ECO 7997 - M.A. Major Paper

EFFECTS ON SELF-PROTECTION: NO-FAULT vs. FAULT AUTOMOBILE INSURANCE

by

Michael Cheng

Supervisor: Julio Silva-Echenique
Defense held: June 23, 1986

UNIVERSITÉ D'OTTAWA

FACULTÉ DES SCIENCES SOCIALES

DÉPARTEMENT DE SCIENCE ÉCONOMIQUE

UNIVERSITY OF OTTAWA

FACULTY OF SOCIAL SCIENCES

DEPARTMENT OF ECONOMICS

EFFECTS ON SELF-PROTECTION:
NO-FAULT vs. FAULT AUTOMOBILE INSURANCE

University of Ottawa
Department of Economics
ECO 7997
M.A. Major Paper

Michael Cheng 049669

June 1986

I would like to express my appreciation to Professor Julio Silva-Echenique for his valued assistance.

"We will not write liability insurance on a railroad train, running on steel rails, through protected property, and under the control of a skilled engineer who has spent years learning how to do it. Then wouldn't we be nitwits writing such insurance covering an automobile travelling at breakneck speed...driven by anybody who has enough money to buy one of the contraptions, whether or not he has enough brains to operate it properly?"

Insurance Executive[1]

ABSTRACT

A ferocious debate on the implementation of no-fault vs. fault automobile insurance is widespread in the popular press, amongst governments, and in the economics and law literature. On May 6, 1986, a comprehensive, 224-page provincial report, made public by David Slater, task force chairman of the Ministry of Consumer and Commercial Relations, included among its many recommendations a call for the Ontario Government to "consider elimination of resort to the tort or litigation (fault) system with respect to personal injury compensation for automobile accidents", or, failing that, "to consider substantially limiting resort to the ... system."[2] (A tort is essentially a private wrong for which a civil suit can be launched.) This paper will, in particular, attempt to analyze several effects of both systems on the level of self-protection taken by a driver. These include the introduction of liability and insurance with perfect information, a comparison of no-fault to fault plans under imperfect information, and the effects of an independent fine system. The models have been adapted from several papers by Michael Hoy, dealing with the income effects of imperfectly categorizing risks.[3]

BACKGROUND

There has been constant pressure for move to a no-fault system of insurance from the system of tort liability which has traditionally existed in accident law. This pressure is, however, not new. The no-fault solution was applied to workmen's compensation after the turn of this century in Ontario (1914).[4] The first no-fault application to automobile insurance in the English speaking world took place in Saskatchewan in 1946.[5] Today, some degree of no-fault insurance is available in all Canadian provinces and is compulsory except in the Maritimes. Quebec has a pure no-fault plan while other provinces have add-on systems where the tort suit option is available. In 1974, New Zealand introduced the most extensive no-fault system of any common law jurisdiction. It covers not only personal injury in automobile accidents but personal injury in all other accidents as well. Between 1971 and 1976, 16 states in the United States have adopted some degree of no-fault insurance.[6]

Proponents of no-fault system cite the high administrative, legal, and investigative costs associated with the fault system as the major reason for a shift. They adduce that the no-fault approach is more efficient, less costly, and looks after the majority of people faster since it generally results in more generous awards and quicker settlements as well as reducing the uncertainty about what dollar value would be put on an injury by the courts.[7] On the other hand, opponents claim that the introduction of no-fault insurance would lead to a decrease in self-protection exercised by drivers. Because the "victim" and the "negligent party" would be compensated equally for the same injuries, more claims will be paid out on an average basis. This moral hazard type reaction, they argue, would likely lead to increased accident costs and higher premiums that society would have to absorb.[8] (Moral hazard refers here to the tendency of automobile insurance to alter the probability of an accident occurring when care cannot be observed. The driver whose car is fully insured takes less precautions than otherwise.)

SELF-PROTECTION AND LIABILITY RULE

Drivers of automobiles face a choice in the level of self-protection during the operation of their vehicle. This choice relates, in some way, their cost of obtaining self-protection to their personal benefits received from increased safety. The level of self-protection chosen is represented by the variable, S. Increasing values of S imply higher levels of safety precautions. Obviously, self-protection is not costless. Resources must be devoted to the provision of care. The costs of care can be easily imagined. They range from the costs associated with regular checks for mechanical fitness, and the need to drive slower and forego other uses of time, to the costs of increasing concentration by neither listening to the radio nor engaging in conversation with fellow passengers. It is likely that the first units of safety can be acquired cheaply while more extreme measures would tend to be increasingly expensive. Given this assumption, we can define the cost function of self-protection as $C(S)$, where $C' > 0$ and $C'' > 0$. $C(S)$ is non-negative.

We must bear in mind that accidents do not take place with certainty. The probability of an accident occurring should depend on the level of self-protection taken by a driver. Intuitively, an increase in self-protective measures should lead to a decrease in the probability of an accident occurring. It seems likely that the law of diminishing returns would apply to this situation. Even for extremely high levels of self-protection, there should still exist some slight probability for an accident to take place. This probability function is represented by P(S), where

$$P' < 0, \; P'' < 0, \text{ and } 0 < P < 1.$$

Suppose for now that all individuals are identical. Let L^o and W^o denote respectively the personal loss incurred by a driver involved in an accident and the value of his initial wealth. It is assumed that $W^o > L^o$ in order to remove any distracting results due to limited liability. Limited liability occurs because no more than the total wealth of a person can be taken away from him. The individual with $L^o < W^o$ faces a smaller income loss than the person with $L^o > W^o$ and may therefore take a less than optimal amount of self-protection. In addition, L^o is a constant since we do not allow defensive driving in our model. A driver involved in an accident is either the "victim" or the "negligent party". We assume that he has no control over his probability of becoming the "victim". Injury by another driver is deemed purely random. This type of loss would then be lump-sum and has no influence whatsoever on self-protection levels.

There are 2 states of the world: the accident and the no-accident state. Their corresponding payoffs are denoted as follows.

$$W_A = W^o - L^o - C(S) \qquad \text{(accident)} \qquad (1)$$

$$W_N = W^o - C(S) \qquad \text{(no-accident)} \qquad (2)$$

Which state will occur is unknown in advance but their worth to the individual is certain. von Neumann and Morgenstern (1947) provided a way of explaining the individual's attitude to risk in terms of a utility function. In our model, we assume that the individual can attach probabilities (subjective) to each of the two states. He receives W_A with probability P and W_N with probability (1-P). It is assumed that the individual's preferences can be ordered and he maximizes expected utility, EU.

$$\max_S EU = P(S)U(W_A) + (1-P(S))U(W_N)$$

$$= P(S)U(W^o-L^o-C(S)) + (1-P(S))U(W^o-C(S)) \qquad (3)$$

The utility function satisfies the property of continuity and independence. It is unique up to an affine transformation.[9] (An affine transformation preserves the expected utility property, $U(PW_A+(1-P)W_N) = PU(W_A)+(1-P)U(W_N)$.) The individual is assumed to be risk averse. Decreasing marginal utility of income follows from this assumption.

