is $\pi \lambda^* H \cdot n$; and the expected accumulation of wealth for the time period is $(\rho^* - \pi)\lambda^* Hn$.

c) The Inclusion of Risk for the Insurance Company

Consider an insurance company, which faces a downward sloping demand curve for its insurance. That is, we will assume that the insurer has de facto some degree of monopoly power. The insurance company may confront a risk of illiquidity even when its expected wealth accumulation in each year is positive. Namely, the actual number of claims by policyholders in any year is unknown in advance, and may exceed the premiums paid by the current policyholders. One way for the company to protect against this eventuality is by maintaining a "liquid reserve" fund[32] above the annual premium payments. In such a case, the profitability of the insurance company will be reduced in favor of liquidity safety. That is, liquid reserves will supposedly earn lower yields than long-term illiquid investments. Of course, if it became necessary to convert quickly relatively illiquid assets into cash, the cost of conversion to the company could also be significant because of transaction expenses and potential asset price uncertainty. With this in mind, one could ask; 1) Is a "liquid reserve" fund consistent with profitability or wealth maximization? 2) How large a reserve fund is optimal to obviate illiquidity risk?

Again using our binary-model as a paradigm, we will examine these issues. First, it is appropriate to delineate explicitly several of the underlying statistical ASSUMPTIONS:

1) the occurrence of claims in different time intervals are stochastically independent;
2) the claims by individuals are mutually independent in the same time interval;
3) the size of each claim is exactly λH (i.e., it is nonstochastic);
4) the number of claims is dependent upon the length of the time interval studied (i.e., stationarity of parameters with respect to time increments);
5) the total number of policy holders is a constant, n.

Our binary model for the sum of the insureds' claims is in the limit a Poisson process for the probability of claims against the insurance company. Consider the n-insured individuals as an n-trial "Bernoulli experiment" with probabilities of π and $(1-\pi)$, such that n is large, π

is small and, therefore, $n\pi$, the expected number of claimants, is virtually constant for small changes in n. The probability that N claims will be made in any year (N=1,2,...i...n) is the Prob $(N=i) = \lim_{n\to\infty} \text{Prob}(N=i) = \binom{n}{i}\pi^{i}(1-\pi)^{n-i} = \frac{n}{n}\,\frac{n-1}{n}\cdots\frac{n-i+1}{n}\,\frac{(n\pi)^{i}}{i!}\,(1-\pi)^{-i}(1-\pi)^{n}$. In the last sequence of terms, the first i terms approach unity, the following term is a constant, the next term is approximately unity for π very small, and the last term is $e^{-n\pi}$. Therefore, Prob $(N=i) = e^{-n\pi}\cdot\frac{(n\pi)^{i}}{i!}$, which is a Poisson probability density function.

The assumed insurance company will have a claim payment distribution function, which will be the Poisson cumulative distribution multiplied by λH, the fixed size of each claim. Define this function, the claims function:

$$F(C) = \sum_{i} \lambda H e^{-n\pi}\,\frac{(n\pi)^{i}}{i!} \quad \text{for } i = 1,2,\ldots n \tag{35}$$

The expected cost of claims is

$$E(F(C)) = n\pi\lambda H$$

A measure of dispersion for the claims cost is the standard deviation, or the second moment around the mean of F(C):

$$\sigma = \lambda H\sqrt{n\pi}$$

Applying the Central Limit Theorem[33] to the Poisson analysis, it is clear that for the presumed large number of policyholders n, $F(C) \sim \varphi\left(\frac{C-n\pi\lambda H}{\lambda H\sqrt{n\pi}}\right)$ where φ represents the Normal distribution function with zero mean and unit variance.

Unless a company's total premium income is greater or equal to the total possible claims, there is always some risk of being unable to meet claimants' demands. Therefore, the insurance company might set some acceptable and arbitrarily small probability of not being able to meet its claims from both liquid assets and current premium payments. Call this "acceptable" probability level α. Premium income is $\rho n\lambda H$, and is assumed to be paid to the company at the beginning of each period, and is, therefore, nonstochastic. In other words, the company's subjective utility function contains some subjective weight for "survival" or avoidance of "ruin." Clearly, more wealth is preferred to less, but not until the safety of liquidity has been guaranteed, up to an arbitrarily small probability α.[34]

The company, therefore, wishes to find the minimum $K \geq 0$, the size of its "liquid reserve" fund, such that

1) Prob $(C > K+\rho n\lambda H)=1-\text{Prob}(C\leq K+\rho n\lambda H)\leq \alpha$; and
2) in order to see if this is consistent with expected wealth maximization, set $\rho = \rho^* > \pi$ from (34) above (with n given).

$$\varphi(Z) \approx \varphi\left(\frac{C-n\pi\lambda H}{\lambda H\sqrt{n\pi}}\right) = 1 - \alpha$$

where Z is the standardized normal distribution table value for

$$Z = \frac{C-n\pi\lambda H}{\lambda H\sqrt{n\pi}}$$

Therefore,

$$C = Z\cdot\lambda H\sqrt{n\pi} + n\pi\lambda H$$

or

$$K = (\pi-\rho)n\pi\lambda H + Z\lambda H\sqrt{n\pi} \qquad (36)$$

The critical value for $K = 0$ occurs when the number of policyholders is n^*

$$n^* = \frac{Z^2}{(\pi-\rho)^2\pi} \qquad (37)$$

Hence, if the actual number of policyholders is $n \geq n^*$, there is no reason to keep a "liquid reserve" fund. In other words, expected wealth maximization with no liquid reserve fund is consistent with liquidity safety for arbitrary α, so long as our assumptions above are satisfied, and n is sufficiently large. In fact, it is possible if n is large enough to set $K < 0$, which implies one can safely invest current premium income without significant illiquidity risks. This really is not surprising if one considers that $\rho = \rho^*$ is actuarially favorable for the company. In insurance nomenclature, ρ^* contains a loading factor equal to $(\rho^* - \pi)$ because π is the expected marginal cost for the last dollar of purchased coverage. From (37) it is clear n^* must increase as the loading factor decreases to maintain $K = 0$. This means that smaller loadings can be compensated for by increments in the size of the pool of policyholders.

The price elasticity of demand for insurance, e, determines the magnitude of the loading in the wealth maximizing model:

$$e = \frac{-\partial\lambda}{\partial\rho} \cdot \frac{\rho}{\lambda} \text{ with } n, \pi, \text{ and } H \text{ as parameters.}$$

Then, by rearranging (34),

$$\frac{\pi}{\rho} = (1 - \frac{1}{e}).$$

Clearly, $\frac{\pi}{\rho}$ will be smaller for small values of e(i.e., $e > 1$). That is, the less elastic the demand curve, the greater will be the loading. We can also observe that the more "competitive" the insurance industry is the greater the importance for a company to expand its market share in order to avoid liquidity risks, and vice versa.

In sum, illiquidity risks are reduced when the insurance company has a large pool of statistically independent policyholders and the demand for insurance is relatively inelastic. In addition, under these very favorable conditions a company may actually keep zero "liquid reserves" above the annual premium payments.

d) Some Qualifying Comments

If the occurrence of thefts during any time period were not independent, as assumed above, it is rational for insurance companies to take into account the simultaneity risk associated with being the insurer of several individuals. An example of this is the inter-related higher probabilities of theft, fire, and so forth, associated with urban-core residence. Therefore, insurance companies might prefer to diversify by having theft insurance policies in many neighborhoods. In that way, it is intended to reduce the degree of statistical dependence among clients. If, in fact, insurance companies cannot find independent risks in large quantities, then we would expect that the optimal premium rate would have to be determined taking into account risk. For instance, a "substandard" risk loading factor might be added to policies written in high crime districts.

The decision would be a portfolio decision similar to a gambler: issuing insurance by the company is equivalent to converting one's certain assets into risky assets. If premium rates are actuarially unfair, the company increases its expected wealth by increasing its risk. The determination of the optimal ρ,[35] would entail some knowledge of the insurance companies' aversion to risk. This is comparable to the portfolio decision of individuals, with the insurance company attempting to maximize its "utility" function subject to its income and capital constraints.

In addition, it can be shown that insurance companies would take all similar risks without adjusting premium rates, gray penciling, red penciling, and so forth if sufficient neighborhood diversification were possible. This is true even if there exists stochastic dependence between insured individuals living in close proximity. More formally, consider (integers) $i \epsilon N > 0$ identical individuals, facing π_i, the probability of theft of his asset H. Moreover, π_i and π_j for i, j ϵ N are stochastically dependent; the dependence vanishes as $|i-j| = \epsilon$ becomes sufficiently large. That is, as the individuals become sufficiently distant neighbors, the fate of i is independent of j.

Define a random variable y_i such that

$$\begin{Bmatrix} y_i = 1, \text{ if H is stolen from the } i^{th} \text{ individual} \\ y_i = 0, \text{ otherwise} \end{Bmatrix}$$

Then, $E(y_i) = \pi = E(y_i^2)$. Examine

$$|E(y_i y_j) - E(y_i)E(y_j)| = |E(y_i y_j) - \pi^2| \leq D(i-j) \qquad (38)$$

where $D(i-j) \leq 1$, the function which denotes the upper bound for statistical dependence for each $\epsilon = |i-j|$. Also, $D(i-j)$ is assumed to vanish for sufficiently large ϵ.

Consider $E\left(\frac{H \sum_{i=1}^{N} y_i}{N}\right) = \pi H$. That is, the average and marginal expected loss from theft for all individuals is constant. In addition, the variance of the average (and marginal) loss is

$Var(\pi H) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} [E(y_i y_j) - \pi^2]$ and from (38), we see that

$$Var(\pi H) \leq \frac{1}{N^2} \sum \sum D(i-j) \qquad (39)$$

It is clear that the variance in (39) will approach zero if a sufficiently large N individuals are insured.

Put differently, the insurance companies' use of red-penciling, and so forth, is in part the result of being unable to pool a sufficient number of insured individuals from diversified locations. If sufficient diversification is possible, the expected costs from each insured client will be $\pi H < \rho H$. Therefore, wealth maximization will be possible, even when stochastic dependence between individuals exists, so long as appropriate locational diversification opportunities exist. Also an individual insurance company may not be able to locationally diversify; but either groups of companies or the government operating from a national purview should be able to achieve the necessary risk spreading.

The other assumptions of our simple model are also subject to significant qualifications. Clearly, the number of policyholders is to some extent controlled by, among other things, the company's pricing policies. Also, we would expect that the world is not immutable, and over time, parameters do change. For example, if crime is on the secular rise, it is not unreasonable for insurance companies to incorporate changes of π into their policy decisions on ρ. Usually, companies utilize their experience from the most recent past to determine new premia structures. Finally, our assumptions that claims are of a fixed size (i.e., non-variable) is unrealistic.[36] However, all of these objections can be handled; the mathematics will become more complex, but the essence of our analysis is not impaired.

The Application of the Theory of Insurance to Urban Problems

I. A Statement of the Problems in Urban Core Insurance[37]

a) The Insured's View

The apparent inadequacy of comprehensive property insurance for fire and extended coverage, vandalism, burglary and theft in our urban core areas has become an issue of public concern. The insurance market is not meeting the needs of individuals for asset security; the demand for insurance outstrips the availability of coverage and/or coverage is available at what is considered excessive rates. Moreover, this problem has ramifications beyond the insurance market. Insurance coverage is crucial for the normal functioning and growth of the economy. Without the security of insurance, banks may not readily make loans to businesses, mortgages may not be obtained easily for residential, mercantile, or commercial property, the revitalization and renewal of urban areas cannot progress, and so forth.

The general restrictive availability and increased premium rates for insurance have been exacerbated by the occurrence of urban riots during the latter half of the 1960's and the fear of future riots. In a recent survey[38] of contiguous-Watts-Los Angeles (1968), twenty per cent of the responding businesses did not have fire and extended coverage insurance policies; and 44 per cent of these businesses did not have this insurance because it was unavailable. Of those who had fire and extended coverage, more than half of the respondents had to pay increased premium rates between 1965 to 1968. Moreover, half of these businesses that paid increased rates, paid 50 per cent or larger increments.

b) The Insurer's View

The reluctant and restrictive policies of the insurance companies have several interdependent dimensions. First, recent adverse experience with urban riots has intensified the insurance companies' belief that core insurance is a potentially catastrophic liability.

Table 1[39]:

Estimated Insured Civil Disorder Losses, 1965-1969

Year		Loss (in millions of $)
1965		38
1966		79
1967		67
1968		1
1969		31
	Total	216

Table 1 shows that the costs of riots to insurance have not been trivial, and are a legitimate basis for insurance companies' apprehension.

