The Geneva Papers on Risk and Insurance Theory, 16:1 45–58(1991)

Reliability of Risk Management: Market Insurance, Self-Insurance and Self-Protection Reconsidered

ERIC BRIYS
Groupe HEC, Jouy-en-Josas, France

HARRIS SCHLESINGER
University of Alabama, Tuscaloosa, AL 35487

J.-MATTHIAS GRAF v.d. SCHULENBURG
University of Hannover, Hannover, FRG

Abstract

This paper examines the three main tools of risk management in a setting where reliability cannot be guaranteed. Thus, for example, insurers might be insolvent, sprinkler systems might be inoperative and alarm systems might be faulty. These types of nonreliability are shown to have significant consequences for risk management. In particular, the relationships between increased risk aversion and the use of the various risk management tools do not carry over from models with full reliability. Moreover, the well-known result of Ehrlich and Becker, that market insurance and self-insurance are substitutes, is shown to fail in the presence of nonreliability risk.

Key words: Default Risk, Reliability, Risk Aversion, Risk Management.

1. Introduction

In their classic paper, Ehrlich and Becker [1972] examine the use of market insurance, self-insurance (reducing the severity of a loss) and self-protection (reducing the probability of loss) as three main tools for personal risk management.[1] An implicit assumption throughout the theory is the reliability of these risk-management tools. Alarms are assumed to sound when break-ins occur, sprinkler systems operate properly during a fire and insurance companies are always standing by with their checkbooks ready to pay claims as they are filed. Unfortunately, these tools of risk management are also risky in their own right. The systems may fail to respond properly during a time of need. An inoperative sprinkler system may be totally destroyed during a fire, thus not only failing in its primary mission of restricting the amount of damages, but actually adding its own value to the damage total. Of course, a faulty sprinkler system might also set itself into operation even though no fire exists, thus causing damages that would not have occurred in its absence.

In this paper, we examine the use of the three main tools of risk management when reliability is not certain. We assume that the potential nonperformance of the tools is known by the consumer, who assigns a probability distribution to the effectiveness of the tools. We also only consider cases where the tools do not work properly in the event of a loss. Under such circumstances, we show that several existing results, which hold for reliable risk-management tools, no longer hold in the present model. Our point is fairly robust: clear-cut monotonic relationships do not obtain when the effectiveness of the loss-control tools is uncertain. In particular, we examine the relationships between risk aversion and the use of risk-management tools, as well as the joint purchase of market insurance and self insurance.

The nonreliability risk is very real. For example, only eight U.S. Property & Casualty insurance companies became insolvent over the period 1981–1983, whereas fifty-five were declared insolvent during the years 1985–1987 (source: National Association of Insurance Commissioners). Given that there were about 2500 U.S. companies in existence over this latter period, the probability of insolvency for a randomly chosen company during this three-year period exceeds two percent, which we consider to be fairly substantial. Over the same period of time, the number of insurers placed on the National Association of Insurance Commissioners' "immediate-attention" list (as showing strong indications of potential default) has risen from 77 during 1981 to 234 during 1987. In fact, over 22 percent of Property & Casualty insurers in the U.S. were in some jeopardy of insolvency in 1987 according to the NAIC.

Even if an insurer remains solvent, claims might not be paid due to certain contract conditions; or if the claims are paid, they might be excessively delayed, thus reducing their present value and, in effect, acting somewhat like a partial default. Clearly then, the risk of insurer solvency is significant and is not likely to be ignored by the rational consumer.