By setting $dEU/dS = 0$, we obtain the first order condition.

$$\frac{dEU}{dS} = \frac{dP}{dS} U(W_A) - \frac{dP}{dS} U(W_N) - P(S) \frac{dU}{dW_A} \frac{dC}{dS} - (1-P(S))\frac{dU}{dW_N} \frac{dC}{dS} = 0$$

Since $W_A = W_N - L^o$ from (1) and (2),

$$-\frac{dP}{dS}\left[U(W_N)-U(W_N-L^o)\right] = \left[\frac{dU}{dW_N} + P(S)\left(\frac{dU}{dW_A} - \frac{dU}{dW_N}\right)\right] \frac{dC}{dS} \qquad (4)$$

$$MB_1 = MC_1$$

The individual chooses S at which marginal benefit, MB_1, equals marginal cost, MC_1. MB_1 is positive. It is the negative of the marginal decrease in expected utility associated with the provision of additional self-protection. MC_1, on the other hand, is composed of two positive terms. The first term is the marginal utility of the marginal cost of self-protection. The second term represents the additional marginal cost that results from incurring self-protection in the accident state.

However, L^o represents only the private cost of an accident. It is normally expected that external costs will be imposed by the "negligent party" on the "victim" in the form of property damage and bodily harm when an accident occurs. As with all externalities, the economically efficient result will be obtained if these costs are internalized. This is the basic economic idea behind tort law. The assignment of liability would discourage those activities which impose social costs greater than their private benefit. In this model, there is no comparative negligence, a system wherein drivers can be found liable for a portion of damages. The "negligent party" is now held responsible for the entire loss to society, as opposed to only his private loss previously. This ensures that the efficient amount of self-protection will be taken. Let L represent the total loss where $L > L^o$ and $L < W^o$. The latter condition is placed to avoid the limited liability problem. Notice that L and L^o are fixed and independent of S.

The individual maximizes expected utility while taking into account the larger loss.

$$\max_S EU = P(S)U(W_A) + (1-P(S))U(W_N)$$

$$= P(S)\, U(W^o-L-C(S)) + (1-P(S))\, U(W^o-C(S)) \qquad (5)$$

The new first order condition is the following.

$$- \frac{dP}{dS}\left[U(W_N)-U(W_N-L)\right] = \left[\frac{dU}{dW_N} + P(S)\left(\frac{dU}{dW_A} - \frac{dU}{dW_N}\right)\right] \frac{dC}{dS} \qquad (6)$$

$$MB_2 = MC_2$$

To compare the levels of self-protection taken under no-liability and liability, we need to consider the changes incurred in both the MB and MC curves. Suppose that they are linear and MB and MC are non-negative. Under a liability rule, MB increases due to a larger total loss, L ($L > L^o$). MC, on the other hand, also increases. Due to a larger loss, L, W_A falls. There is an increase in the value of dU/dW_A found in the second term of MC. This follows the assumption of decreasing marginal utility of income. In order to determine the different magnitudes of the increases in MB and MC, it is necessary to analyze the nature of these increases. We know that the rise in MC can be attributed to decreasing marginal utility of income. But decreasing marginal utility also leads to a similar increase in MB over the interval $L-L^0$. Moreover, MB increases from the growth of possible total loss ($L > L^0$).This would occur even if the utility function were linear. Therefore, the rise in MB exceeds the rise in MC. This results in a greater amount of self-protection taken by the individual (See Figure 1).

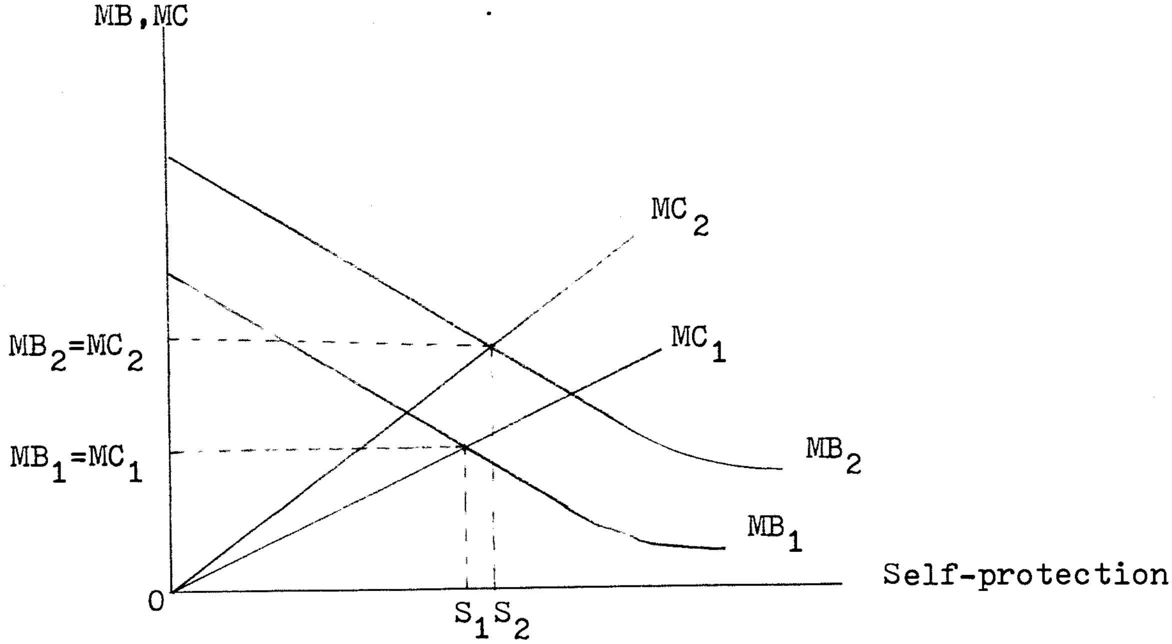

Figure 1. The level of self-protection rises with liability.

Finally, there is a further effect to consider. The increase in the total cost of self-protection, which results from the higher level of self-protection, has lowered all levels of ex-post wealth. The individual now falls into a region of higher marginal utility of income. This will lead to a further increase in both MB and MC. However, these increases are offsetting so that there is no further effect on the level of self-protection. In short, self-protection remains at a higher level under a liability rule.

SELF-PROTECTION AND RISK TYPES IDENTIFICATION

Suppose that there are 2 risk groups: the high risk (HR) and the low risk (LR) group. The high risk group tends to be more accident-prone than its low risk counterparts. This difference is due to some different combination of physical and mental capabilities. As a result, the high risk group faces higher costs of self-protection. To achieve the same level of self-protection, an individual from the high risk group may need to turn off the radio while a low risk individual does not. In addition, the high risk individual may need to drive more slowly due to poor reflexes. Another explanation could be the higher opportunity cost of time for members of the high risk group than those of the low risk group. In any case, the basic difference between the 2 groups is that for any additional unit of self-protection desired, the cost is less to members of the low risk group.

The cost functions of both risk types can be represented as $C_{HR}(S) \geq 0$ and $C_{LR}(S) \geq 0$, where $C'_{HR} > 0$, $C'_{LR} > 0$, $C''_{HR} > 0$, $C''_{LR} > 0$, and $C'_{HR}(\bar{S}) > C'_{LR}(\bar{S})$. (At any given level of S, for example, $\bar{S}$, $C'_{HR} > C'_{LR}$.)

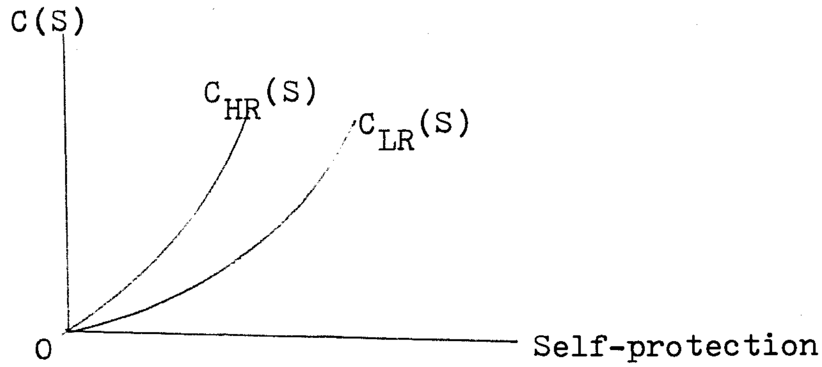

Figure 2. The high risk group faces higher costs of self-protection.