Underwriting, the process by which the company determines to accept or reject an insurance application, is intended to filter out undesired risks in order to create a profitable and safe portfolio of insurance contracts for the company. Insurance policies for urban core areas are generally considered to be inferior risks because they are frequently in blighted areas, high crime districts, poorly maintained or older neighborhoods. An underwriter's compensation ultimately depends upon the profitability of the business he accepts. Therefore, the higher risk application from the urban core is more likely to be declined. This underwriting effect is reinforced by the current marketing techniques utilized by insurance companies. There appear to be relatively few insurance agents, virtually none of whom are black[40], who deal in urban core business. These factors systematically militate against insuring in the core areas.

Insurance companies typically reinsure policies in order to protect themselves from excessive losses. The supply of capital for reinsurance has traditionally and essentially been uncontrolled by the public sector and sensitive to rate of return it can earn. Urban core insurance has become less desired by reinsurers, and thereby further discouraging primary insurers from issuing such policies. That is, without adequate reinsurance, primary insurers must behave more conservatively to protect their own capital. This clearly is a major economic hurdle; one which may be solved only with significant intervention by the public sector.

All of these factors interact and contribute to the insurance companies' overall problem of finding adequate profitability and reduced risk of excessive loss or insolvency. The fire and extended coverage, vandalism, theft and burglary lines of insurance are significantly less profitable than other types of insurance; add the uncertain, but potentially large losses because of riots and it is clear that insurance companies will not actively expand urban core business.

II. Elements of the Optimal Insurance Market Solution

a) The Welfare Maximizing Quantity of Insurance and the Private Insurer

With the above practical considerations understood, there is much to be learned about the welfare implications of insurance from simplified economic models. Earlier we argued the private insurer, under special conditions pertaining to the number of policyholders and the distribution of claims, will maximize expected wealth (i.e., in the short run this implies the maximization of net-revenue). We shall now generalize several important elements of that analysis in order to study the welfare effects of insurance in the urban core.

We assume that the insurer has some degree of monopoly power, and therefore a downward sloping demand curve. Also, for simplicity, we assume that the pitfalls of "moral hazard" are not significant for the insurer. Consider an insurer of similar types or classes of risk in the urban core insurance market; assume each policyholder wishes to insure (at least a proportion of) a similar valued piece of property, H. For example, the insured items might be similar dry-cleaning stores or similar groceries. The expected total value of the insurance in force in each time period is I, and is related to λ, the proportion of H that is insured by each policyholder, and n, the expected number of clients. In general, $I = I\ (n,\lambda)$.[41]

The expected value or quantity of insurance sold will vary according to the premium rate, ρ. Previously, it was demonstrated that as ρ decreases, the proportion of H insured increases. Also, we will expect that as ρ decreases, the number of policyholders will increase, as either new people find insuring optimal or people switch companies. This is embraced in the demand function

$$\rho = \rho(\lambda,\ n,\bar{\bar{H}}) = \rho(I)$$

with $\frac{d\rho}{dI} = \rho'(I) < 0$, or the demand function is assumed to be normal-downward sloping.

The expected cost function for the insurer is positively related to the value of insurance in force in each period. However, as discussed above, insurance companies accept (at a given premium rate) the low risk superior-clients before the high risk inferior clients, and statistical dependence increases as you increase your number of clients or saturate the market from, say, one area (e.g., the urban core clientele are statistically dependent risks because of such events as riots). We will presume that the expected cost curve will be upward sloping in the relevant range:

$$C = C(\lambda, n, \bar{H}) \tag{40}$$

with $C'(I) > 0$.

The wealth maximization quantity of insurance and premium rate are I^* and ρ^*, respectively, and satisfy the marginal condition[42] (41):

$$\frac{dE(W)}{dI} = 0 = \rho + \rho'(I) \cdot I - C'(I)^{43} \tag{41}$$

That is, the expected marginal revenue will equal the expected marginal cost in equilibrium.

The Pareto optimal quantity of insurance I^p and premium rate ρ^p can be found as the solution of

$$\rho^p = C'(I^p) \tag{42}$$

The difference between the Pareto and Private Insurance Solution can be seen in Figure 7 as A and B, respectively.

The government might attempt to correct the apparent deficiency in the quantity of insurance by a sales subsidy of g to the company. The magnitude of g can be discovered by assuming wealth maximization by the insurer:

$$\rho^p + I^p \cdot \frac{d\rho^p}{dI} + g - C'(I^p) = 0$$

Upon rearrangement and substitution this will yield

$$g = \rho^p - \text{marginal revenue (at } I^p\text{)}$$

In other words, the subsidy shifts the marginal revenue curve, as seen by the insurer, such that it passes through the point of intersection of the marginal cost curve and the demand function. From a social point of view, the marginal policy-holder is paying exactly what he costs (i.e., his loading is zero).

Of course, the subsidy increases the total wealth of the insurance company. This can be found by analyzing the components in its wealth change from before and after the subsidy, ΔW:

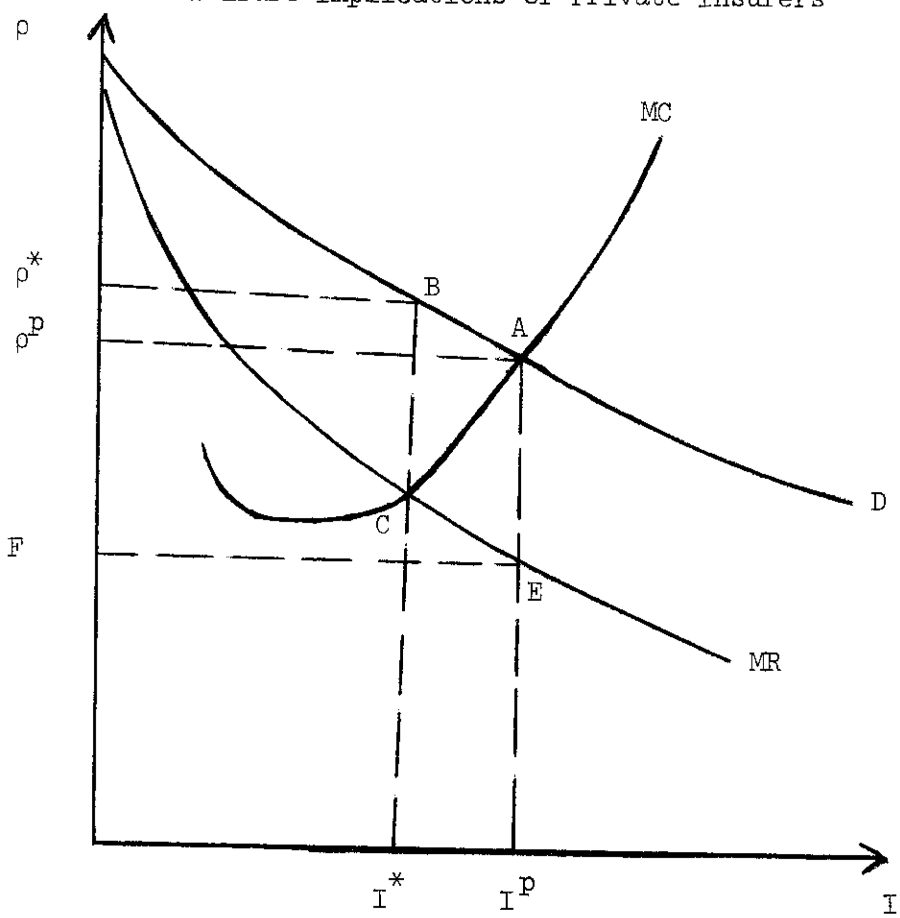
$$\Delta W = W^p - W* = \int_{I*}^{I^p} (\rho + \rho' \cdot I - C')\, dI + g \cdot I^p$$

The subsidy $g \cdot I^p$ is represented in Figure 7 as ρ^p AEF, the integral is negative and represented by ACE. Clearly $\Delta W > 0$ (i.e., area of ρ^p ACEF). A general lump-sum tax of ΔW on the insurance company's wealth, combined with the aforementioned subsidy g, will not alter private profitability. The total quantity of insurance will be increased to the Pareto optimal point by the subsidy taxation program, but there is a net cost of ACE to society. This may be financed by taxing those who benefitted from the new insurance program.

If we assume that the marginal utility of wealth for each individual is constant over the relevant range of wealth,[44] then the areas under the demand curve for changes in insurance measure the exact number of dollars necessary for retaining the original level of utility. (This

Figure 7

The Welfare Implications of Private Insurers



might be called in a Hicksian spirit the wealth-compensated demand curve.)

$$\Delta C = \int_{I^*}^{I^p} (\rho(I) - \rho(I) - \rho'(I)\cdot I)dI = \int_{I^*}^{I^p} -\rho'(I)\cdot IdI$$

ΔC is the total wealth the consumers could give up, after the Pareto optimal insurance point is achieved, and be indifferent to the original insurance arrangement. That is, the consumers at the very most could lose everything under the demand function in Figure 7 between I^p and I^*, but through insurance coverage only pay the area under the marginal revenue curve in that interval. The net gain is BCEA, which is necessarily larger than ACE. Therefore, a consumer taxation scheme is possible to reimburse the insurance company ACE (the net cost of the subsidy-general tax), and still have the consumers of insurance better-off.

Tentatively, we may conclude that the government could increase the quantity of insurance available in markets with a current deficiency (in a Pareto sense) by subsidizing such insurance sales, taxing insurance companies more heavily in general, and employing general income or wealth taxes on consumers (to finance any deficits arising in this operation.

b) The Welfare Implications of the Suburban-Urban Core Insurance Mix

The welfare implications of insurance can not be fully understood for the urban core without some discussion of its relationship to the suburban areas. Theoretically, the full range of risk discrimination is necessary for insurance to attain its maximum social benefit. That is, properties that are inferior risks should pay higher premiums.

Urban core property, <u>ceteris paribus</u>, represents an inferior insurance risk, and should pay a higher premium rate from an efficiency point of view. This may be felt to be inequitable, even if efficient.

Using the theft-no-theft binary model, H would have a greater probability of theft in the intrinsically more risky urban core than if owned by a suburbanite. That is $\pi_c > \pi_s$, where the subscripts c and s represent the urban core and suburbs, respectively. The expected losses would be $\pi_c H > \pi_s H$. If there are N people living in c with a piece of property H, and M people living in s with a piece of property H, the total expected loss would be $N\pi_n H + M\pi_s H$. Combining both markets, and forming an actuarially fair premium π would yield:

$$\pi = \frac{N_{\pi_c} + M_{\pi_s}}{N+M}$$

c) Equity, Efficiency and the Uniform Premium Rate

The suburbs under this scheme would be subsidizing the urban core insurance business. That is, the uniform premium rate, $\pi = \pi_s + \ell$ for the suburbs, where ℓ is an implicit positive loading above the acturial risk, yields an expected utility maximization of s, assuming all individuals have the same utility functions, initial wealth endowments and so forth, of:

$$E(U) = \pi_s\ U(A+\lambda H(1-\pi))+(1-\pi_s)U(A+H(1-\lambda\pi))$$

where λ is the optimal proportion of H insured (assuming $0 < \lambda < 1$). Therefore, the total premium is $\pi\lambda H = \lambda H(\pi_s+\ell)$.

If we assume that the suburban expenditures exactly subsidize the urban core business, it is possible to show that a lump sum tax equal to $\lambda H \ell$, and actuarial insurance pricing for the suburbs will increase welfare. That is, the urban core will be subsidized exactly as under the uniform premium plan, but the financing will be done by a lump-sum tax rather than an increased premium rate.

In the suburbs, with the lump sum tax $\ell\lambda H$:

$$E(U) = U(A+H(1-\pi_s) - \ell\lambda H).$$

(Note that an actuarial fair premium, π_s, will induce full-insurance coverage.) Also under the new plan the expected wealth is identical to the original plan:

$$A+H(1-\pi_s)-\ell\lambda H \equiv \pi_s(A+\lambda H(1-\pi))+(1-\pi_s)\ (A+H(1-\lambda\pi))$$

However, our new plan represents <u>certainty</u> of wealth by virtue of full-insurance coverage, and is in utility (and social welfare) terms preferred. This result is true in general for variable loss-insurance arrangements: that is, if equity dictates that a sub-market of insurance be subsidized, it is welfare-superior to do this by financing the subsidy through a lump-sum tax, and maintaining actuarially fair premium rates elsewhere.[45] Additionally, under the lump-sum plan the quantity of insurance in the suburban market is expanded relative to the core market. Also, this aspect of the insurance problem can be solved only by a governmental authority who can tax and subsidize in the entire SMSA. That is, "fiscal federalism" appears to be the only rational way to achieve our welfare maximization-equity solution.