The inclusion of nonreliability risk has only recently begun to find its way into the risk-management literature. Papers by Tapiero, Kahane and Jacque [1986], Doherty and Schlesinger [1990], and Schlesinger and Schulenburg [1987], all consider the demand for insurance in the presence of default risk. The paper by Tapiero et al. considers aggregate demand for an all-or-nothing (i.e. either full coverage or no coverage) insurance decision while the latter two papers are in the same genre as the current paper; they consider the individual's decision as to the level of insurance to purchase. Our insurance model, in section 2, extends the models of Schlesinger and Schulenburg [1987] and Doherty and Schlesinger [1990] by allowing for a continuous distribution of full vs. partial default. Furthermore, Schlesinger and Schulenburg examine aversion to risk using standard Arrow-Pratt measures. Since nonreliability itself adds a second risk to the wealth prospect of the individual, we consider Ross' stronger measure of risk aversion [Ross, 1981], which is more appropriate in multiple-risk settings. The Schlesinger/Schulenburg and Doherty/Schlesinger results are then seen as special cases of our model. We next turn our attention to risky self-insurance and risky self-protection, which, except for some recent work by Hiebert [1989] has not received any attention in the literature. Our model extends Hiebert's results on risk aversion to use Ross' stronger measure. Furthermore, we provide the intuition for the ensuing result, which derives from second-degree stochastic dominance.

As in the case of market insurance, nonreliability is a real-world problem for other loss control mechanisms. For example, a self-insurance mechanism, such as a sprinkler system, might not operate properly during a fire due to a faulty sensor, a temporary lack of water pressure, a clogged water pipe, and numerous other reasons. In addition to these types of hidden hazards, the sprinkler system would also be inoperative if it were temporarily shut down for routine maintenance. Self-protection mechanisms, such as burglar alarms, are subject to similar types of failures. Given that real-world risk management tools themselves are often risky, it is perhaps not too surprising to find that a more risk averse consumer could possibly decide to use less of these risky tools.

In sections 2–4, we consider the optimal purchases of market insurance, self-insurance and self-protection respectively. We also examine the relationship between consumers' risk aversion and the level of investment in each of the three risk-management tools. As pointed out by Ehrlich and Becker [1972], the "riskiness" of final wealth—which in their model is measured by the variance—is not necessarily reduced by self-protection, although it is reduced by the purchase of reliable market insurance and by reliable self-insurance. Such is not the case, however, when market insurance and self-insurance are not fully reliable, as we show in the paper.

We next turn our attention, in section 5, to the joint purchase of market insurance and other risk-management tools. A key result in Ehrlich and Becker [1972] is that, although market insurance and self-insurance are always substitutes, market insurance and self-protection may be either substitutes or complements. This result is fairly significant because, as pointed out by Ehrlich and Becker [1972, pg. 623], "[it] challenges the notion that 'moral hazard' is an inevitable consequence of market insurance, by showing that under certain conditions the latter may lead to a reduction in the probabilities of hazardous events." In section 5 of the current paper, we extend this result by showing that market insurance and risky self-insurance might also be complements. Thus, a higher level of market insurance does not necessarily preclude the individual from taking more care to reduce the size of a potential loss.

2. Risky market insurance

Consider a risk-averse individual with preferences given by the von Neumann-Morgenstern utility function U(W), where $U' > 0$ and $U'' < 0$. The individual has initial wealth A which is subject to a loss of size L, $0 \leq L \leq A$. The loss occurs with probability q. To protect against the loss, the individual may purchase an insurance policy which, in consideration of a specified insurance premium, contracts the insurer to pay a fraction, α, of the ensuing loss. The individual is free to choose the level of insurance, α, and might or might not be restricted to purchasing no more than one-hundred percent coverage, $\alpha \leq 1$. For simplicity, and in order to focus on the "risky" nature of the insurance, sections 2–4 consider only the case of actuarially-fair insurance premiums (i.e. the premium equals the expected insurance indemnity). Extensions to premiums that include a loading are straightforward and don't seem to add any additional insight [cf. Doherty and Schlesinger, 1990].

The premium for insurance level α is given by

$$P(\alpha) \equiv \alpha q \bar{z} L, \tag{1}$$

where, for now, $\bar{z}$ is assumed to equal one. It follows immediately that the individual's expected-utility-maximizing level of insurance is full coverage, $\alpha = 1$ [cf. Mossin, 1968; and Ehrlich and Becker, 1972].