Due to the different cost conditions of safety for the 2 groups, there is no longer one single optimal level of self-protection. The marginal costs of both groups must be substituted into equation (6) individually to find the appropriate level of self-protection. Assume that this has been solved for the low risk group. The introduction of the high risk group's marginal cost at the low risk group's equilibrium level of self-protection would result in an inequality. MC_{HR} would exceed MB. This is because C'_{HR} is greater than C'_{LR} at a given S. A reduction in the level of self-protection taken by the high risk group would result in equilibrium since dC/dS would fall and -dP/dS would rise. It is, however, not clear which group will spend more on self-protection. While the high risk group chooses less self-protection, it has to pay a higher price per unit than the low risk group. Nonetheless, it is assumed that the difference between their total cost of care is negligible and will thus be ignored. Both groups take the efficient amount of self-protection by equating social MB to the relevant individual MC. Since both groups face the same probability-of-an-accident-occurring function, members of the high risk type have a higher P in equilibrium. This is due to the choice of a smaller amount of self-protection than the low risk group (See Figure 3).

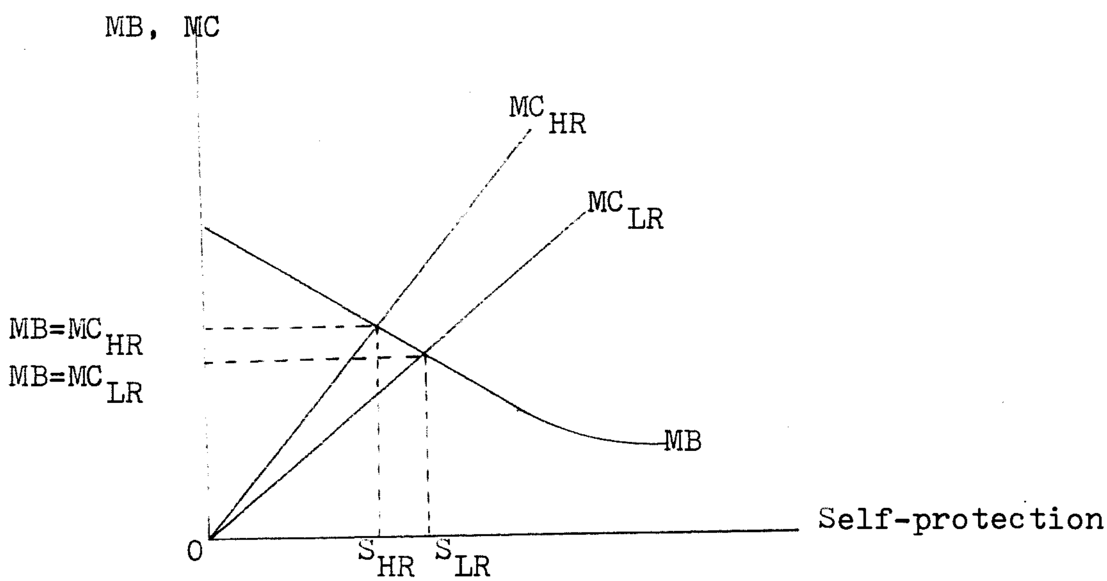

Figure 3. The high risk group chooses a lower level of self-protection.

INSURANCE WITH PERFECT INFORMATION

We can now turn our attention to insurance policy. It will be shown that even the introduction of insurance with perfect information would have an effect on the level of self-protection taken by the 2 risk groups.

Suppose that the level of self-protection can be observed costlessly. This allows an individual to alter the premium he pays by changing his level of self-protection. The individual may also choose the amount of insurance coverage he desires in a manner which maximizes his utility. In this situation, the individual faces 2 choice variables: the level of self-protection and the amount of coverage.

Firms, on the other hand, are assumed to be risk-neutral and perfectly competitive. Under perfect information, they are able to distinguish between the risk types of their clients costlessly. Given constant returns to scale, insurance will be sold at the actuarially fair price so that risk-neutral insurers make expected zero profit. The zero profit condition can be expressed as follows.

$$\beta(S)\,(1-P(S)) = \alpha(S)\,P(S) \qquad (7)$$

$\beta(S)$ and $\alpha(S)$ represent respectively the premium paid by an individual in the no-accident state and the compensation received in the accident state. Payoffs in the 2 states are the following.

$$W_N = W^o - \beta(S) - C(S) \qquad \text{(no-accident)} \qquad (8)$$

$$W_A = W^o - L + \alpha(S) - C(S) \qquad \text{(accident)} \qquad (9)$$

From (7), $$\beta(S) = \frac{\alpha(S)P(S)}{1-P(S)} \qquad (10)$$

Substitute (10) into (8),

$$W_N = W^o - \frac{\alpha(S)P(S)}{1-P(S)} - C(S) \qquad (11)$$

The individual's problem is to maximize expected utility, EU.

$$\begin{aligned} \max_{S,\alpha} EU &= P(S)U(W_A) + (1-P(S))U(W_N) \\ &= P(S)U(W^o-L+\alpha(S)-C(S)) + \\ &\quad (1-P(S))U\left[W^o - \frac{\alpha(S)P(S)}{1-P(S)} - C(S)\right] \end{aligned} \qquad (12)$$

$$\begin{aligned} \frac{\partial EU}{\partial S} &= \frac{dP}{dS}U(W_A) + P(S)\frac{dU}{dW_A}\left[\frac{d\alpha}{dS} - \frac{dC}{dS}\right] \\ &\quad + (1-P(S))\frac{dU}{dW_N}\left[\frac{-(d\alpha/dS)P(S)(1-P(S)) - (dP/dS)\alpha(S)}{(1-P(S))^2} - \frac{dC}{dS}\right] \\ &\quad - \frac{dP}{dS}U(W_N) \\ &= 0 \end{aligned}$$

$$-\frac{dP}{dS}\left[U(W_N)-U(W_A)\right] + P(S)\ \frac{dC}{dS}\left[\frac{dU}{dW_N} - \frac{dU}{dW_A}\right] - P(S)\ \frac{d\alpha}{dS}\left[\frac{dU}{dW_N} - \frac{dU}{dW_A}\right]$$

$$-\frac{dU}{dW_N}\ \frac{(dP/dS)\alpha(S)}{1-P(S)} - \frac{dU}{dW_N}\ \frac{dC}{dS} = 0 \qquad (13)$$

$$\frac{\partial EU}{\partial \alpha} = P(S)\ \frac{dU}{dW_A} + (1-P(S))\ \frac{dU}{dW_N}\ \frac{-P(S)}{1-P(S)} = 0$$

$$P(S)\left[\frac{dU}{dW_A} - \frac{dU}{dW_N}\right] = 0$$

By assumption, $P(S) > 0$

Therefore, $\frac{dU}{dW_A} = \frac{dU}{dW_N}$ (14)

Since the individual is assumed to be risk averse and marginal utility of income is strictly decreasing, (14) implies that $W_A = W_N$ and full insurance is optimal.