III. Private Market Approaches to the Solution

a) Institutional Elements of the Private Insurers' Solutions

The cost of an insurance policy supposedly reflects the nature of the risk. Theoretically, the insurance underwriter classifies the risk, and thereby establishes the appropriate premium rate. As a practical consideration, many underwriters must use only the standard rate (i.e., a rate established by state regulations for certain classes of risks), and therefore accept or reject applications according to their perceived riskiness. In the case of urban core property insurance, the standard rate classification may apply to an entire city or a county. That is, the standard rate will be relevant to a wide spectrum of risks. Therefore, underwriters, with a limited capacity to accept risks, will insure the least risky properties, or in general those outside of the urban core.

Insurance companies have introduced rate and other types of flexibility into their insuring processes in order to increase the magnitude of their market and increase their compensation for accepting greater risks. Some of the techniques employed by insurers in order to create and/or clear insurance markets are:

1) Substandard risk ratings, by which the company increases the premium by a surcharge, according to the precise nature of the substandard risk;
2) Limitation of the term of insurance to one year, with no guarantee of renewal;
3) Restricted perils, such as no glass coverage in riot prone areas;
4) Reduced commissions to agents, thereby increasing the effective premium to the company; and
5) Schedule Excess Rates.[46]

b) The Excess Rates Case: An Example

Take the last of these as an example: the schedule excess rating plans permit the underwriter to charge higher rates based upon hazards or conditions not taken into account by the standard rates. Such conditions include block congestion, high valued buildings without adequate fire protection services, and buildings with large open areas susceptible to the spread of fire and riot exposure. Responsible owners of well-maintained properties may be penalized because of the inherent riskiness of his neighborhood. (This phenomenon is called Environmental Hazards.)[47] The underwriter in such cases places primary importance upon the location of the risk and will refuse to insure the property, even though the risk is beyond the control of the owner. The net result may be a disincentive to improve or perhaps maintain one's property in blighted areas. If we wish to revitalize our urban core areas, the operation of private insurers, using excess rate plans, may from a social point of view result in an inadequate supply of insurance for the urban core.

c) Excess Rate Schedules as a Sub-optimal Allocation of Insurance

The Excess Rate Schedules represent a real-world alternative to our elaborate subsidy-taxation plan discussed earlier. The Excess Rating affects individuals who might be thought of as occupying the upper parts of the expected Marginal Cost Curve in Figure 7. A wealth maximizer with some monopoly power charges more than expected marginal cost. Therefore, excess rate schedules would charge profit-loadings to individuals on the upper part of the marginal cost curve. (Also, no insurance will be sold to customers at (ρ^p, I^p) because no loading can be extracted from them.) This will, in effect, cause a reduction in the Pareto optimal level of insurance, and is clearly undesirable.

It is clear that more readily available reinsurance for urban core insurers may effectively reduce the expected costs for each insurer, and thereby encourage an increase in the supply of insurance. Also, it is possible to increase the supply of insurance by reducing the "effective" cost function through alternative mechanisms: instead of reinsurance, the pooling of all private insurance companies' portfolios reduces individual company risk; or "catastrophe" reserves can be accumulated more rapidly by reduced taxation rates on profits derived from riskier urban core insurance business.

For the moment, if it is assumed that the private property-liability insurance industry can ultimately bear the major component of, say, urban riot losses, and we wish at the individual firm level to relieve the potential of catastrophic losses, reinsurance may be a potential solution. Stated differently, if the insurance industry can handle "riot" insurance problems through the appropriate spreading of the expected costs of risks, a private market solution is feasible. However, as observed earlier, there has been an increasing reluctance among private reinsurers to handle urban core business. Therefore, while the private insurance industry may have the financial capability for undertaking significant levels of urban core insurance, a governmental reinsurance program may be a necessary vehicle for actual market operation.

Basically, the federal riot reinsurance program, a derivative of the Urban Protection and Reinsurance Act,[49] is an attempt to provide the reinsurance needed by the private market to do insurance business in the urban core. The program's principle strengths[50] are (1) it provides assistance to individual firms against catastrophic losses and (2) in the case of macro-catastrophic losses such as major riots it can provide either state or federal aid to the industry as a whole. Simultaneously, the federal reinsurance program has several intrinsic weaknesses: (1) the program is organized on a state-by-state basis - Clearly, reinsurance calls for risk-spreading at the widest possible level (i.e., at the national level); and (2) it still may be reasonable to assume that private property-liability insurance in urban areas, even with the availability of reinsurance, which the private insurer ultimately pays for through reinsurance premiums, may not be as profitable as other lines of and/or other localities for insurance. Therefore, the private supply of insurance may not expand rapidly with the introduction of some degree of premium pricing flexibility such as excess rate schedules and guaranteed reinsurance.

In conclusion, the private solution to the urban core insurance situation, even if supplemented by reinsurance, will be sub-optimal in terms of the "idealized" solution. That is, without more complete and multi-faceted governmental intervention, the Pareto optimal insurance solution (i.e., Point A in Figure 7) will not be achieved.

IV. Existing Public Programs as Solutions

From the two preceeding sections of this essay, it is clear that the government must ultimately play a crucial role if urban insurance problems are to be solved equitably and efficiently. Towards this end, it would be interesting to examine the effects of current government regulations and programs on the equilibrium of the insurance industry.

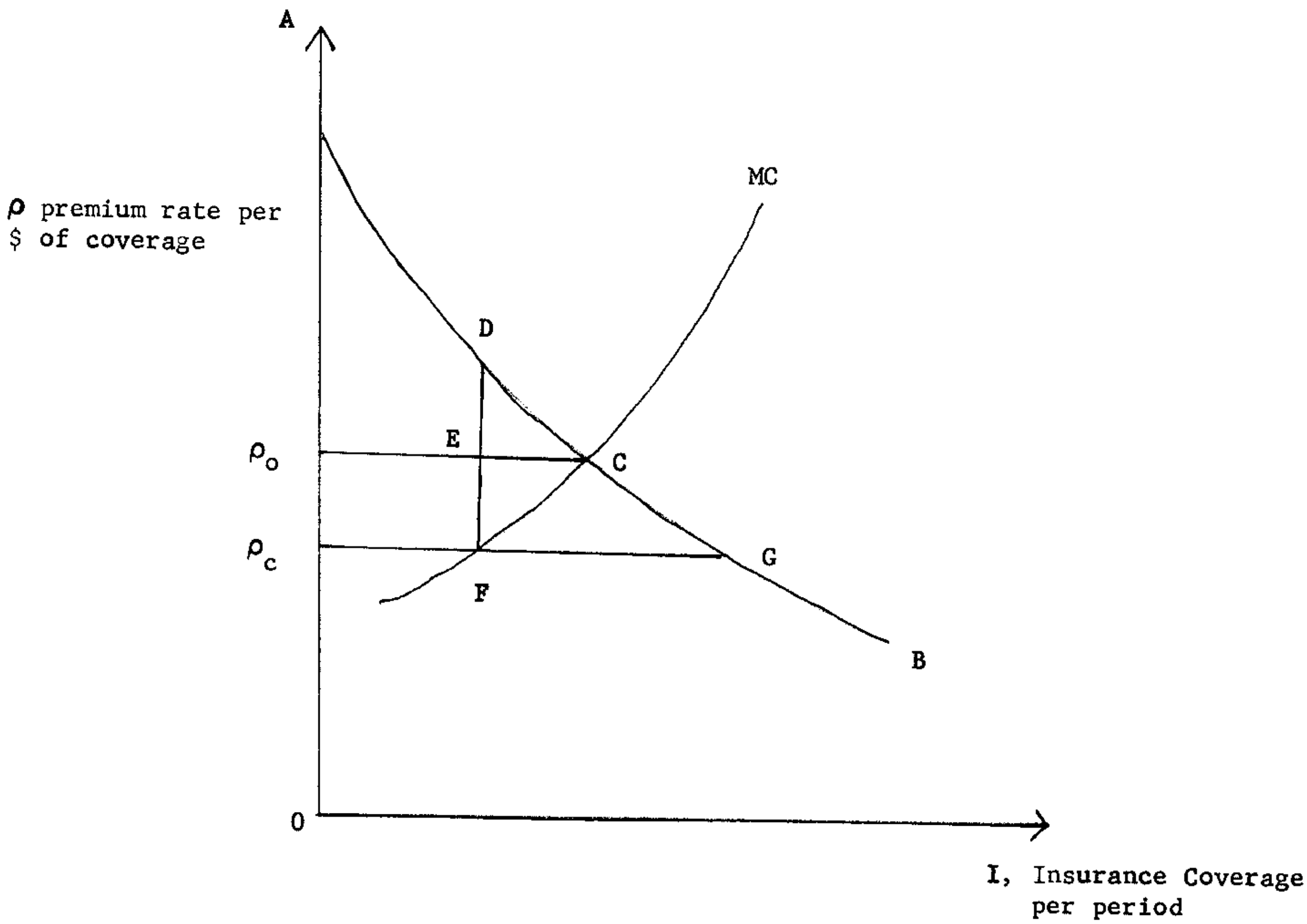
a) Insurance Rate Regulation and the Suburban-Urban Core Insurance Mix under Perfect Competition

A framework will now be developed for determining the effects of governmental insurance rate regulation on the equilibrium of the property insurance market and on the urban-suburban insurance mix. If the urban and suburban insurance markets were perfectly competitive and there were no externalities, an "effective" rate ceiling clearly would disrupt the normal operation of the market. Moreover, if we assume the urban core property insurance market has, <u>ceteris paribus</u>, a relatively and sufficiently inelastic demand function, the effective rate ceiling, appraised in traditional consumers' surplus terms, would reduce the well-being of the urban community more than a market with a relatively elastic demand for insurance.

This measurement of the welfare-change induced by a rate ceiling in a competitive market is illustrated in Figure 8. If the market were unconstrained by regulatory action, the equilibrium price per dollar of insurance coverage would be ρ_o with an implied consumers' surplus of $AC\rho_o$. The introduction of an "effective" ceiling ρ_c will reduce the quantity of insurance coverage sold in each period as well as engendering an excess demand (i.e., a market disequilibrium) of FG. The consumers' surplus which results from the rate ceiling ρ_c will equal the area of the quadralateral $ADF\rho_c$. This may be larger, equal to, or smaller than the unregulated consumers' surplus depending upon if $\rho_o\rho_c FE \gtreqless DEC$. (The

Figure 8

Rate Regulation in a Competitive Market



AB is the demand function
MC is the marginal cost or industry supply curve
ρ_c is the maximum premium rate ceiling under regulation

ρ_o is the market equilibrating premium rate producers' surplus will be disregarded in our analysis.) Therefore, a priori one can not state the welfare effects of a rate ceiling. However, the more inelastic the demand curve (and the more elastic the marginal cost curve) the larger DEC will be relative to $\rho_o\rho_c$FE. If AB, the demand for insurance, is sufficiently inelastic, a net loss of well-being will be experienced because of the rate ceiling. In other words, under our assumptions of perfect competition and no externalities, it is clear that denying a group with a sufficiently inelastic demand for insurance (i.e., a need for insurance as measured by willingness to pay) free access to a market will necessarily reduce the groups' welfare.

This elementary analysis breaks down if one considers that the property insurance market, while possessing competitive elements, is not as assumed above strictly competitive. For example, different companies offer subtly differentiated insurance coverage, such as better or faster claim payment procedures, or better "local" sales agents, and so forth. Also, the perfect competition assumption of free entry, in general, will be untrue because of licensing regulation and, in general, significant capital start-up costs. The analysis will, therefore, consider rate regulation effects in the context of "imperfect" competition.

More precisely, a monopolistic competition[51] model will be assumed with each insurance company adjusting its operation in two separable markets, the urban core and the suburban area. The objective of each firm in line with our earlier presentation will be to maximize expected wealth, appropriately adjusted for risk, by choosing pricing and marketing policies that do not violate legal constraints, such as rate ceilings. In general, we will assume that the insurer perceives, *ceteris paribus*, the urban core market as riskier than the suburban insurance market. We will proceed by analyzing the behavior of a "typical" property insurance company, tacitly assuming that this represents the actions we would expect the total market to be undertaking. Also, since in a monopolistic competition model the insurance companies are interdependent, particularly from the demand side of the market, we would expect, and henceforth will assume, that there will be market stability once equilibrium is determined.

b) Monopolistic Competition without Rate Regulation

First, we will consider the equilibrium solution if there were no rate regulation. The demand functions for insurance by each area for each insurer, *ceteris paribus*, are assumed to be downward sloping, independent between market areas, and interdependent with firms' behavior. (We will subsume that the behavioral interdependence among firms is known and therefore incorporated into the "effective" demand function each firm uses when optimizing.) The functions denoting the expected costs of supplying insurance coverage for each firm for each area are not assumed to be independent between market areas or among firms.[52] (Again, we will subsume that the behavioral interdependence among firms is known and therefore incorporated into the cost curves.) The interdependence between insurance market areas arises because of the possibility of neighborhood diversification by the insurer by appropriately choosing among policyholders, or in other words insurance company portfolio risk spreading among potential claimants. We have argued at length in subsections above that an insurer may reduce his expected costs by diversifying his client portfolio over several neighborhoods in such a way as to delimit the stochastic dependence among policyholders. Put differently, if the insurer can find sufficient neighborhood diversification, under the assumption that covariation of losses declines as the distance between policyholders increases, the adjustment for risk due to interdependence among policyholders in the expected cost function for an area may be reduced.