Unfortunately, insurance contracts are not always fulfilled. Insurers become insolvent; they dispute claims; they delay claims payments, and so on. In each of these situations, the individual does not receive the full indemnity specified in the insurance contract. In the case of delayed payments, at least the present value of the indemnity is reduced. To capture the effect of this potential contract nonfulfilment, which we will call "default," we suppose that the individual actually receives an insurance indemnity of $z\alpha L$ following a loss, where z is a realization of the random variable Z. The support of Z is a subset of [0,1]. A degenerate distribution, in which z always equals one, corresponds to the no-default models one usually sees in the literature. We would expect, in reality, a mixed distribution (continuous/discrete) for Z with probability-mass points at $z = 1$ and $z = 0$.

The model we consider is essentially an extension of Doherty and Schlesinger [1990], who examine the case where the support of Z consists of only two points, zero and one, so that only a total default is possible. For values of z strictly between zero and one in our model, the insurer has only a partial default. We assume full information about the distribution of Z and that this information is taken into account in the insurance premium. The "actuarially-fair" insurance premium is thus given by equation (1), where $\bar{z}$ is now interpreted as the mean of Z.

The individual's expected utility, following the purchase of an insurance contract is given by

$$EU = (1 - q)U(A - P) + qE[U(A - P - L + Z\alpha L)], \tag{2}$$

where E denotes the expectation operator. The first-order condition for maximizing (2) is

$$\frac{dEU}{d\alpha} = -(1 - q)q\bar{z}LU'(W_1) + qE[U'(W_2) \cdot (ZL - q\bar{z}L)] = 0 \tag{3}$$

where

$$W_1 \equiv A - P$$

and

$$W_2 \equiv A - P - L + Z\alpha L.$$

The second-order condition is easily verified. From equation (3), it is easy to show that full coverage is optimal whenever the distribution of Z is degenerate with $z = 1$, thus confirming the earlier results of Mossin [1968].

For any degenerate distribution on the support of Z, except for the case where $z = 0$, it follows from (3) that the optimal level of insurance is given by $\alpha^* = 1/z$. For example, if $z = \frac{1}{2}$, then the insurer always pays exactly one-half of the contracted indemnity payment. Assuming no restrictions on α, the optimal level of insurance coverage in this case is $\alpha^* = 2$. Note that such a contract guarantees "full coverage;" that is, the loss is totally indemnified. For small enough ε, $\varepsilon > 0$, it follows that $\alpha^* \approx 2$ for nondegenerate distributions of Z on support $(\frac{1}{2} + \varepsilon, \frac{1}{2} - \varepsilon)$.

It also is possible that partial coverage is optimal for nondegenerate distributions of Z, $\alpha^* < 1$. For example, in the case where $z \in \{0,1\}$—so that only a total default is possible—Doherty and Schlesinger [1990] show that less-than-full coverage is optimal, $\alpha^* < 1$. In general, we can evaluate equation (3) at $\alpha = 1$ to see whether the optimal α^* should be higher or lower. Doing so, we obtain

$$\left.\frac{dEU}{d\alpha}\right|_{\alpha = 1} = (1 - q)q\bar{z}L\{E[U'(W_2)] - U'(W_1)\} + qLCov[U'(W_2),Z]. \tag{4}$$

The sign of (4) is ambiguous. The first term on the right-hand side is nonnegative while the second term is nonpositive. Thus in general, α^* can be higher or lower than one.

We now turn our attention to changes in the individual's degree of risk aversion. We consider an individual who is strongly more risk averse in the sense of Ross [1981]. The strongly-more-risk-averse preferences can be expressed as

$$V(W) \equiv bU(W) + G(W) \tag{5}$$

where $b > 0$, $G' < 0$, $G'' < 0$ [see Ross, 1981].