Therefore, $U(W_A) = U(W_N)$ (15)

Substitute (14) and (15) into (13),

$$-\frac{dU}{dW_N}\ \frac{(dP/dS)\ \alpha(S)}{1-P(S)} - \frac{dU}{dW_N}\ \frac{dC}{dS} = 0$$

$$-\frac{dP}{dS}\ \frac{\alpha(S)}{1-P(S)} = \frac{dC}{dS} \qquad (16)$$

With full insurance,

$$\beta = L - \alpha \qquad (17)$$

Substitute (17) into (7).

$$(L - \alpha)(1-P) = P$$

$$L(1-P) - \alpha = 0$$

$$L = \frac{\alpha}{1-P} \qquad (18)$$

Substitute (18) into (16).

$$-\frac{dP}{dS}\frac{dU}{dW_N} L = \frac{dU}{dW_N}\frac{dC}{dS}$$

$$-\frac{dP}{dS} L = \frac{dC}{dS} \qquad (19)$$

$$MB_3 = MC_3$$

The above result differs from that obtained from the no-insurance case. Here, the individual maximizes income rather than utility. According to Shavell, insurance leads to a decrease in the level of self-protection taken. When an individual is risk averse, it is desirable for him to avoid risk. The gain from his risk avoiding will be greater the more his actual level of effort serves as an incentive in the premium.[10] Therefore, even with a drop in the level of self-protection, this results in the first best solution and society is better off.

For this drop in the level of self-protection to occur, there must have been some changes in the relative positions of the MC and MB curves. Recall equation (6). The absence of the second term of MC_2 in MC_3 contributes a slight downward movement to MC. The disappearance of the utility terms from MB and the first term from MC must also play a role. The first term of MC_2 is influenced by dU/dW_N . MB_2 includes the utility term $U(W_N)-U(W_N-L)$. Utility would have a larger value in this term as each point over the interval L is affected by a decrease in marginal utility greater than dU/dW_N due to decreasing marginal utility of income. With the removal of these utility terms, MB would decrease by more than that of MC and a lower amount of self-protection is chosen (See Figure 4).

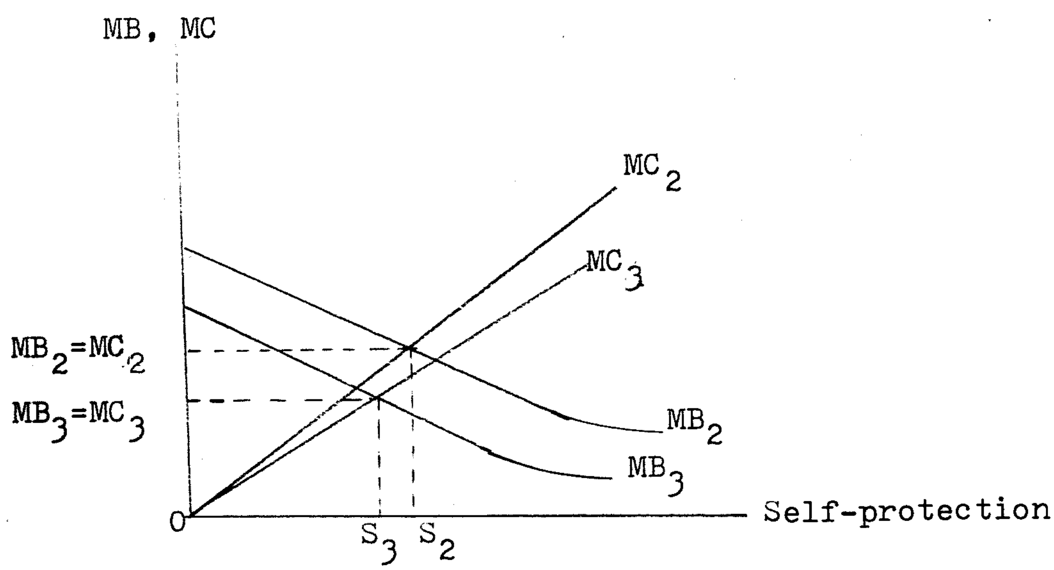

Figure 4. Insurance leads to a fall in self-protection.

Both risk types are expected to take different levels of self-protection due to different marginal costs. Figure 5 depicts the individual's opportunity locus, where his wealth in the no-accident state is measured along the horizontal axis and his wealth in the accident state on the vertical axis. E represents the endowment (no-insurance) point. It is assumed that firms operate with expected zero profit and zero administrative costs. From the zero profit condition (7), the slope of the fair odds line for each risk group is determined to be (1-P)/P. U_{HR} and U_{LR} denote the indifference curves for HR and LR. Under perfect information, both groups could be identified and their utility is maximized by purchasing full insurance at α_{HR} and α_{LR}.

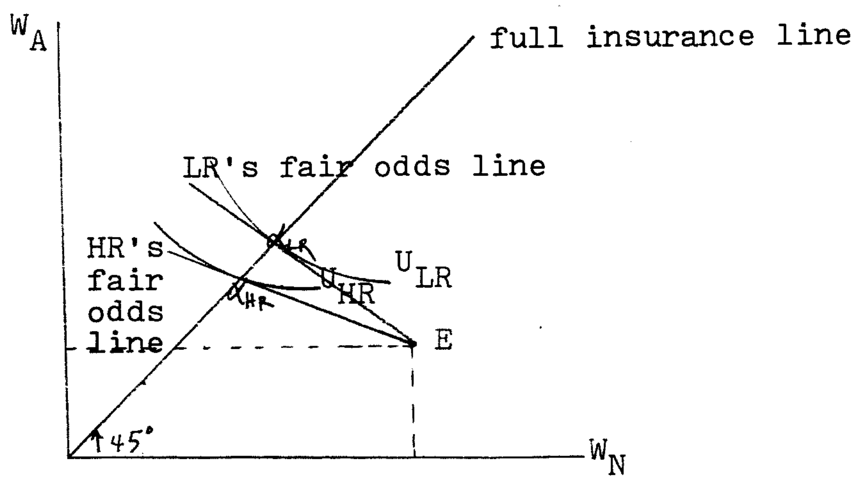

Figure 5. Under perfect information, risk types purchase full insurance at their actuarially fair prices.

VOLUNTARY INSURANCE WITH ASYMMETRIC INFORMATION

Insurance using perfect information does not exist in reality due to high costs of collecting information. Self-selection, a process whereby various risk types identify themselves, may be used. The use of imperfect information enables firms to correlate some observable personal characteristics, for example, age and sex, with various risk types. Nevertheless, the problem of adverse selection limits the ability of insurance companies to offer efficient contracts to the low risk group. (Adverse selection occurs when the risk group that purchases insurance differs adversely from that anticipated by the insurer.) If a firm offers a contract to low risk individuals at their fair rate, high risk individuals will also purchase it so that the firm will make expected losses.[11]

In Figure 6, it is demonstrated that the high risk group may purchase full insurance at the higher fair HR rate. For the low risk group, only partial coverage may be bought at the lower fair LR rate. In addition, partial coverage can be offered to Q, slightly below the indifference curve U_{HR}^{Max}. Obviously, the high risk group would still prefer full coverage at the fair HR rate. If Q were to be located above U_{HR}^{Max}, members of the high risk group would prefer less than full coverage at the fair LR rate. Consequently, the fair LR rate could no longer be offered without incurring a loss. This is because expected losses would exceed premiums paid when individuals from the high risk group purchase insurance at the fair LR rate. In general, prices per unit of coverage increase with the amount of coverage purchased. This is how a system of voluntary insurance would work.