In addition, the expected cost function for each area is assumed to be a non-decreasing function of insurance coverage in that area, given the other area's quantity of insurance. The expected marginal cost function for each area's insurance, given the other area's coverage (in the ranges relevant to our analysis), is assumed to be upward sloping because the insurer will (1) expose himself to increased risk through expansion of operation without further neighborhood diversification, (2) expose himself to marginally inferior risks within the area by expanding business and (3) increase his marketing costs disproportionately through area sales "saturation." Also, it is assumed that the urban core property insurance market is more risky business than the suburban area business because of the possibility of claims from collective catastrophic losses by the urban policyholders due to, say, riots, as well as higher "normal" non-catastrophic losses from, say, fires.

Let superscripts c and s symbolize the urban core and suburban areas, respectively. Mathematically, the demand functions for each area are

$$\rho^c = \rho^c(I^c)$$

$$\rho^s = \rho^s(I^s)$$

The expected cost function for each area will be

$$F^c = F(I^c, I^s)$$

$$F^s = F^s(I^c, I^s)$$

In the light of our assumptions and discussion above, where subscripts represent the partial derivatives,

$$\rho^c_{I^c} < 0 \text{ and } \rho^s_{I^s} < 0$$

$$F^c_{I^c} > 0 \text{ and } F^c_{I^s} < 0 \quad .$$

$$F^s_{I^s} > 0 \text{ and } F^s_{I^c} < 0$$

This information is also illustrated graphically in Figure 9 for the typical insurer.

In our one-period model, the expected wealth maximizing firm will optimize by maximizing the difference between all revenues and all expected costs (risk adjusted in both markets). That is, the firm's objective function will be

$$\pi = \rho^c \cdot I^c - F^c + \rho^s \cdot I^s - F^s \qquad (43)$$

The first order or marginal conditions for optimization are

$$\frac{d\rho^c}{dI^c} \cdot I^c + \rho^c = F^c_{I^c} + F^s_{I^c} \qquad (44)$$

$$\frac{d\rho^s}{dI^s} \cdot I^s + \rho^s = F^c_{I^s} + F^s_{I^s}$$

The marginal conditions (44), if satisfied and assuming second order conditions are met, will yield the optimizing values for price and insurance coverage in each market, which will maximize the objective function (43). The left-hand side of each equation of (44) represents the marginal revenue and the right-hand side represents the expected marginal cost for each market area. In equilibrium, the marginal revenue will equal the expected marginal cost in each market. Figure 9 depicts such an equilibrium for ρ^{c*}, I^{c*} and ρ^{s*}, I^{s*}.

The second order conditions for maximization are

$$\frac{d^2\rho^c}{dI^{c2}} \cdot I^c + 2\,\frac{d\rho^c}{dI^c} - F^c_{I^cI^c} - F^s_{I^cI^c} = J < 0$$

$$\frac{d^2\rho^s}{dI^{s2}} \cdot I^s + 2\,\frac{d\rho^s}{dI^s} - F^c_{I^sI^s} - F^s_{I^sI^s} = K < 0$$

and

$$J \cdot K - (F^c_{I^cI^s} + F^s_{I^cI^s})^2 > 0$$

The first two second order conditions will be met under our assumption of downward sloping demand curves, and further assuming that the marginal cost curves for each area, given the other area's coverage, are increasing at increasing rates. The remaining second order condition will be met if the advantage of further diversification between market areas does not compensate for the rate of increase of costs within markets. (We would also expect that $F^c_{I^cI^s} < 0$ and $F^s_{I^cI^s} < 0$.)

In order to discover the urban-suburban property insurance mix we will return to equations (44). Each of these equations could be called a "partial-optimization" function for each market area. That is, given the other market area's insurance coverage, each equation of (44) yields the maximizing behavior for the insurer in that market area. This system of equations (44) should not be interpreted as Cournot or Stackelberg Reaction functions because the insurer, in fact, has control of all variables simultaneously. It is therefore a static equilibrium analysis.

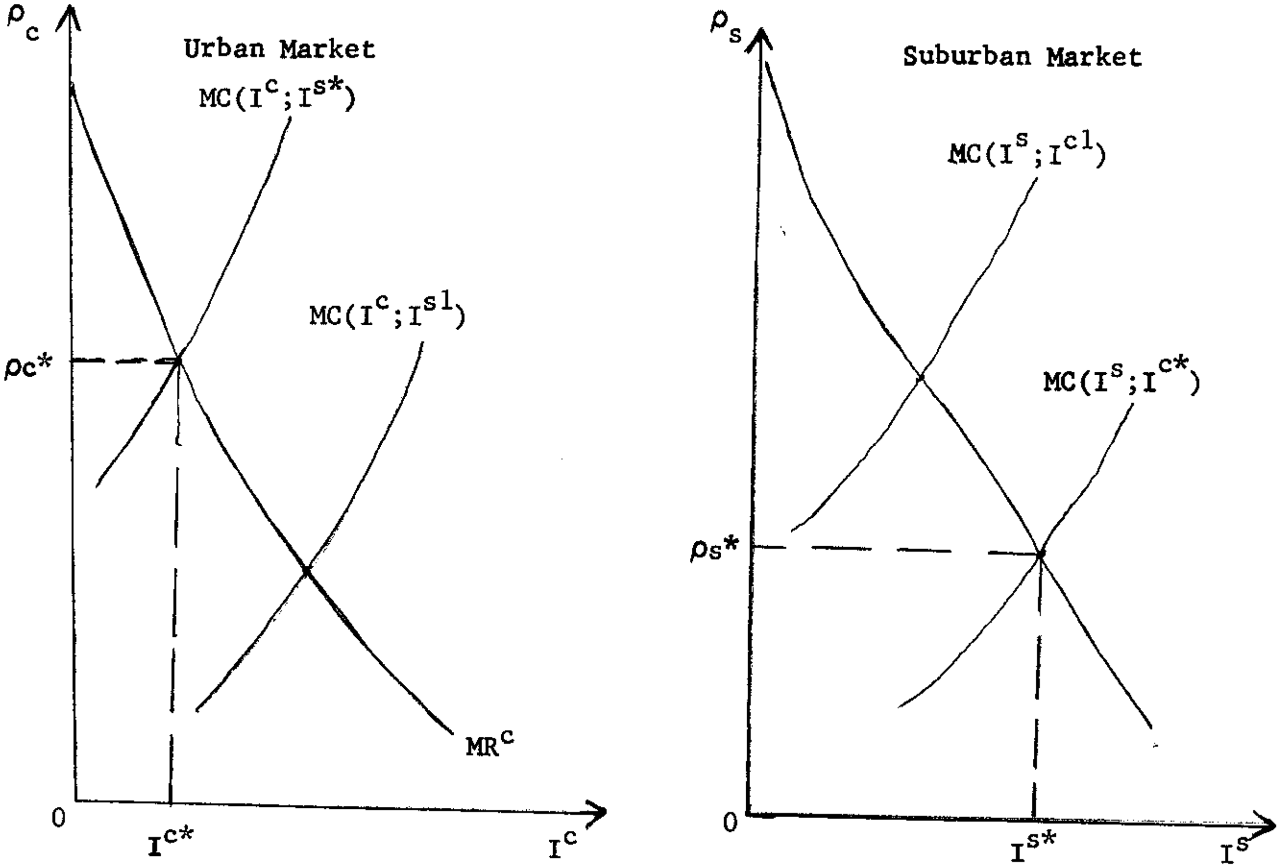
Figure 10 depicts the partial-optimization functions for each market $R^c(I^s)$ and $R^s(I^c)$, and the static equilibrium at their intersection, I^{c*} and I^{s*}. If the second order conditions are met, the intersection of the partial-optimization functions represents the mix of insurance between the urban core and suburban area which will maximize the expected net revenue for the insurer.

It is possible, however, that R^c and R^s will touch on the same axis and/or not intersect in the I^s - I^c plane. This case could occur if, for instance in Figure 9, MC^c did not intersect the MR^c for $0 \leq I^s \leq I^{s'}$. In such a case the optimal I^c would be zero for $I^s \leq I^{s'}$. This is not entirely implausible if one believes that the urban market is significantly riskier than the suburban market. It could also occur if the MR^c were small because of low levels of urban core insurance demand. If either or both cases exist, the profit maximizing equilibrium without regulation could occur when $I^c = 0$ and $I^s > 0$.

b) Monopolistic Competition and Rate Regulation

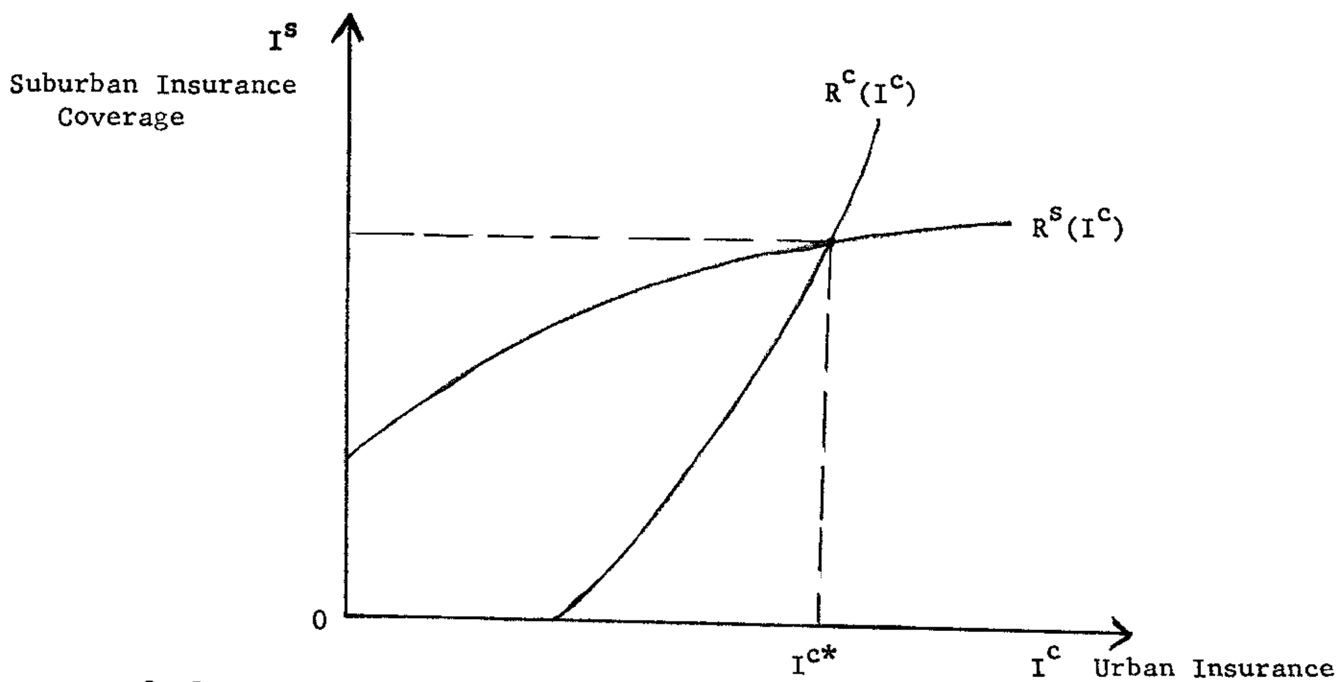
As discussed earlier, there exist significant inflexibilities in the insurance rates a company can charge. The urban-suburban monopolistic competition insurance analysis will now incorporate the effects of maximum rate ceiling regulation. In particular, we will assume that the maximum regulated price ceiling will be lower than (at least) the equilibrium insurance rate that would persist in the urban core market if left unregulated. Figure 11 demonstrates that a price ceiling in the urban core insurance market creates a new marginal revenue curve, with a discontinuity

Figure 9
Urban and Suburban Insurance Markets --
Demand and Cost Data for the Typical Insurance Firm



MC is the expected marginal cost function for each market
MR is the marginal revenue function for each market

Figure 10
Partial-Optimization Functions for the Urban-Suburban Insurance Markets