Using the optimality of α^* for U, we obtain

$$\left.\frac{dEV}{d\alpha}\right|_{\alpha^*} = q(1 - q)\bar{z}L\{E[G'(W_2)] - G'(W_1)\} + qLCov[G'(W_2),Z]. \tag{6}$$

Once again, we obtain an ambiguous sign. With higher risk aversion, α^* can be either higher or lower. Since Ross' measure of risk aversion also implies that V is more risk averse than U in the usual Arrow-Pratt (A-P) sense, we see that higher levels of A-P risk aversion do not necessarily imply higher levels of insurance, contrary to the usual result (in models that do not take insolvency into account). Schlesinger and Schulenburg [1987], using a model with a two-point support for Z, provide a numerical example to show that a positive relationship does not always exist between risk aversion and the level of insurance.

3. Risky self-insurance

We consider here the same model as in the preceding section, except that we no longer allow for market insurance. Instead, the individual is now able to invest in self-insurance to reduce the severity of a loss (also called the "size" of a loss) should it occur. We let x denote the level of self-insurance and assume that the size of the loss is now given by the differentiable function L(x), where $L' < 0$. We introduce risk into the self-insurance model by assuming that the actual loss, if it occurs, is of size L(zx), where z is a realization of the previously-defined random variable Z. If $z = 0$, for example, the self-insurance is ineffective and the loss severity is the same as would have occurred if no investment in self-insurance had been made. If $z = 1$, the self-insurance works "as advertised."

Since self-insurance activities are costly, the individual stands to lose his or her expenditures on self-insurance activities, in addition to suffering a loss of nonreduced size. Thus, the worst possible fate of the individual is now deteriorated. In essence, the risky self-insurance attacks one risk, but creates another; namely, the risk of "wasting good money" on self-insurance goods and services that don't work. Since the real test of workability comes only during the loss experience, the individual cannot be certain whether or not the self-insurance will be effective until a loss is experienced. A more risk-averse individual may conceivably decide to reduce the investment in self-insurance, so as to improve the worst possible fate. Indeed, even if the cost of nonreliable self-insurance is lowered enough to preserve the mean wealth level, we have neither a mean-preserving spread nor a mean-preserving contraction of the wealth distribution. Therefore, some risk averters would prefer self-insurance and others none.[2] We now show this more formally.

The expected utility of the individual is now given by

$$EU = (1 - q)U(A - c(x)) + qE[U(A - c(x) - L(Zx))], \tag{7}$$

where c(x) denotes the cost of self-insurance level x, c(0) = 0, c′ > 0. The first-order condition for maximizing (7) is

$$\frac{dEU}{dx} = -c'(x)[EU'] - qE[Z \cdot L'(Zx) \cdot U'(A - c(x) - L(Zx))] = 0. \tag{8}$$

The second-order condition is trivial, and a solution with $x^* > 0$ occurs if

$$c'(0) < \frac{-\,qE(Z)L'(0)U'(A - c(x) - L(0))}{EU'}. \tag{9}$$

We assume that the optimal x^* is finite and positive as well, which follows, for instance, if we also have $c'' \geq 0$. The two terms in equation (8) are respectively the expected marginal costs and benefits (in terms of utility) for increasing investment in self insurance.

We now consider strongly-more-risk-averse preferences, V, defined as in the previous section, and evaluate dEV/dx at the optimal self-insurance level for U, x^*.

$$\left.\frac{dEV}{dx}\right|_{x^*} = -c'(x)[EG'] - qE[Z \cdot L'(Zx) \cdot G'(A - c(x) - L(Zx))]. \tag{10}$$

The conditions on G are not strong enough to sign (10) and, hence, the optimal level of self-insurance can be either higher or lower under V than under U. Of course ambiguity under "strongly-more-risk averse preferences" implies ambiguity using Arrow-Pratt risk aversion as well. This contradicts the results of Dionne and Eeckhoudt [1985] and of Briys and Schlesinger [1990], who have shown that self-insurance is monotonically related to Arrow-Pratt risk aversion in the case where self-insurance is one-hundred percent reliable.

To illustrate this ambiguity in the present model and to show that the case of a negative relationship between risk aversion and the level of self-insurance is not vacuous, we consider the following example.