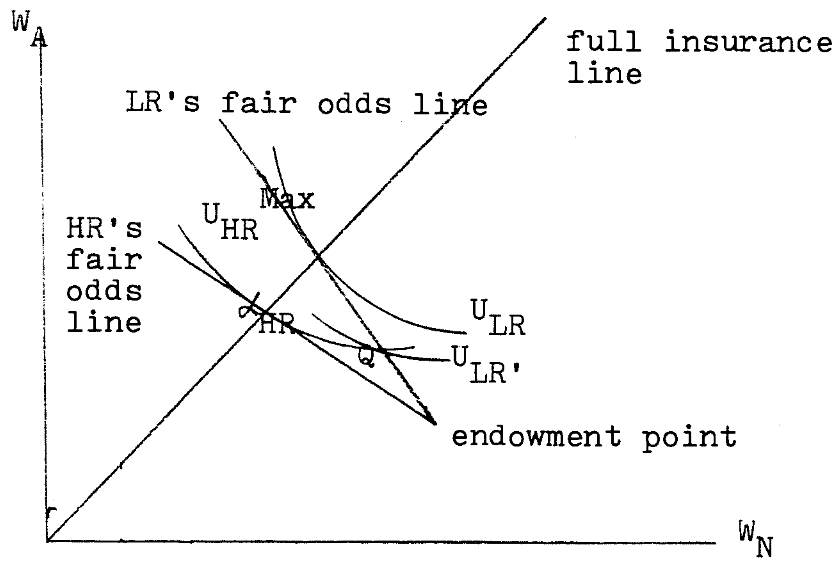

Figure 6. Voluntary insurance gives rise to a competitive equilibrium that is characterized by high risk individuals receiving full coverage at their fair price and low risk individuals receiving partial coverage at their fair price.

COMPULSORY INSURANCE WITH ASYMMETRIC INFORMATION: NO-FAULT vs. FAULT

Automobile insurance is, however, not generally voluntary. In all Canadian provinces and territories, third-party liability insurance is mandatory up to certain levels of coverage. These range from $50,000 to $200,000. Most provinces require first-party benefits and a few provinces require collision as well.[12]

There are several reasons to explain why mandatory insurance has come about. In any society, there are risk lovers and risk-neutral individuals who will not buy insurance. There exist also many people in a position of limited liability. With little to lose personally, they may not purchase insurance. Returning to the model, it is therefore possible that the level of coverage available to members of the low risk group under the two-priced scheme is negligible. If a large loss is suffered by any of the previously-mentioned groups, they and their victims could be reduced to the poverty level so that they are dependent on the state for aid. In this context, the move to compulsory insurance can be seen as an attempt by the government to internalize part of the cost of automobile use. Such a move would have a strong appeal to risk-averse individuals, preventing their chances of becoming uncompensated victims of uninsured drivers. Compensation for all seems to be the goal of mandatory insurance.

The first type of mandatory insurance to be considered is a one-rate-for-all pure no-fault system. Insurance will be offered at the zero-profit pooled rate to both risk groups. In other words, firms charge each individual a price that is the weighted average of the fair prices for both risk groups.

$$(\alpha-2d)\left[\theta P(S_{LR})+(1-\theta)P(S_{HR})\right] = \beta\left[\theta(1-P(S_{LR}))+(1-\theta)(1-P(S_{HR}))\right] \quad (20$$

θ represents the proportion of the population that belongs to the low risk group and $(1-\theta)$ is the proportion belonging to the high risk group. β is the premium paid in the no-accident state and α is the compensation distributed in the accident state to both the victim and the negligent party. The term d represents a deductible amount. Since we assume that each accident has one victim and one negligent party, 2d appears in the zero profit condition (20). As no distinction is made between the victim and the negligent party under a no-fault scheme, both parties must make a deduction to their insurance company after an accident. This amount paid by the victim is a lump sum because we assume that he cannot take action to prevent the accident from occurring. The insurance company is assumed to know θ and have perfect foresight which enables it to predict the equilibrium values of S_{HR} and S_{LR}. These assumptions eliminate the need for gradual adjustment to equilibrium over several periods.

In Figure 7, the mandatory insurance is covered with a deduction of compulsory size. Members of both risk group must purchase the same amount of coverage at the zero-profit pooled rate. The size of the deduction must be made mandatory to ensure this. The high risk group prefers full coverage at B while the low risk group prefers less than full coverage at some point, D. Since losses are of a fixed size, the distance DB may be considered as a deduction. Depending on the nature of the low risk type's indifference curves and the slope of the pooled rate, these individuals may prefer a deduction that reduces the level of coverage to zero. Besides ensuring zero profit, the mandatory deduction under this scheme also plays a role in the determination of the level of self-protection. This would be set at some point between D and B so that the deduction would be small relative to the loss and the goal of compensation is met.

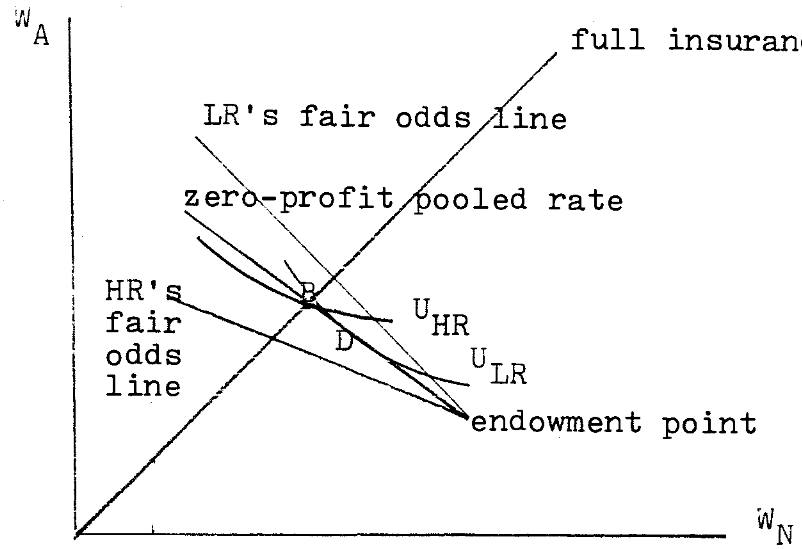

Figure 7. At the actuarially pooled price, both risk groups are offered a single contract.

The industry's zero profit condition will hold if firms correctly estimate S_{LR} and S_{HR} as it has been assumed. For the individual, however, the fair insurance rate, $(\alpha - d)P(S) = \beta(1-P(S))$, is no longer valid. The premium, β, is based on the average care taken by all individuals and is no longer a function of the level of self-protection taken by one single individual. Payoffs in the 2 states are as follows.

$$W_N = W^o - \beta - C(S) \qquad \text{(no-accident)} \qquad (21)$$

$$W_A = W^o - L^o + \alpha^o - C(S) \qquad \text{(accident)} \qquad (22)$$

Since an individual is never found at fault, he cannot be held responsible for the entire loss, L. He suffers personal loss, L^o, and receives a compensation, α^o ($L^o = \alpha^o + d + \beta$). The firm also pays ($L - L^o - d$) to the victim in order to compensate his loss.