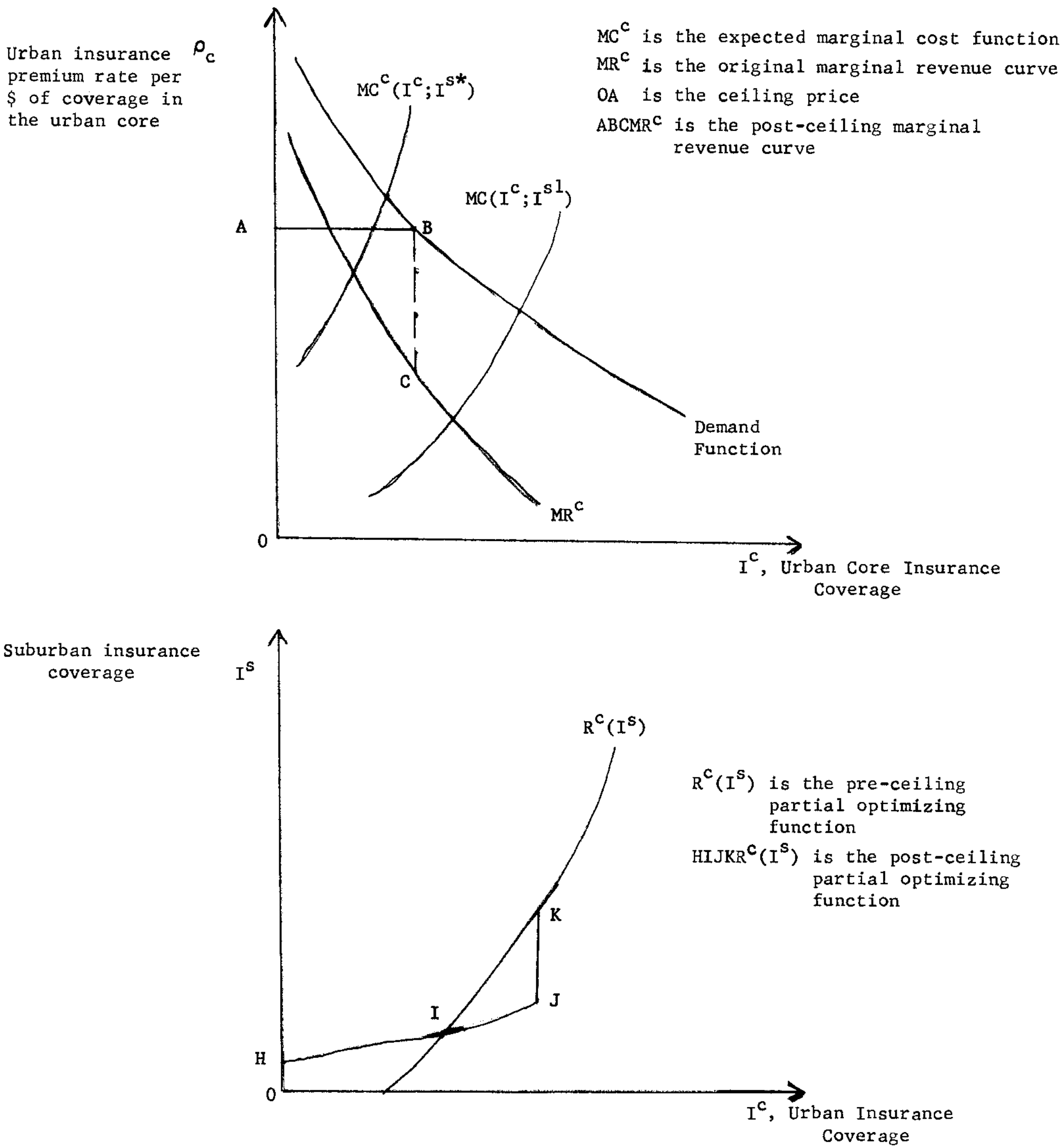


$I^c = R^c(I^s)$ is the partial-optimization function for the urban core market area

$I^s = R^s(I^c)$ is the partial-optimization function for the suburban market area

Figure 11

The Effects of a Price Ceiling on the Urban Core Insurance Market and the Partial Optimizing Function



where the price ceiling intersects the demand curve, and simultaneously an intersection with the original marginal revenue curve. The effects of the rate ceiling are traced out in Figure 11 for the partial optimization function. For I^c greater than the point where the ceiling price intersects the demand function, the marginal revenue curves for post- and pre-ceiling regulation coincide, and, likewise, the partial optimization functions. It is also clear that the post-ceiling equilibrium insurance coverage in the urban core market could be greater than, less than, or equal to the pre-ceiling equilibrium solution, depending upon the new intersection of the two partial-optimization functions. (Moreover, the price ceiling analysis similarly can be applied to the suburban insurance market.)

It is generally believed that the risk differential based upon expected losses for the urban core market is significantly greater than that in the surburban insurance market, implying a higher risk-adjusted expected marginal cost curve at each level of insurance coverage. Furthermore, current rate regulations are reputed to be an effective ceiling in the urban core market, causing an excess demand at the ceiling price (i.e., a market disequilibrium). On the other hand, the maximum regulated rate, in general, is not necessarily considered to be an effective ceiling in terms of the equilibrium price for insurance coverage charged in the suburban market. Furthermore, the gains from diversification, in terms of the "suburban" partial optimization function, may be minimal. In other words, $R^s(I^c)$ is relatively constant with a virtually zero slope in the $I^c - I^s$ plane. _A priori_ one can not state if the imposed maximum rate ceiling will alter the equilibrium quantity of insurance in each market.

The pre- and post-regulation insurance mix depends upon the relationship of the demand functions, cost functions, and the imposed rate and the unregulated equilibrium rates in each market.

In general, if the regulated ceiling rate is <u>much lower</u> than the pre-regulatory market clearing rate in the urban market, it will eliminate the urban core insurance coverage altogether, and reduce the quantity of insurance sold in the suburban market.[53] If the regulated ceiling is <u>very close to</u> the original market clearing rate in the urban core market, both markets will expand with lower equilibrium market rates for each.

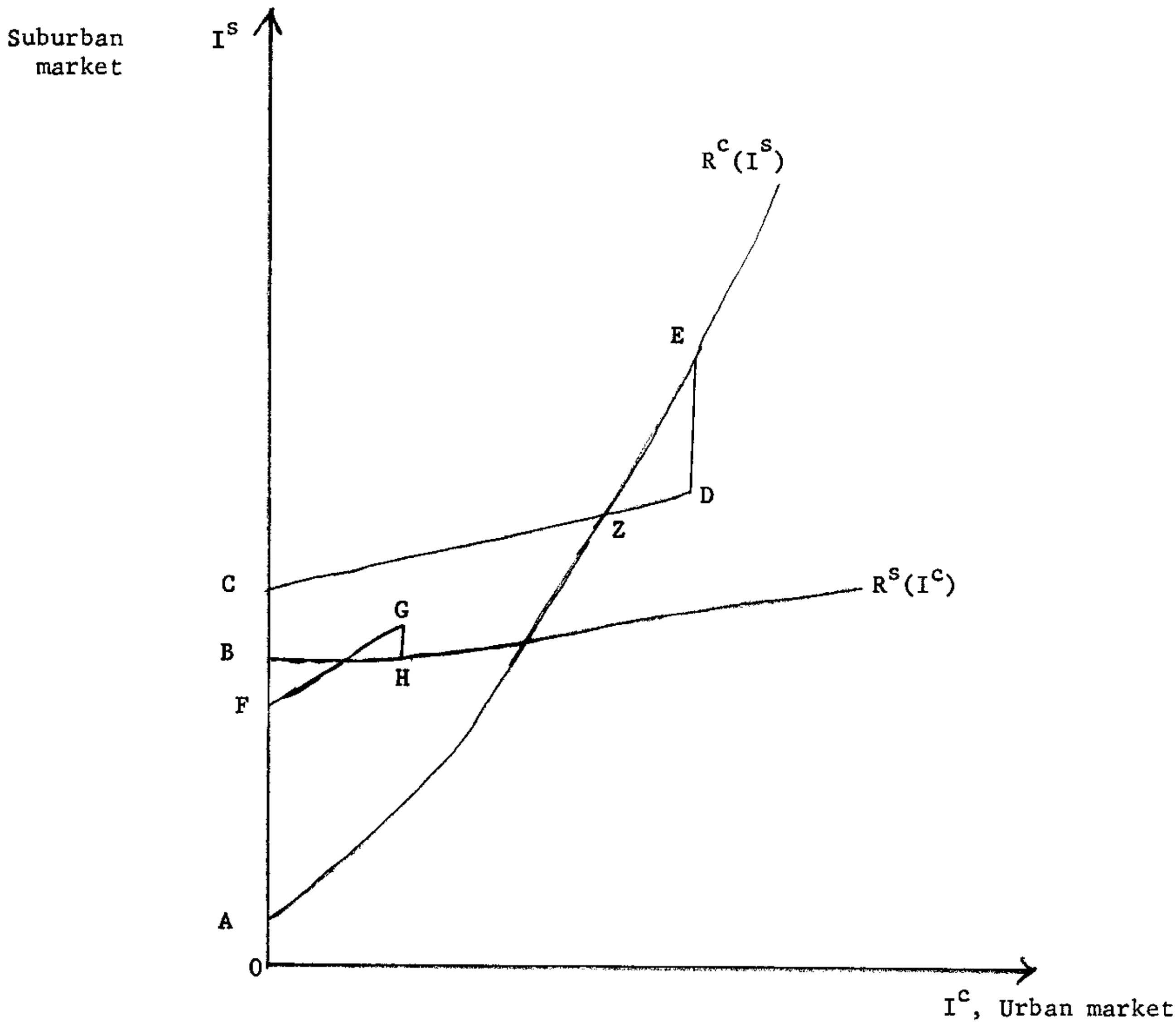
Figure 12 depicts the likely relationship which currently exists in the urban-suburban markets. An unregulated equilibrium would occur at point Z. The maximum regulated ceiling creates partial optimization functions, represented by OCDE in the urban core market and FGH in the suburban market. The post-ceiling equilibrium will occur at point F, with no insurance sold in the urban core and OF insurance coverage sold to the suburban area (a reduction from the unregulated equilibrium).

c) Pricing Regulation and Policy Implications[54]

It is possible that insurance price regulation when compared to the unregulated market may produce (in equity terms) an undesired reduction in urban core insurance coverage. Furthermore, with the existence of imperfect competition, socially non-optimal quantities of insurance will be sold in all markets. If, on equity grounds and/or societal externalities grounds, urban core insurance coverage should be increased, premium rate regulation should be rejected in favor of the more efficient combination of an insurer subsidy-tax and the insured's lump-sum tax as discussed in Section II above.

Figure 12

Urban-Suburban Insurance Mix, Assuming High Urban Expected Marginal Costs and Minimal Expected Marginal Cost Reduction in the Suburban Market by Diversification



$AR^c(I^s)$ is the non-regulated partial optimization function for the urban market

$BR^s(I^c)$ is the non-regulated partial optimization function for the suburban market

$CDER^c$ is the partial optimization function for the urban market with a premium rate ceiling

$FGHR^s$ is the partial optimization function for the suburban market under a premium rate ceiling

d) Other Government Programs

Another problem with the insurance market is the imperfections of access by individuals to insurers. This is particularly true with respect to urban core property. To combat this imperfection several property inspection plans have been initiated; whereby an individual, who cannot find an insurer, can submit his property for inspection by an "impartial" government agency, and upon passing the inspection, will be guaranteed insurance. This type of plan goes under the rubric of FAIR[55] (Fair Access to Insurance Requirements).

FAIR plans increase available information to the insurer and insured at little or no cost to them. However, the social costs may be significant. The set-up and maintenance of an inspection bureau is costly. Therefore, one can think of the increment of information as a force, (coupled with our subsidy-taxing scheme) which moves the market closer to the Pareto optimal; but also reduces the surplus left to consumers. In fact, it is possible that the inspection plan costs will be so expensive that it will be on balance welfare-reducing.[56]

In addition, FAIR plans may have a practical implementation problem. The issuance of insurance under FAIR plans depends ultimately upon the cooperation from agents and brokers operating in urban core areas. In many instances the broker has the option of placing coverage through FAIR plans or substandard rates plans (e.g., excess rates schedules). Typically, the FAIR premium rates and the commission rates received by the broker are significantly lower than those in alternative insurance coverage. Clearly this represents a pecuniary disincentive at a most important level in the operation of FAIR plans.