Example 1. Consider an individual whose preferences exhibit constant absolute risk aversion (CARA), $U(W) = -\exp(-\beta W)$, where the exponential coefficient β, $\beta > 0$, denotes the individual's (Arrow-Pratt) degree of absolute risk aversion. We make the following additional assumptions:

$$A = 4$$
$$q = .2$$
$$c(x) \equiv x$$
$$L(x) \equiv L(0)[\exp(-x)]$$
$$L(0) = 2$$
$$z \in \{0,1\};\ \text{prob}(z = 0) \equiv p.$$

We now consider the optimal value of x for various values of β and with several values for the parameter p, the probability that the self-insurance fails to perform properly. Note that in this simple example, self-insurance either performs perfectly (z = 1) or not at all (z = 0).

Results of computer calculations are plotted in Figure 1. It is easy to show, as evident from Figure 1, that $dx^*/dp < 0$. For the usual case of nonrisky self-insurance, the earlier result of Dionne and Eeckhoudt [1985] is confirmed; there is a monotonically increasing relationship between risk aversion and the level of self-insurance. For cases where self-insurance does not always perform properly, we see that self-insurance is decreasing in risk aversion for high enough levels of β. For the cases drawn (p = .10, p = .20 and p = .30), the level of self-insurance is negatively related to risk aversion for high enough levels of risk aversion. While the examples clearly show the non-monotonicity of self-insurance, we must admit that a negative relationship exists only at levels of risk aversion that seem rather high empirically.

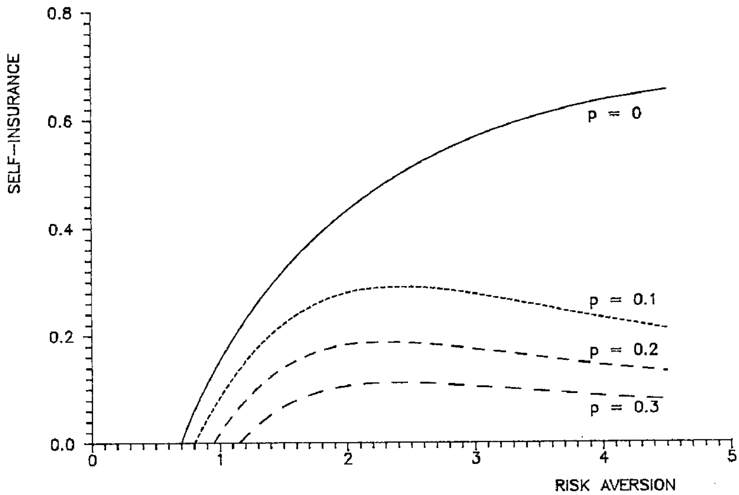

Figure 1. Effects of risk aversion on self-insurance.

4. Risky self-protection

We now assume that the individual can invest in self-protection activities which reduce the probability of a loss, but do not affect the size of a loss should it occur. The size of the loss is fixed at L, but the probability of loss is now a function of the level of self-protection activity, y. We assume that $q'(y) < 0$ and $q''(y) > 0$. The cost of self-protection is given by c(y), where $c' > 0$ and $c'' \geq 0$. Once again, we interject riskiness by use of the random variable Z and assume that the true probability of loss is an unobservable realization of q(Zy). Of course, if at most one loss occurs, the individual may never know which value of Z is realized. However, if aggregate loss information is available among consumers, there may be enough information to ascertain the distribution of Z. Besides, it is the individual's perception of the distribution of Z that is important for our model. The individual's expected utility is now given by

$$EU = \{1 - E[q(Zy)]\}U(A - c(y)) + E[q(Zy)]U(A - c(y) - L). \tag{11}$$

The first-order condition for maximizing (11) is

$$\frac{dEU}{dy} = -c'(y)[EU'] - E[Zq'(Zy)][U(W_1) - U(W_2)] = 0, \tag{12}$$

where $W_1 \equiv A - c(y)$ and $W_2 \equiv A - c(y) - L$.

The two terms in (12) are easily seen to be marginal costs and marginal benefits (both in utility terms) for increasing the level of self-protection. We once again assume an interior solution with $0 < y < \infty$.