(22) may now be rewritten as follows.

$$W_A = W^o - \beta - d - C(S) \qquad (23)$$

The individual maximizes expected utility with respect to S and sets dEU/dS = 0.

$$\max_S EU = P(S)U(W_A) + (1-P(S))U(W_N)$$

$$= P(S)U(W^o-\beta-d-C(S)) + (1-P(S))U(W^o-\beta-C(S)) \qquad (24)$$

$$\frac{dEU}{dS} = \frac{dP}{dS}\left[U(W_A)-U(W_N)\right] - \frac{dU}{dW_N}\frac{dC}{dS} - P(S)\frac{dC}{dS}\left[\frac{dU}{dW_A} - \frac{dU}{dW_N}\right] = 0$$

$$-\frac{dP}{dS}\left[U(W_N)-U(W_N-d)\right] = \left[\frac{dU}{dW_N} + P(S)\left(\frac{dU}{dW_A} - \frac{dU}{dW_N}\right)\right]\frac{dC}{dS} \qquad (25)$$

$$MB_4 = MC_4$$

Drivers equate their MB with MC and arrive at the equilibrium level of self-protection. Due to higher marginal cost, the high risk group exercises a lower level of self-protection than the low risk group. As a result, they will be involved in more accidents.

Examining the first order condition (25), we notice that as d approaches 0, the level of self-protection will fall. When d equals 0, MB and dC/dS must also equal 0. The only point where this may occur is at the level of zero self-protection. If a decreasing d leads to lower level of self-protection, it is then safe to assume that an increasing d would lead to higher level of self-protection.

The situation in which the driver takes no care is the result of the theoretical construction of this model. Compensation is considered perfect. When the deduction equals 0, the driver is indifferent between having an accident and not having one. In the real world, however, deductible policies are offered ($d > 0$) so that drivers cannot be fully insured. This encourages self-protection although it is not clear that a deductible policy gives the appropriate incentives to offer some optimal level of care. In addition, no-fault systems do not normally compensate as fully as fault systems do. In the case of physical impairment and non-pecuniary losses, awards are generally less than those received under fault schemes.[13] Upper limits are often put on the amount of this compensation. In fact, this is equivalent to an increase in the deduction. Lower awards under a no-fault plan may be intended to ensure greater self-protection.

The introduction of fault into insurance policy should lead to an increase in the level of self-protection. According to Shavell, even a system under imperfect information would be valuable to society although the magnitude of its value may vary.[14] While one wonders whether the implementation of no-fault insurance does lead to a decrease in self-protection levels, this backward approach has been taken for the sake of clarity in the development of the model.

In the real world, insurance companies are not able to measure self-protection directly. They use some imperfect indicators that are correlated with the individual's choice of self-protection. These include residential location, use of vehicle, age, sex, and marital status. However, such classifications have little effect on the level of self-protection exercised. Instead, they are more concerned with the equitable distribution of losses.[15] Driving record is another important indicator used in the assignment of rates and this should have an effect on self-protection.

For the purpose of our analysis, firms use driving record as a basis for determining rates. There are 2 groups of individuals, namely HR and LR, but they cannot be identified. It is known that the high risk individuals face a higher probability of having an accident than their low risk counterparts. As a result, there would be some correlation between those who are high risk and those who have poor driving records. This can be used as the classification category.

A 2-period approach is introduced as follows. Those individuals who have accidents during the first period will be placed in the high rate category, H, for the next period. Those who do not have an accident will be placed in the low rate group, L. Both groups are charged different premiums based on the probability of the average individual in the respective rate group's probability of having an accident so that the zero profit condition is met. The breakdown of the rate groups can be found by Bayesian methods.[16] For the high rate category, we have the following.

$$P(S^*_{LR})\,\theta = \lambda_{LR,H} \qquad (26)$$

$$P(S^*_{HR})(1-\theta) = \lambda_{HR,H} \qquad (27)$$

$$\lambda_{LR,H} + \lambda_{HR,H} = \lambda_H \qquad (28)$$

For the low rate category, we have the following.

$$(1-P(S^*_{LR}))\,\theta = \lambda_{LR,L} \qquad (29)$$

$$(1-P(S^*_{HR}))(1-\theta) = \lambda_{HR,L} \qquad (30)$$

$$\lambda_{LR,L} + \lambda_{HR,L} = \lambda_L \qquad (31)$$

λ represents the size of each group. S^*_{LR} and S^*_{HR} are the respective amount of care taken in the first period by the risk groups. They know the premiums the different rate groups would face under this scheme.

The following zero profit rate is charged.

High Rate:

$$(\alpha - d)\left[\frac{\lambda_{LR,H}}{\lambda_H} P(S^*_{LR}) + \frac{\lambda_{HR,H}}{\lambda_H} P(S^*_{HR})\right]$$

$$= \beta_H\left[\frac{\lambda_{LR,H}}{\lambda_H} (1-P(S^*_{LR})) + \frac{\lambda_{HR,H}}{\lambda_H} (1-P(S^*_{HR}))\right] \qquad (32)$$

Low Rate:

$$(\alpha - d)\left[\frac{\lambda_{LR,L}}{\lambda_L} P(S^*_{LR}) + \frac{\lambda_{HR,L}}{\lambda_L} P(S^*_{HR})\right]$$

$$= \beta_L\left[\frac{\lambda_{LR,L}}{\lambda_L} (1-P(S^*_{LR})) + \frac{\lambda_{HR,L}}{\lambda_L} (1-P(S^*_{HR}))\right] \qquad (33)$$

Premium β_H is always higher than β_L because no matter what percentage of high risk individuals exists in the population, high risk types will compose a larger proportion of the high rate group than they do of the low rate group. To avoid several periods of adjustment, it is assumed that firms can predict equilibrium premium levels.

In the initial period, drivers cannot influence the premium they face. This has been determined by the outcome of the previous period. However, individuals can influence the likelihood of their having to pay the higher premium, β_H, in the next period. Drivers with poor records will be penalized by higher premiums and they can avoid this by adjusting their levels of self-protection. Wealth in the 2 states can be expressed as follows.

$$W_N = W^o - \beta_i - C(S) \qquad \text{(no-accident)} \qquad (34)$$

$$W_A = W^o - \beta_i - d - \frac{\beta_H - \beta_L}{1+r} - C(S) \qquad \text{(accident)} \qquad (35)$$

β_i , i = H,L.

Insurance with a set deduction is still mandatory. The new term, $\frac{\beta_H - \beta_L}{1+r}$, represents the discounted loss resulting from the fact that individuals pay the higher premium in the period following an accident. The individual maximizes expected utility with respect to his choice variable, S.

$$\begin{aligned} \max_S EU &= P(S)U(W_A) + (1-P(S))U(W_N) \\ &= P(S)U(W^o-\beta_i-d-(\beta_H-\beta_L)/(1+r)-C(S)) \\ &\quad + (1-P(S))U(W^o-\beta_i-C(S)) \qquad (36) \end{aligned}$$

The first order condition is obtained by setting dEU/dS = 0.

$$-\frac{dP}{dS}\left[U(W_N)-U(W_N-d-\frac{\beta_H-\beta_L}{1+r})\right] = \left[\frac{dU}{dW_N} + P(S)(\frac{dU}{dW_A} - \frac{dU}{dW_N})\right]\frac{dC}{dS} \qquad (37)$$

$$MB_5 = MC_5$$

We can compare the above result with the no-fault case. The addition of the term, $\frac{\beta_H-\beta_L}{1+r}$, to the left hand side of the equation has an effect equivalent to increasing the size of d. As noted earlier, this leads to an increase in the level of self-protection. MB increases with the addition of this term. On the MC side, there is a small increase due to the decrease in the wealth of the accident state. This occurs in the second term of MC. dU/dW_A has a larger value due to decreasing marginal utility of income, which has also led to a similar increase in MB. However, MB rises further due to the addition of the term, $\frac{\beta_H-\beta_L}{1+r}$. This would occur even if the utility function were linear. Therefore, the increase in MC should be small relative to the increase in MB.