Another program for crime insurance which attempts to avoid the costs of information associated with extensive on-site inspections has been organized under the auspices of H.U.D. This program will issue coverage at a uniform premium within a pre-specified area, as described in the following:

> "The Government announced today that it would begin selling crime insurance August 1, 1971 ... where private insurance is not available at reasonable rates. ... Premium rates (will reflect) ... crime statistics ... for each metropolitan area. ... The department adopted a provision of the regulations earlier that makes premium rates uniform throughout a Standard Metropolitan Statistical Area, ... an area (being) characterized (as) a central city and its suburbs."[57]

As was discussed above in Section II in the analysis of the optimal suburban-urban insurance mix, the H.U.D. uniform premium rate and a subsidy lump sum tax plan alternative would be premised on an equity judgment. The latter is, however, a more efficient technique to achieve any particular level of wealth redistribution. The social value of the redistribution is entirely a value-judgment. However, there may exist other arguments for uniform premium rates. If the government is to consider the general well-being of society, and, over time, areas of a city change from superior to inferior risks, and vice versa, the uniform premium rate plan will protect the insured against state changes of areas, which might lead to undesirable uncertainty about future premium rates. As mentioned in relation to the FAIR plans, an idealized insurance schedule would charge each policyholder according to his property's risk. However, this information may be <u>costly</u> to ascertain precisely; a uniform premium rate may save significant information gathering costs.

Overall, it is evident that none of the insurance panaceas currently in use or proposed here or elsewhere will bring about the desired effects without government intervention. Moreover, the government intervention required will be strenuous and multifaceted; it will tax the consumer, tax and subsidize the insurer, and may set-up several regulatory and inspection agencies.

FOOTNOTES

*Assistant Professor of Finance, The Wharton School, University of Pennsylvania. The author wishes to thank the Rodney L. White Center for Financial Research for financial support and Professors Martin Feldstein and Wassily Leontief for their intellectual support. The author also wishes to thank his colleagues of the Finance Department of the University of Pennsylvania for their innumerable helpful comments. This paper represents a significant revision of an earlier paper of the same title. Of course, any errors are the responsibility of the author.

[1]Three seminal studies investigating the role of expectations in the theory of asset preferences are:

a. John R. Hicks, "A Suggestion for Simplifying the Theory of Money," Economica, New Series, Vol. II (1935), pp. 1-19.

b. Harry Markowitz, "Portfolio Selection," Journal of Finance, Vol. VII, No. 1 (March, 1952), pp. 77-91.

c. James Tobin, "Liquidity Preference as Behavior Towards Risk," Review of Economic Studies, Vol. XXV, Nos. 66, 67 and 68 (1957-1958), pp. 65-86.

[2]Examples of previous analysis related to the theory of insurance are:

a. K. J. Arrow, "Uncertainty and the Economics of Medical Care," American Economic Review, Vol. LIII, No. 5 (December, 1963), pp. 941-973.

b. Karl Borch, "Some Elements of a Theory of Reinsurance," Journal of Insurance, Vol. XXVIII, No. 3 (September, 1961), pp. 35-43.

c. Isaac Ehrlich and Gary S. Becker, "Market Insurance, Self-Insurance and Self-Protection," Center for Mathematical Studies in Business and Economics, Report 7110 (Chicago: Department of Economics and Graduate School of Business, University of Chicago, February 1971), 52 p.

d. John P. Gould, "The Expected Utility Hypothesis and the Selection of Optimal Deductible for a Given Insurance Policy," Journal of Business, Vol XLII, No. 2 (April, 1969), pp. 143-151.

e. Jan Mossin, "Aspects of Rational Insurance Purchasing," Journal of Political Economy, Vol. LXXVI, No. 4 Pt. 1 (July/August, 1968), pp. 553-568.

f. B. Peter Pashigian, Lawrence L. Schkade, and George H. Menefee, "The Selection of an Optimal Deductible for a Given Insurance Policy," Journal of Business, Vol. XXXIX, No. 1, Pt. 1 (January, 1966), pp. 35-44.

g. John Pratt, "Risk Aversion in the Small and in the Large," Econometrica, Vol. XXXII, No. 1-2 (January-April, 1964), pp. 122-136.

h. Vernon L. Smith, "Optimal Insurance Coverage," Journal of Political Economy, Vol. LXXVI, No. 1 (January-February, 1968), pp. 68-77.

[3]Frank K. Knight's distinction between "risk" and "uncertainty" will not be maintained in this essay (Risk, Uncertainty and Profit [New York: Harper and Row, 1921]). The essence of Professor Knight's idea is captured in the following quote from page 46:

> "The fact that while a single situation involving a known risk may be regarded as 'uncertain,' this uncertainty is easily converted into effective certainty; for in a considerable number of such cases the results become predictable in accordance with the laws of chance, and the error in such prediction approaches zero as the number of cases is increased. Hence it is simply a matter of an elementary development of business organization to combine a sufficient number of cases to reduce the uncertainty to any desired limits. This is, of course, what is accomplished by the institution of insurance."

The distinction is semantic rather than economic. In particular, insurance policies have been contracted even though in a considerable number of cases risk spreading is not possible; or a probability table can not be constructed for the particular case. This point is discussed further in a slightly different context in section I-C and n.8 below.

[4]Also, all random losses are not insurable even if "probability tables" are constructable. For example, if one could calculate the probability of the sun becoming a nova, it would be clearly an uninsurable event. Also, and more important if the art of insuring causes behavior changes in the insured, the insurer will be facing "moral hazard." Clearly, moral hazard will represent one of the limits of insurability. For example, if people are more likely to commit suicide to transmit resources to their beneficiaries, or if businessmen have a proclivity to set afire their businesses when the firm's sales slump in order to collect insurance, using pre-insured historical data to determine premiums will significantly diminish expected profits. This type of issue will arise in varied contexts in our subsequent discussion. See in particular n.20 below.

[5]An excellent intuitive discussion of the assumptions and implications for an expected utility analysis can be found in R. Duncan Luce and Howard Raiffa, Games and Decisions: Introduction and Critical Survey (New York: John Wiley & Sons, 1957), especially pages 19-31.

More technical discussions can be found in the classic work of John Von Neumann and Oskar Morgenstern, The Theory of Games and Economic Behavior (Princeton: Princeton University Press, 1944), and in I.N. Herstein and John Milnor, "An Axiomatic Approach to Measurable Utility," Econometrica, Vol. XXI, No. 2 (April, 1953), pp. 291-297.

[6]A premium rate of $1.00 per $100 of protected value is in our example the actuarially fair rate. However, insurance companies also

have administrative costs as well as loading charges in order to secure actual profits that will be greater than all actual costs. Therefore, one would believe that the premium would be set above the $1.00 rate.

[7] Pratt, op. cit., and Mossin, op. cit., show these assumptions about the utility function to be fairly standard. Also, we assume, for the moment, that there exists some $\rho^* > 0$ for which full coverage insurance is optimal for the insured.

[8] As observed in n.3 above, we do not maintain the Knightian distinction between risk and uncertainty as a suggestive device. Implicity it should be understood that the distinction with respect to the analysis of expected utility optimization is fallacious. It should also be noted that this issue has engendered a prolific controversy.

a. Daniel Ellsberg, "Risk, Ambiguity, and the Savage Axioms," Quarterly Journal of Economics, Vol. LXXV, No. 4 (November, 1961), pp. 643-669.

b. William Fellner, "Distortion of Subjective Probabilities as a Reaction to Uncertainty," Quarterly Journal of Economics, Vol. LXXV, No. 4 (November, 1961), pp. 670-689.

c. Howard Raiffa, "Risk, Ambiguity, and the Savage Axioms: Comment," Quarterly Journal of Economics, Vol. LXXV, No. 4 (November, 1961), pp. 690-694.

d. Also, see subsequent discussion of this topic in the Quarterly Journal of Economics, Vol. LXXVII, No. 1 (February, 1963), pp. 159-161; No. 2 (May, 1963), pp. 327-342; No. 4 (November, 1963), pp. 676-690; Vol. LXXIX, No. 4 (November, 1965), pp. 657-663.

e. Selwyn W. Becker and Fred O. Brownson, "What Price Ambiguity? Or the Role of Ambiguity in Decision Making," Journal of Political Economy, Vol. LXXII, No. 1 (February, 1964), pp. 62-73.

[9] These can be generalized to m states, each occurring with a subjective probability P_m and a theft probability π_m. The total expected utility in the uninsured situation is

$$E(U) = \sum_m P_m \left(\pi_m U(A) + (1 - \pi_m)\, U\,(A + H)\right)$$

with the side constraint that $\sum_m P_m = 1$

As before, the maximum premium rate, ρ^*, can be found by equating the expected utility of the uninsured circumstances with the utility of the certain situation:

$$E(U) = U(A + H\,(1 - \rho^*))$$

In other words, if we subsume the expected utility hypothesis, we use a prior distribution (the P's) on the theft probability distribution (the π's) in order to find the maximum ρ^*.

[10]Total differentiation of (6) yields

$$\frac{-(U(A) - U(A + H))}{H \cdot U'(A + H(1 - \rho^*))} = \frac{d\rho}{d\pi} > 0$$

This is true on the interval $[A, A + H]$ because under our assumptions the numerator and denominator are both positive.

[11]The change in the value of ρ^*, the maximum premium rate per dollar of coverage, is ambiguous. This is found by examining

$$\frac{d\rho^*}{dH} = \frac{-1}{H \cdot U'(A + H(1 - \rho^*))} \left[(1 - \pi)U'(A + H) - (1 - \rho^*)U'(A + H(1 - \rho^*))\right]$$

The unbracketed term on the right hand side for $\rho \epsilon [0,1]$ is clearly negative. Therefore, the effect of a change of H upon ρ^* depends upon the sign of the bracketed terms.

To discover the sign of the bracketed terms, we define

$$B(\rho) = (1 - \pi)\, U'(A + H) - (1 - \rho)\, U'(A + H(1 - \rho))$$

for $\rho \epsilon [0,1]$. From equation (6) we substitute in for $(1 - \pi)$, yielding

$$B(\rho) = \frac{U'(A + H) \cdot U(A)}{U(A) - U(A + H)} - \frac{U'(A + H) \cdot U(A + H(1 - \rho))}{U(A) - U(A + H)} - (1 - \rho) \cdot U'(A + H(1 - \rho))$$

Note that $B(\rho = 0) = B(\rho = 1) = 0$. Therefore, assuming the appropriate differentiability of $B(\rho)$, by the Mean Value Theorem of the Calculus, we know that an extrema exists for $B(\rho)$ on the interval $\rho \epsilon (0,1)$.

$$B'(\rho) = U'(A+H(1-\rho)) \left[\frac{H \cdot U'(A+H)}{U(A)-U(A+H)} + 1-(1-\rho) \cdot \left[\frac{-U''(A+H(1-\rho))}{U'(A+H(1-\rho))}\right]\right]$$

The bracketed term within the brackets is the risk aversion of the individual (see Pratt, _op. cit._), and is assumed to a decreasing function of wealth. Hence, it is positive and monotonically increasing for $\rho \epsilon (0,1)$. Therefore, for an extreme to occur a side condition is necessary:

$\frac{H \cdot U'(A + H)}{U(A + H) - U(A)} < 1$. The economic significance of the side condition is that H is small relative to total wealth implying that the marginal utility of A + H is also small. If this condition holds $B(\rho)$ will vanish at, say, ρ^* on the open interval (0,1) and $B(\rho^*)$ will be a maximum because the bracketed terms will initially be positive for $\rho < \rho^*$, zero at ρ^*, and finally negative for $\rho > \rho^*$. Therefore $B(\rho)$ is positive. Hence subsuming the side condition the total derivative $\frac{d\rho^*}{dH}$ is negative.

However, it is also clear that the side condition may not hold if the risky portion of one's wealth, H, is very large. In such cases, increased H will be accompanied by increased ρ^*. However, the total premium bill one is willing to pay for full coverage will always increase as H increases. Note that the change in total premium payment can be represented as:

$$\frac{d}{dH}(H \cdot \rho^*) = \rho^* + H\frac{d\rho^*}{dH}$$

By rearranging the derivative of ρ^* with respect to H, under our assumptions, we will discover that:

$$\frac{d}{dH}(H \cdot \rho^*) = 1 - \frac{(1-\pi)\,U'(A+H)}{U'(A+H(1-\rho^*))} > 0$$

[12] Substitute from equation (6) for π in the numerator of equation (9). Upon rearrangement, the numerator of (9) will be a function of ρ^*:

$$f(\rho^*) = -U'(A)\frac{[U(A+H(1-\rho^*)) - U(A+H)]}{U(A) - U(A+H)} - U'(A+H)\frac{[U(A)-U(A+H(1-\rho^*))]}{U(A)-U(A+H)} + U'(A+H(1-\rho^*))$$

Note that $f(\rho^*=0) = f(\rho^*=1) = 0$. That is, the limit values of ρ^* will cause the numerator of (9) to vanish. By the Mean Value Theorem of the Calculus, we know that an extreme for $f(\rho^*)$ will exist on the interval $0<\rho^*<1$, assuming appropriate differentiability of $f(\rho^*)$:

$$f'(\rho^*) = \left[\frac{U'(A) - U'(A+H)}{U(A) - U(A+H)}\right] H \cdot U'(A+H(1-\rho^*)) - U''(A+H(1-\rho^*)) \cdot H$$

The fraction in the first term on the right hand side is a negative constant, -C. Therefore $f'(\rho^*)$ can be rearranged as

$$f(\rho^*) = H \cdot U'(A+H(1-\rho^*)) \left[-C - \frac{U''(A+H(1-\rho^*))}{U'(A+H(1-\rho^*))}\right]$$

The second term in the brackets has been called by Pratt, <u>op. cit.</u>, the risk aversion of the individual. In particular, if we assume <u>the</u> person has decreasing risk aversion as a function of wealth, as ρ^* increases the risk aversion term will monotonically increase.