It is straightforward to show that the optimal level of self-protection, y, may be either higher or lower under strongly-more-risk-averse preferences, and hence under more-risk-averse preferences in the Arrow-Pratt sense as well.

This result is not surprising. Indeed, even under nonrisky self-protection (i.e. $z \equiv 1$), Dionne and Eeckhoudt [1985] have shown that the level of self-protection is not necessarily monotonically linked to the degree of risk aversion. Briys and Schlesinger [1990] give an explanation for this seemingly counterintuitive behavior by showing that self-protection is not a risk-reducing activity, if risk is measured as proposed by Rothschild and Stiglitz [1970]. Obviously, adding riskiness to self-protection's effectiveness only tends to heighten the possibility of a negative relationship between the degree of risk aversion and the optimal level of self-protection. This is illustrated in the following example:

Example 2. We consider the same utility function as in Example 1. We further assume that

$$A = 4$$
$$L = 2$$

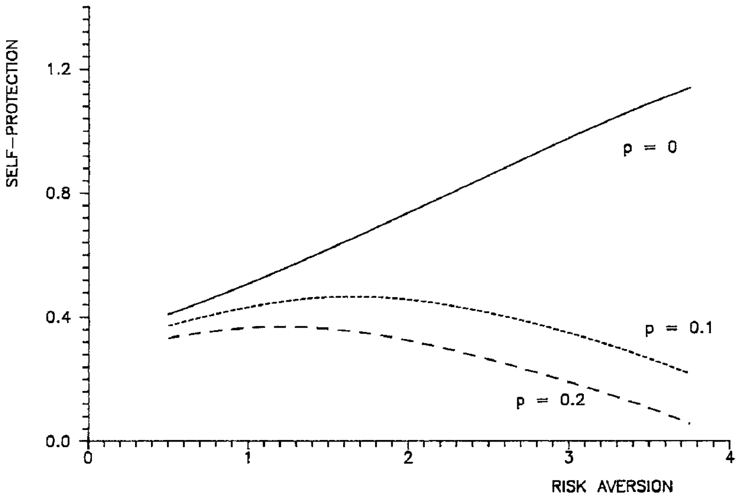

Figure 2. Effects of risk aversion on self-protection.

$$c(y) \equiv y$$
$$q(0) = .5$$
$$q(y) \equiv q(0)[\exp(-y)]$$
$$z \in \{0,1\};\ \text{prob}(z = 0) \equiv p.$$

The optimal level of self-protection for various levels of risk aversion, β, and various probabilities of a totally ineffective self-protection, p, are plotted in Figure 2. Although the level of self-protection is increasing in risk aversion for $p = 0$ in our example, this need not be the case in general. Indeed, Dionne and Eeckhoudt [1985] provide an example with $p = 0$ illustrating a negative relationship between the level of risk aversion and the level of self-protection.

5. Interactions between market insurance and self-insurance

A key result of Ehrlich and Becker [1972] is that, although market insurance and self-insurance are always substitutes, market-insurance and self-protection might be either substitutes or complements.[3] Conditions under which they are complements would tend to lessen the moral-hazard problem, or even lead to what Colwell and Wu [1988] call "moral imperative"—a higher level of insurance leading to the consumer's taking more care. In this setting, "care" is interpreted as self-protection activities. In the case of risky self-insurance, we show that "care" can also be interpreted as self-insurance activities. That is, market insurance and self-insurance may also be complements in our model, where the effects of self-insurance are not perfectly reliable.