As a result, the levels of self-protection taken by both risk groups would increase, leading to higher total costs of self-protection. This, along with the assignment of new premium rates, would cause changes in all levels of ex-post wealth. However, MB and

MC would be altered by the same amount and there should be no effect on the individual's choice of self-protection. This implies that individuals from the same risk group would choose the same amount of self-protection regardless of the rate group to which they are assigned. Changes in premium would not affect utility in a manner that alters the choices of self-protection.

The fault system does, therefore, provide additional incentives for self-protection. But these incentives come only at the price of high legal and administrative costs. Many critics believe that the substantial costs of maintaining the fault system may outweigh the benefits of decreased losses. Others believe that the abolition of fault system would lead to no change in accidents.[17] To say this, however, one must assume away rationality. It has just been shown that even with the veil of insurance protecting the liable individual, the threat of higher insurance premiums should induce more self-protection.

Still, it may be the case that the costs of administration are greater than reduced losses. The administrative cost per dollar of coverage has been estimated for Blue Cross, health insurance, and automobile insurance. The figures are, respectively, 7 cents, 17 cents, and 56 cents.[18] It is thought that a move to no-fault insurance would reduce administrative cost significantly.

Elizabeth Landes (1982) studied the effects on the level of fatal accidents in those U.S. states which placed restrictions on tort liability, i.e., these states adopted some form of no-fault coverage. Although fatal accidents are ruled out of our model,[19] the results are interesting to note. She found that those states which imposed severe restrictions on tort liability had an increase of 10% to 15% in the number of fatal accidents.[20] The reason for using fatal accidents was that they were always reported. The increase in fatal accidents could be roughly indicative of all accidents.

This study has been criticized by many proponents of no-fault insurance.[21] If 56 cents is the portion of each dollar of insurance coverage used for administrative and legal costs, then, presumably, 44 cents is used for compensation. Using Landes' worst estimate to indicate a 15% increase in all accidents with the move to a high degree of no-fault coverage, there must be a corresponding 15% increase in the compensation paid. Compensation would now require 51 cents on what was previously a dollar's coverage. If administrative and legal costs were to fall by more than 7 cents per dollar of coverage, society may actually become better off since the price of insurance coverage has decreased. But society cannot be better off when insurance does not compensate all losses. Increase in the costs of uninsured losses must also be calculated to make the analysis complete.

While this simple exercise has not been conclusive, it does illustrate the point that administrative costs of the fault system are high. If the move from fault to no-fault automobile insurance were to result in a large decrease in these costs, it would seem likely that only a significant increase in the number of accidents could make society worse off. It also suggests that funds used in the administration of the fault system could be used more productively elsewhere to create incentives for self-protection.

EFFECTS OF A FINE SYSTEM

It is believed by some that the fault system does not provide any incentives of self-protection that could not be provided by other means.[22] One possible method to induce safety is through the introduction of a fine system. Many individuals see the insurance structure's role as purely compensatory and consider the incentive for safety to fall within the domain of criminal law.[23] Such regulation does exist and this fine structure could be considered an upgrading of the prevailing system.

The system of fines to be described here is added onto the no-fault system. There are 3 states of the world: no-accident, fined, and accident. Individuals are not fined in the accident state since the costs involved in determining fault are prohibitive. Wealth in the 3 states are the following. (f represents the fine.)

$$W_N = W^o - \beta - C(S) \qquad \text{(no-accident)} \qquad (38)$$

$$W_F = W^o - \beta - f - C(S) \qquad \text{(fined)} \qquad (39)$$

$$W_A = W^o - \beta - d - C(S) \qquad \text{(accident)} \qquad (40)$$

The probability of being fined is expressed by a function of self-protection similar to the probability-of-an-accident-occurring function: $P_F(S)$, $P'_F < 0$, $P''_F < 0$, and $0 < P_F < 1$. The reason behind this is that S cannot be perfectly measured by an observer, but he can distinguish actions that are related to the cause of accidents and low levels of S. Just as the individual at any level of S faces a probability of having an accident, he will face a probability of detection and being fined by the observer. In addition, the probability-of-being-fined function can be affected by the amount of resources devoted to observation. An increase in expenditure spent on observation would raise the probability of detection at any given level of self-protection.

As before, the individual maximizes expected utility with respect to the choice of self-protection.

$$\max_S EU = P(S)U(W_A) + (1-P(S))\left[P_F(S)U(W_F) + (1-P_F(S))U(W_N)\right] \quad (41)$$

$$= P(S)U(W^o-\beta-d-C(S)) + (1-P(S))\left[P_F(S)U(W^o-\beta-f-C(S)) + (1-P_F(S))U(W^o-\beta-C(S))\right]$$

The first order condition is as follows.

$$-\frac{dP}{dS}\left[U(W_N)-U(W_N-d)-P_F(S)\left[U(W_N)-U(W_N-f)\right]\right]$$

$$-\frac{dP_F}{dS}(1-P(S))\left[U(W_N)-(W_N-f)\right]$$

$$=\left[\frac{dU}{dW_N}+P(S)\left[\frac{dU}{dW_A}-\frac{dU}{dW_N}\right]\right.$$

$$\left.+(P_F(S)-P(S)P_F(S))\left[\frac{dU}{dW_F}-\frac{dU}{dW_N}\right]\right]\frac{dC}{dS} \qquad (42)$$

$$MB_6 = MC_6$$

W_N and W_A can be treated as the same as in the no-fault case, i.e., the familiar terms have the same values. The new terms in MB measure the increased marginal benefit of self-protection due to its role in decreasing the probability of incurring the loss associated with the fine. MC also increases but by less than that of MB. The new term on the MC side represents the increase in the cost of care related to the fined state due to decreasing marginal utility of income. But decreasing marginal utility of income also comes into effect on the MB side. Moreover, MB increases further due to the effect of imposing the fine,f. This would occur even if the utility function were linear. As a result, a fine system induces higher level of self-protection if the fine is greater than zero and P_F exists. As a driver takes more self-protection, there will be adjustments to wealth in all states due to lower insurance premiums and more spending on safety measures. But any changes in wealth would result in offsetting changes in both sides of the equation and would not affect the level of self-protection. In short, the level of self-protection remains higher under the fine system. Due to marginal cost conditions, the high risk group takes less self-protection than the low risk group.

The basic premise of the fine scheme combined with no-fault insurance is that it may induce self-protection at lower costs than the fault system. It is often suggested that low levels of enforcement with high fines is the proper method of instituting such a system.[24] High levels of enforcement are costly while fines are just a transfer and do not represent a loss to society. However, high fines are not without problems. It is not desirable to set fines so high that it encourages people to avoid payment. Higher fines would make individuals devote more resources to their defence in court.[25] Conversely, individuals may have to bear the onus of proof.[26] Acquiring such evidence as well as higher defence costs incur a cost to society.

The above considerations seem to indicate that there would be an upper limit on the magnitudes of fines.

On the other hand, it is possible to increase the probability of detection without increasing costs if new methods and technology are introduced. For example, the breathalizer and the roadside breathalizer had both reduced the cost of detecting impaired drivers, replacing the use of blood tests previously. A future innovation would place responsibility on the owner of a vehicle for traffic violations. Notices of fines could be mailed to owners rather than maintaining the time-consuming procedure of stopping a vehicle and giving a ticket to the driver. It is obvious that there exists many possibilities of designing a fine system to replace the fault scheme as both the probability of detection and the fine can be controlled.