In order for the derivative of f to vanish on the open interval $(0<\rho^*<1)$, the second term in the brackets must vanish as ρ^* increases. Hence, the sum of the bracketed terms must be initially negative, monotonically increasing, and vanishing somewhere on the open interval, which implies that $f'(\rho^*) = 0$ is a minimum point. Therefore, $f(\rho^*)$ on the internal $0 < \rho^* < 1$ is negative and

$$\frac{d\rho^*}{dA} < 0.$$

[13]The introduction of δ and ε reduces the original Range of Wealth, H: $R=H - \delta-\varepsilon < H$. For all $A + \delta \leq A+H-\varepsilon$.

[14]Total differentiation of (10) is used to find the appropriate values of ε for each δ. That is,

$$\pi\, U'(A+\delta)\, d\delta - (1-\pi)\, U'(A+H-\varepsilon)\, d\varepsilon = 0$$

will be used to find the relevant ε and δ. The first term of the equation is the additional expected utility gained by an increase of δ, the second is the decrease in utility caused by the reduction of A+H by ε. In order to maintain the same level of satisfaction for changes in δ and ε the two terms sum to zero.

[15]In this analysis, additional indifference maps are derived by assuming π and A remain unchanged. We vary the value of H, the risk asset, and then proceed to derive a utility frontier such as B E in figure 3.

[16]Analogous results may be derived by using the standard deviation of wealth as the risk measure rather than the range. I believe it is clear that in our example with only two possible outcomes (i.e., theft-no theft) the range is a more appropriate measure of dispersion and, therefore, risk. If you wished to use the standard deviation as your risk measure, equation (11) in the text remains as it is; equation (12) is rewritten as (12'):

$$\sigma_W = (\pi(1 - \pi)\,)^{\frac{1}{2}}\, H(1 - \lambda)$$

The new opportunity line derived by substitution will be (13'):

$$E(W) = A + H\,(1-\rho) + (\rho - \pi)\,(\pi\,(1 - \pi)\,)^{-\frac{1}{2}}\,\sigma_W$$

Therefore, all subsequent results in sections c and d using the range of wealth as the index of risk could be reworked in terms of the standard deviation of wealth.

[17]This equation is derived by total differentiation of the utility frontier with respect to λ, the choice variable of the individual. Also, note that this treatment is equivalent to the analysis in terms of R since $R = (1-\lambda)H$. That is, E(U) = constant =

$$\pi U(A+(H-R)\,(1-\rho)) + (1-\pi)\,U(A+H(1-\rho) + \rho R)$$

[18]More correctly, a maximum occurs even if a tangency does not when

$\frac{dE(U)}{d\lambda} \geq 0$, which, of course, will lead one to increase λ up a maximum of unity.

$$\frac{d^2E(U)}{d\lambda^2} = \pi H^2(1-\rho)^2\ U''(A+\lambda H(1-\rho)) + (1-\pi)H^2\rho^2 U''(A+H(1-\lambda\rho)) < 0$$

The satisfaction of the second order condition is implied by $U''(W) < 0$ with $0 \leq \pi \leq 1$ and $0 \leq \rho \leq 1$. Unless otherwise stated we will assume these conditions are satisfied.

[19]Evaluating (16) at $\rho = 1$ and $\rho = 0$ yields interesting information.

$$\frac{dE(U)}{d\lambda} = (1-\pi)HU'(A+H(1-\lambda)) < 0 \text{ for } \rho = 1$$

and $1 > \pi \geq 0$. This is true for all $\lambda \in [0,1]$, and signifies that no insurance will be purchased. Stated in words, one would never purchase insurance if the premium rate equals the value of the protected object.

$$\frac{dE(U)}{d\lambda} = \pi\ H\ U'(A+\lambda H) > 0 \text{ for } \rho = 0,\ 0 < \pi \leq 1, \text{ for all } \lambda \in [0,1].$$

This denotes that if insurance is a costless commodity, one will accept full-coverage. Therefore, the ensuing analysis assumes $0 < \rho < 1$.

[20]For a discussion of "Moral Hazard" see

a. K. J. Arrow, "Uncertainty and the Welfare Economics of Medical Care," American Economic Review, Vol. LIII, No. 5 (December, 1963), pp. 961-962.

b. __________, "The Economics of Moral Hazard: Further Comment," American Economic Review, Vol. LVIII, No. 3, Pt. 1 (June, 1968), pp. 537-539.

c. __________, Aspects of the Theory of Risk-Bearing (Helsinki: Yrjö Johanssonin Säätiö, 1965), pp. 55-56.

d. Mark V. Pauly, "The Economics of Moral Hazard: Comment," American Economic Review, Vol. LVIII, No. 3, Pt. 1 (June, 1968), pp. 531-537.

[21]We have chosen to make the random variable loss model discreet rather than continuous because it is possible that certain value points of L_i will have probabilities associated, an impossibility for a continuous probability density function. However, as we create finer partitions of L_i and π_i, the analysis approaches a "continuous" random variable loss model.

[22]

$$\frac{dE(U)}{d\lambda} = 0 \text{ and}$$

$$\frac{d^2E(U)}{d\lambda^2} = \rho^2 H^2 \Sigma \pi_i U'' (A+H(1-\lambda\rho))$$

$$+ H^2(1-\rho) \Sigma \pi_h U'' (A+H(1+\lambda(1-\rho))-L_h) < 0$$

for $0 < \lambda < 1$ are the necessary and sufficient conditions for optimization equilibrium.

[23]In economics terminology, the corner solutions occur when

$$\frac{dE(U)}{d\lambda} \leq 0 \text{ for } \lambda = 0 \text{ and}$$

$$\frac{dE(U)}{d\lambda} \geq 0 \text{ for } \lambda = 1$$

[24]We shall not derive the results here in order to save space and maintain simplified mathematical presentations.

[25]A random loss model will yield the same results, with somewhat more complicated mathematics. The expected utility will be

$$E(U) = \Sigma \pi_i U (A+H(1-\rho(d)\lambda) - L_i) +$$

$$\Sigma \pi_h U (A+H(1-\rho(d)\lambda) - d) +$$

$$\Sigma \pi_j U (A+H(1+\lambda(1-\rho(d)) - d - L_j)$$

where $i \in 0 \leq L_i \leq d$

$h \in d < L_h \leq \lambda H$

$j \in \lambda H < L_j \leq H$

The mode of analysis proceeds identically to the text.

[26]Arrow, "Uncertainty and the Welfare Economics of Medical Care," mathematical appendix, proposition 1, gives the proof.

[27]Ehrlich and Becker, loc.cit. Also see n. 20.

[28]Ehrlich and Becker, loc.cit., pp. 18-26.

[29]Many of my ideas and notions for this section of the essay have evolved from work done by Karl Borch, such as:

a. Karl Borch, "A Utility Function Derived from a Survival Game," Working Paper No. 66 (Los Angeles, California: Western Science Management Institute, Graduate School of Business, University of California at Los Angeles, February 1965), 12 pp.

b. __________, "Equilibrium in a Reinsurance Market," Econometrica, Vol. XXX, No. 3 (July, 1962), pp. 424-444.

c. __________, "Reciprocal Reinsurance Treaties," The Astin Bulletin, Vol. I (1960), pp. 170-191.

d. __________, "Some Elements of a Theory of Reinsurance," Journal of Insurance, Vol. XXVIII, No. 3 (September, 1961), pp. 35-43.

e. __________, "The Optimal Management Policy of an Insurance Company," Proceedings of the Casualty Actuarial Society, Vol. LI, Pt. 11, No. 96 (1964), pp. 182-197.

Other discussions which were helpful in formulating my analysis in this section were:

a. R. E. Beard, T. Pentikainen and E. Pesonen, Risk Theory (London: Methuen & Company, 1969), pp. 7-40.

b. Hilary Seal, Stochastic Theory of a Risk Business (New York: John Wiley & Sons, 1969), pp. 4-132.

[30]This proof is developed along similar lines to Robert Dorfman's "Mathematical Models: The Multistructure Approach," (Maass, Arthur, et al., Design of Water Resource Systems [Cambridge: Harvard University Press, 1966], pp. 494-539.

[31]Administrative costs can be incorporated into the analysis. For example, assume administrative costs were a function of the number of policyholders the company had: $A = A(n)$. One could continue the analysis suggested by adding this to the cost function, and then choosing not only the optimal premium rate but the optimal number of clients (i.e., optimal size of market share).

[32]We will not discuss the technique of risk reduction by reinsurance, though it is an appropriate tool for a risk averse insurance company. You may assume that either the company is the ultimate insurer, such as a government agency, or reinsurance is unavailable. For an excellent discussion of reinsurance, see n. 29, references a, c and d, above. Also, the notion of risk-liquidity-survival is discussed in reference c, n. 29 above.

[33]For example, see Samuel S. Wilks, Mathematical Statistics, (New York: John Wiley & Sons, 1962), pp. 257-258.

[34]See David H. Pyle and Stephen J. Turnovsky, "Safety-First and Expected Utility Maximization in Mean-Standard Deviation Portfolio Analysis," Review of Economics and Statistics, Vol. LII, No. 1 (February,

1970), pp. 75-81 for an interesting discussion of the "ruin avoidance" behavior as a form of expected utility maximization. Also, our discussion will focus upon the one period risk of illiquidity or ruin. In fact, an insurance company must concern itself with this as well as a multi-period probability ruin.

[35]In extreme cases, the insurance company does not solely determine the premium rate ρ. It may refuse to underwrite any policy from a particular area, or contract policies of only low risks from certain districts; the former policy is called "red-pencil" insuring, and the latter is known as "gray-pencil" insuring. The net result is that insurance companies control ρ, n, and the quality of risks.

[36]Claims of random size can be introduced into the problem of total claims by using a Generalized Poisson Distribution. For a sophisticated discussion of this mathematical device, see D. R. Cox and H. D. Miller, The Theory of Stochastic Processes (New York: John Wiley & Sons, 1968), pp. 146-56. In the case for variable claims, one assumes that there is a probability distribution function for claims, h(C) which is independent of the number of claims and is mutually independent for the size of each claim. Then the Generalized Poisson function is

$$F(C) = \sum_{i=0}^{k} (h(C))^i \, e^{-n\pi} \frac{(n\pi)^i}{i!}$$

[37]Meeting the Insurance Crisis of Our Cities, Report by the President's National Advisory Panel on Insurance in Riot-Affected Areas (Washington: U. S. Government Printing Office, 1968). This document has the best all-inclusive discussion of the problems of urban area insurance. The introductory comments in this section of the paper are ostensibly adapted from it.

[38]J. J. Launie, "The Supply Function of Urban Property Insurance," The Journal of Risk and Insurance, Vol. XXXVI, No. 3 (June, 1969), pp. 269-283. See particularly pages 271-273.

[39]John R. Lewis, "A Critical Review of the Federal Riot Reinsurance System," The Journal of Risk and Insurance, Vol. XXXVIII, No. 1 (March, 1971), p. 29.

[40]Meeting the Insurance Crisis of our Cities, op. cit., p. 25.

[41]H also may be a variable: The arguments discussed will remain true, with the mathematics becoming more complex.

[42]Assume that second order conditions are satisfied.

[43]Since I is a function of λ and n, the optimal I and ρ depend upon the optimal λ and n. This is equivalent to our analysis in (41) because

$$dE(W) = 0 = d(\rho \cdot I) - d(C) = \rho' \cdot I + I' \rho - C'$$

where $\rho=\rho(\lambda,n)$, $C = C(\lambda,n)$, and $I = I(\lambda,n)$
Let f_x be the partial derivative of f with respect to X, then

$$dE(W) = 0 = (I \cdot \rho_\lambda + \rho \cdot I_\lambda - C_\lambda)d\lambda + (I \cdot \rho_n + \rho \cdot I_n - C_n)dn$$

For the total differential of wealth to be zero (i.e., an extrema for wealth), each term in the parenthesis must vanish. In terms of economics, this means that the expected marginal revenue equals the expected marginal cost for the last accepted policy-holder and the last dollar of protection purchased by each policy-holder.

[44] This is indeed a strong assumption, and implies that there is a zero "wealth effect." Note that a zero "wealth effect" does not imply constancy of the marginal utility of wealth.

[45] The generalization to the variable loss case follows the exact argument presented in the text for the binary model. If $f_c(L)$ and $f_s(L)$ represent the probability density functions for variable losses, L, for the core (c) and the suburbs (s), respectively.

$$E(L_c) = \int_0^H L_c \, f_c(L) \, dL > \int_0^H L_s \, f_s(L) \, dL = E(L_s)$$

The premium rate for the two submarkets depends upon the λ's chosen in each market as well as N and M.

$$\pi = \pi(\lambda_c, \lambda_s) = \frac{E(L_c|\lambda_c) + E(L_s|\lambda_s)}{N+M}$$

The actuarial rate in each market depends upon the size of λ. That is $\pi(\lambda_s, \lambda_c) = \pi_s(\lambda_s) + \ell(\lambda_s, \lambda_c)$ for the suburbs where $\ell > 0$, and $\pi(\lambda_s, \lambda_c) = \pi_c(\lambda_c) + b(\lambda_c)$ for the urban core where $b(\lambda_c) < 0$ is the subsidy.

$$E(U_s) = \int_0^{\lambda H} U_s(A + H(1-\lambda\pi) \, f_s(L)dL + \int_{\lambda H}^{H} U_s(A + H(1+\lambda\,(1-\pi))-L_s)f_s(L)\,dL$$

is the expected utility under the uniform premium rate plan. The lump sum tax plan will induce full coverage:

$$E(U_s) = U_s \, (A + H\,(1-\pi_s(\lambda_s=1)) - \ell\,(\lambda_s,\lambda_c)\cdot\lambda_s\cdot H)$$

Moreover, $A + H\,(1-\pi_s) - \ell\lambda H \equiv A + H - \lambda H(\pi_s + \ell) - \int_{\lambda H}^{H} (L_s - \lambda H)\, f_s(L)dL.$

Therefore, the lump sum tax plan is preferred to the uniform premium plan as above.

[46] Andrew S. Whitman and C. Arthur Williams, Jr., "FAIR Plan and Excess Rate Plans in Minnesota," The Journal of Risk and Insurance, Vol. XXXVIII, No. 1 (March, 1971), pp. 43-59.

[47] Andrew S. Whitman and C. Arthur Williams, Jr., "Environmental Hazards in Rating Urban Core Properties," ibid., Vol. XXXVII, No. 3 (September, 1970), pp. 419-436.

[48] An excellent discussion of supply problems not covered in this paper are in:

a. Robert S. Allen, "The Supply Function of Urban Property Insurance: Comment," and "Author's Reply," The Journal of Risk and Insurance, Vol. XXXVII, No. 3 (September, 1970), pp. 459-464.
b. J. J. Launie, "The Supply Function of Urban Property Insurance," The Journal of Risk and Insurance, Vol. XXXVI, No. 3 (June, 1969), pp. 269-283.

[49] U. S. Congress, Public Law, 90-448, Stat. 476, Title XI (August 1, 1968).

[50] John R. Lewis, "A Critical Review of the Federal Riot Reinsurance System," The Journal of Risk and Insurance, Vol. XXXVIII, No. 1 (March, 1971), pp. 29-42.

[51] More correctly, the insurance market is "imperfectly" competitive. We will use the notion of monopolistic competition as an approximation for our analysis.

The classic references about market morphology are:

a. E. H. Chamberlin, The Theory of Monopolistic Competition (Cambridge: Harvard University Press, 1933).
b. William Fellner, Competition Among the Few: Oligopoly and Similar Market Structures, (1st. ed.; New York: A. A. Knopf, 1949).
c. Joan Robinson, The Economics of Imperfect Competition (London: Macmillan Co., 1933).

Other more recent references are:

a. P. W. S. Andrews, On Competition in Economic Theory (New York: St. Martin's, 1964).
b. William J. Baumol, Business Behavior, Value and Growth (Rev. ed.; New York: Harcourt, Brace & World, 1967).
c. Kalman J. Cohen and Richard M. Cyert, Theory of the Firm (Englewood Cliffs, N. J.: Prentice-Hall, 1965).

We will examine the effects of price regulation upon the monopolistic competition solution. Obviously, there are other types of regulation which could be studied, such as quality of service, marketing practices, cash-reserve requirements, and so forth, which could alter the monopolistic competition solution. A side issue which will arise later in the analysis is that marketing practices such as red-penciling or gray-penciling may emerge as a resultant from price regulation.

The most complete discussions of the effects of price regulation upon market equilibrium can be found in:

a. Rudolph C. Blitz and Millard F. Long, "The Economics of Usury Regulation," Journal of Political Economy, Vol. LXXIII, No. 6 (December, 1965), pp. 608-619.

b. Maurice Goudzwaard, "Price Ceilings and Credit Rationing," Journal of Finance, Vol. XXIII, No. 1 (March, 1968), pp. 177-185.

c. Maurice Goudzwaard, "Discussion," ibid.; Vol. XXV, No. 2 (May, 1970), pp. 526-528.

[52] We are also assuming that the demand function is independent of the cost function. In a Galbraithian spirit, this is clearly not true. Marketing techniques can alter demand functions among competitors as well as expand the market absolutely. However, we assume independence to simplify our analysis; it is not crucial to our conclusions.

[53] Again insurers must take care to avoid the potential pitfalls of moral hazard. It will be assumed that insurers incorporate the possibilities of moral hazard explicitly in their policy decisions.

[54] The analysis of the two markets case can be expanded to include any number of risk-location separable markets. We could trace out in n-dimensional space partial optimization surfaces, and the effects of regulation, and so forth. However, for our purposes, the two market case is sufficient to demonstrate the potential results from regulation in imperfectly competitive markets.

[55] Meeting the Insurance Crisis of Our Cities, op. cit., particularly Chapter 4.

[56] Whitman and Williams, FAIR Plan ... in Minnesota," op.cit, believe that inspection costs are excessively high, and therefore impractical.

[57] The New York Times, July 7, 1971, p. 21.

This essay attempts to explore the demand for and supply of insurance through the use of simplified mathematical models. The theoretical analysis uses an expected utility-portfolio approach in order to derive implications for static equilibrium in the insurance market. In addition to the theoretical models, there are suggested applications of the theory for urban core insurance problems. Some of the major conclusions of the essay are (1) given the apparent risks involved in urban core areas, the government will probably be called upon to be the insurer or re-insurer for significant quantities of property insurance, (2) governmental price regulation in the insurance market is likely to be a partial cause for inadequate urban core coverage, and (3) Fair Access to Insurance Requirements plans (FAIR) and Excess Rates plans either currently in effect or proposed will be non-optimal and inadequate solutions to urban property insurance problems.

The essay is presented in two principal parts, The Theory of Insurance and The Applications of the Theory of Insurance to Urban Problems. These parts are divided in several main sub-sections:

<u>The Theory of Insurance</u>

I - Optimal Property Insurance for an Individual

II - Expected Utility and Variable Coverage

III - Insurance Company Behavior

<u>The Applications of the Theory of Insurance to Urban Problems</u>

I - The Statement: Urban Core Insurance Problems

II - Elements of the Optimal Solution

III - Private Market Approaches to the Solution

IV - Existing Public Programs as Solutions

The Theory of Insurance

Economic theorists recently have lucubrated about the relationship between asset preferences and expectations regarding future events or future opportunities.[1] Expectations are multi-faceted; a world in which individuals have expectations about future states or conditions generally can not be described by a simple "parametric" index. We are not affected merely by what we suspect is the most probable outcome of a set of possible outcomes, or even by the average outcome (i.e., expected value); our behavior is determined by our expectations about less probable, or even the most improbable of possible outcomes, as witnessed by the public's demand for Irish sweepstakes tickets or certain Canadian uranium mining companies' stock. One of the most important aspects demonstrating the interconnectedness of expectations and asset preferences in the real world is the demand for insurance protection, and the resulting robustness of insurance companies' wealth. Unfortunately, the theory of insurance demand and supply are somewhat neglected in the literature.[2] This paper will explore, through the use of simplified models, the theoretical determinants of the demand for and the supply of insurance, and the implications for static equilibrium in the insurance market. The theory will be applied in subsequent sections to several problems associated with urban areas and insurance arrangements.[3, 4]

Overall, it is evident that none of the insurance panaceas currently in use or proposed here or elsewhere will bring about the desired effects without government intervention. Moreover, the government intervention required will be strenuous and multifaceted; it will tax the consumer, tax and subsidize the insurer, and may set-up several regulatory and inspection agencies.

FOOTNOTES

*Assistant Professor of Finance, The Wharton School, University of Pennsylvania. The author wishes to thank the Rodney L. White Center for Financial Research for financial support and Professors Martin Feldstein and Wassily Leontief for their intellectual support. The author also wishes to thank his colleagues of the Finance Department of the University of Pennsylvania for their innumerable helpful comments. This paper represents a significant revision of an earlier paper of the same title. Of course, any errors are the responsibility of the author.

[1]Three seminal studies investigating the role of expectations in the theory of asset preferences are:

a. John R. Hicks, "A Suggestion for Simplifying the Theory of Money," Economica, New Series, Vol. II (1935), pp. 1-19.
b. Harry Markowitz, "Portfolio Selection," Journal of Finance, Vol. VII, No. 1 (March, 1952), pp. 77-91.
c. James Tobin, "Liquidity Preference as Behavior Towards Risk," Review of Economic Studies, Vol. XXV, Nos. 66, 67 and 68 (1957-1958), pp. 65-86.

[2]Examples of previous analysis related to the theory of insurance are:

a. K. J. Arrow, "Uncertainty and the Economics of Medical Care," American Economic Review, Vol. LIII, No. 5 (December, 1963), pp. 941-973.
b. Karl Borch, "Some Elements of a Theory of Reinsurance," Journal of Insurance, Vol. XXVIII, No. 3 (September, 1961), pp. 35-43.
c. Isaac Ehrlich and Gary S. Becker, "Market Insurance, Self-Insurance and Self-Protection," Center for Mathematical Studies in Business and Economics, Report 7110 (Chicago: Department of Economics and Graduate School of Business, University of Chicago, February 1971), 52 p.
d. John P. Gould, "The Expected Utility Hypothesis and the Selection of Optimal Deductible for a Given Insurance Policy," Journal of Business, Vol XLII, No. 2 (April, 1969), pp. 143-151.
e. Jan Mossin, "Aspects of Rational Insurance Purchasing," Journal of Political Economy, Vol. LXXVI, No. 4 Pt. 1 (July/August, 1968), pp. 553-568.
f. B. Peter Pashigian, Lawrence L. Schkade, and George H. Menefee, "The Selection of an Optimal Deductible for a Given Insurance Policy," Journal of Business, Vol. XXXIX, No. 1, Pt. 1 (January, 1966), pp. 35-44.
g. John Pratt, "Risk Aversion in the Small and in the Large," Econometrica, Vol. XXXII, No. 1-2 (January-April, 1964), pp. 122-136.
h. Vernon L. Smith, "Optimal Insurance Coverage," Journal of Political Economy, Vol. LXXVI, No. 1 (January-February, 1968), pp. 68-77.

the monopolistic competition solution. A side issue which will arise later in the analysis is that marketing practices such as red-penciling or gray-penciling may emerge as a resultant from price regulation.

The most complete discussions of the effects of price regulation upon market equilibrium can be found in:

a. Rudolph C. Blitz and Millard F. Long, "The Economics of Usury Regulation," Journal of Political Economy, Vol. LXXIII, No. 6 (December, 1965), pp. 608-619.

b. Maurice Goudzwaard, "Price Ceilings and Credit Rationing," Journal of Finance, Vol. XXIII, No. 1 (March, 1968), pp. 177-185.

c. Maurice Goudzwaard, "Discussion," ibid.; Vol. XXV, No. 2 (May, 1970), pp. 526-528.

[52] We are also assuming that the demand function is independent of the cost function. In a Galbraithian spirit, this is clearly not true. Marketing techniques can alter demand functions among competitors as well as expand the market absolutely. However, we assume independence to simplify our analysis; it is not crucial to our conclusions.

[53] Again insurers must take care to avoid the potential pitfalls of moral hazard. It will be assumed that insurers incorporate the possibilities of moral hazard explicitly in their policy decisions.

[54] The analysis of the two markets case can be expanded to include any number of risk-location separable markets. We could trace out in n-dimensional space partial optimization surfaces, and the effects of regulation, and so forth. However, for our purposes, the two market case is sufficient to demonstrate the potential results from regulation in imperfectly competitive markets.

[55] Meeting the Insurance Crisis of Our Cities, op. cit., particularly Chapter 4.

[56] Whitman and Williams, FAIR Plan ... in Minnesota," op.cit, believe that inspection costs are excessively high, and therefore impractical.

[57] The New York Times, July 7, 1971, p. 21.