Craig Brown (1984) found in his study of the New Zealand experience with the no-fault system that there was a trend to decreased levels of accidents after the 1974 adoption of a pure no-fault system. He noted many other influences which, over the past decade, might have contributed to these changes. These include changes in driving habits due to fluctuations in oil prices, changes in vehicle design, and changes in highway regulations.[27]

SUMMARY AND CONCLUSION

The question remains: what role does the removal of tort liability play? We know that its removal should reduce the levels of self-protection, but the magnitude of the fault system's influence relative to other factors is not known.

Perhaps, it is safe to say that the move from fault to no-fault insurance cannot be examined in isolation. Other alternatives must also be considered. First, insurance is not complete. Individuals who have been disfigured by accidents would not necessarily trade their health for the cash settlement received.[28] The threat to life and physical defect would be a major source of inducement for the provision of care. Second, there is a fine and punishment structure in existence. This would provide additional incentive for self-protection. Third, another alternative would be a "modified" fault/no-fault system under which law suits can still be brought in the case of serious and permanent injuries.[29]

In conclusion, the influence of the fault system must be looked upon as part of a range of tools available to encourage self-protection. Fault and liability do seem appealing to one's sense of justice as the proper way to eliminate the externalities of automobile accidents, for the guilty party pays for its actions. However, the massive transaction costs associated with the maintenance of this system vs. the no-fault scheme may well exceed the social benefits received. To make our study complete, a thorough cost-benefit analysis of both systems is thus apparent.

ENDNOTES

[1]James A. Wickman, Cars, Drivers, and Accidents: The Environment of Automobile Insurance (Seattle: University of Seattle Press, 1967), p.23.

[2]Regina Hickl-Szabo, "Call for No-Fault Insurance Worries Ontario Lawyers", The Globe and Mail, May 7, 1986.

[3]See Hoy(1981), Hoy(1982), Hoy(1984).

[4]Robert M. Solomon, B.P. Feldthusen, & S.J. Mills, Cases and Materials on the Law of Torts (Toronto: Carswell, 1982), p.606.

[5]Ibid., p.614.

[6]Elizabeth M. Landes, "Insurance, Liability, and Accidents: A Theoretical and Empirical Investigation of the Effects of No-Fault Accidents", Journal of Law and Economics, April 1982, p.49.

[7]Solomon, p.49.

[8]Ibid., p.51.

[9]H.R. Varian, Microeconomic Analysis (London: W.W. Norton, 1984), p.159.

[10]Steven Shavell, "Risk Sharing and Incentives in the Principal and Agent Relationship", The Bell Journal of Economics, Spring 1979, p.64.

[11]M. Rothschild & J. Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", Quarterly Journal of Economics, November 1976, p.296.

[12]Solomon, p.616.

[13]Ibid.,p. 603.

[14]Shavell, p.64.

[15]Craig Brown, "Deterrence and Accident Compensation Schemes", University of Western Ontario Law Review, 111, 1978-79, p.142.

[16]John Hey, Uncertainty in Microeconomics (Oxford: Martin Robertson, 1979), p.58-60.

[17]Solomon, p.49.

[18]Ibid., p.605.

[19]Kenneth J. Arrow, Aspects of the Theory of Risk-Bearing (Helsinki: The Academic Book Store, 1965), p.54.

[20]Landes, p.50.

[21]Jeffrey O'Connell & Saul Levmore, "A Reply to Landes: A Faulty Study of No-Fault's Effect on Fault", Missouri Law Review, vol. 48.

[22]Brown, p.127.

[23]Richard A. Posner, Economic Analysis of Law (Toronto: Little and Brown, 1977), p.157.

[24]Gary S. Becker, "Crime and Punishment: An Economic Approach", Journal of Political Economy, March 1968, p.180.

[25]A.M. Polinsky and S. Shavell, "The Optimal Trade-Off Between Fines and Probability", American Economic Review, December 1979, p.884.

[26]R.C. Cranton, "Driver Behaviour and Legal Sanction: A Study of Deterrence" Michigan Law Review, 67, 1968, p.429.

[27]Craig Brown, "Deterrence in Tort and No-Fault: The New Zealand Experience", Economics and Law Workshop, 84-09, University of Western Ontario, p.40.

[28]Solomon, p.608.

[29]Hickl-Szabo.

REFERENCES

Arrow, Kenneth J., Aspects of the Theory of Risk-Bearing, Helsinki: The Academic Book Store, 1965.

Becker, Gary S., "Crime and Punishment: An Economic Approach", Journal of Political Economy, March 1968.

Brown, Craig, "Deterrence and Accident Compensation Schemes", University of Western Ontario Law Review, 111, 1978-79.

------------- "Deterrence in Tort and No-Fault: The New Zealand Experience", Economics and Law Workshop, University of Western Ontario, 84-09.

Calabresi, G. The Cost of Accidents: A Legal and Economic Analysis, New Haven: Yale University Press, 1970.

Cranton, R.C., "Driver Behaviour and Legal Sanction: A Study of Deterrence", Michigan Law Review, 67, 1968.

Hey, John, Uncertainty in Microeconomics, Oxford: Martin Robertson, 1979.

Hickl-Szabo, Regina, "Call for No-Fault Insurance Worries Ontario Lawyers", The Globe and Mail, May 7, 1986.

Hoy, Michael, "Discrimination and Risk-Rated Insurance", Economics and Law Workshop, University of Western Ontario, 81-26.

------------- "Categorizing Risks in the Insurance Industry", Quarterly Journal of Economics, April 1982.

------------- "The Impact of Imperfectly Categorizing Risks on Income Inequality and Social Welfare", Canadian Journal of Economics, August 1984.

Johnson, Joseph E. & Flanigan, George B., "The Outlook for Insurance Rates Based on Driving Records", Insurance Legal Journal, January 1975.

Landes, Elizabeth M. , "Insurance, Liability, and Accidents: A Theoretical and Empirical Investigation of the Effects of No-Fault Accidents", Journal of Law and Economics, April 1982.

Malinvaud, E., Lectures on Microeconomic Theory, New York: North-Holland, 1972.

O'Connell, Jeffrey & Levmore, Saul, "A Reply to Landes: A Faulty Study of No-Fault's Effect on Fault", Missouri Law Review, vol. 48.

Pauly, M., "The Economics of Moral Hazard: Comment", American Economic Review, June 1968.

Polinsky, A.M. & Shavell, S., "The Optimal Trade-Off Between Fines and Probability", American Economic Review, December 1979.

Posner, Richard A., Economic Analysis of Law, Toronto: Little & Brown, 1977.

Rothschild, M. & Stiglitz, J., "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", Quarterly Journal of Economics, November 1976.

Shavell, Steven, "Risk Sharing and Incentives in the Principal and Agent Relationship", The Bell Journal of Economics, Spring 1979.

---------------- "On Moral Hazard and Insurance", Quarterly Journal of Economics, November 1979.

Solomon, Robert M., Feldthusen, B.P., & Mills, S.J., Cases and Materials on the Law of Torts, Toronto: Carswell, 1982.

Stigler, George J., "Optimum Enforcement of Laws", Journal of Political Economy, June 1970.

Varian, H.R., Microeconomic Analysis, London: W.W. Norton, 1984.

von Neumann, J. & Morgenstern, O., Theory of Games and Economic Behaviour, Princeton University Press, 1947.

Wickman, James A., Cars, Drivers, and Accidents: The Environment of Automobile Insurance, Seattle: University of Seattle Press, 1967.

Wilson, C., "A Model of Insurance Markets with Incomplete Information", Journal of Economic Theory, June 1977.