We assume that the potential unreliability of the self-insurance is known both to the individual and to the insurance company, which accounts for this possibility in its insurance premiums. We use a model with nonrisky insurance, for the sake of simplicity. The individual's expected utility is given by

$$EU \equiv (1 - q)U(A - c(x) - P) + qE_zU[A - c(x) - P - (1 - \alpha)L(Zx)], \quad (13)$$

where E_z denotes expectations over Z and

$$P \equiv \alpha mqE_z[L(Zx)]. \quad (14)$$

Unlike in section 2, we do not maintain an assumption of actuarially-fair insurance prices and examine the effects of changes in the price of insurance. To this end, let m denote the "price" of insurance (equal to one plus a so-called "loading factor"). Thus, $m = 1$ denotes an actuarially-fair price while $m > 1$ denotes a positive premium loading. Also note how both the self-insurance and its nonreliability are reflected in the insurance premium. A higher investment in self-insurance and/or a higher degree of reliability will lower the insurance premium. The individual must simultaneously choose both α and x to maximize expected utility. The first-order conditions for this maximization are:

$$\frac{\partial EU}{\partial \alpha} = -qm\bar{L}[EU'] + qE_z[L(Zx) \cdot U'(W_2)] = 0 \quad (15)$$

and

$$\frac{\partial EU}{\partial x} = -\{c'(x) + m\alpha qE_z[Z \cdot L'(Zx)]\}[EU'] - q(1 - \alpha)E_z[Z \cdot L'(Zx) \cdot U'(W_2)] = 0 \quad (16)$$

where

$$\bar{L} \equiv E_z[L(Zx)]$$
$$W_2 \equiv A - c(x) - P - (1 - \alpha)L(Zx).$$

The first term in equation (15) represents the marginal cost (in utility terms) of increasing the premium. The second term represents the expected marginal benefit of a higher level of coverage, α. The actual benefit is random due to randomness of the loss size, which in turn stems from the nonreliability. For the case where loss size is fixed, equation (15) is easily seen to reduce to (3). The first and second terms in equation (16) are easily seen to represent the expected cost of increasing the level of self-insurance and the expected benefit respectively. This equation generalizes the optimality condition with no insurance, equation (8), and indeed the latter condition is easily obtained by setting $\alpha = 0$ above. When $\alpha > 0$, a higher x will also lower the insurance premium, thus reducing the net cost. However, since only a fraction $(1 - \alpha)$ of the incurred loss is "out-of-pocket," a positive level of insurance also lessens the benefit of an increase in self-insurance.

To examine whether self-insurance and market insurance are substitutes or complements, we examine consumer behavior due to changes in the "price" of insurance, m. Unfortunately, the comparative-static effects of changing the insurance are not as easily calculated as they are in the case of fully reliable self-insurance (see Ehrlich and Becker, [1972]). To show that the effect is indeed ambiguous, we consider the following example.

Example 3. We consider an example similar to Example 1, but with the allowance for the purchase of market insurance. The utility function is the same as in previous examples, but with β fixed at $\beta = 0.3$. Additionally, we assume

$$A = 4$$
$$q = 0.4$$
$$c(x) \equiv (0.2)x$$
$$L(x) \equiv L(0)[\exp(-x)]$$
$$L(0) = 2$$
$$z \in \{0,1\};\ \text{prob}(Z = 0) \equiv 0.375.$$

We do not maintain an actuarially-fair price of market insurance. Rather, we let m vary from $m = 1$ to $m = 1.35$ (i.e., from a zero "loading" to a 35% loading). As m rises, α^* falls from $\alpha^* = 1.0$ to $\alpha^* = 0.01$. However, as α^* falls, the corresponding maximal value of self-insurance, x^*, at first rises but then also falls. This is depicted in Figure 3, which shows the locus of optimal (α^*, x^*) pairs for the various values of m. For $m > 1.3$, market insurance and self-insurance are seen to be complements, contrary to the usual case under full reliability, where they are always substitutes.

The intuition behind this result may be somewhat obscure, but might be best understood by focusing on the worst possible outcome for the consumer. This occurs in the state of nature where both a loss occurs and the self-insurance fails. In this case, the consumer not only suffers the higher loss, L(0), but also loses the investment in self-insurance, c(x). At higher values of m, less insurance is purchased and so more of the loss is borne "out of pocket."[4] By decreasing the investment in self-insurance, the consumer can at least improve the worst possible state of the world—which is exactly what the consumer does for $m > 1.3$ in our example.

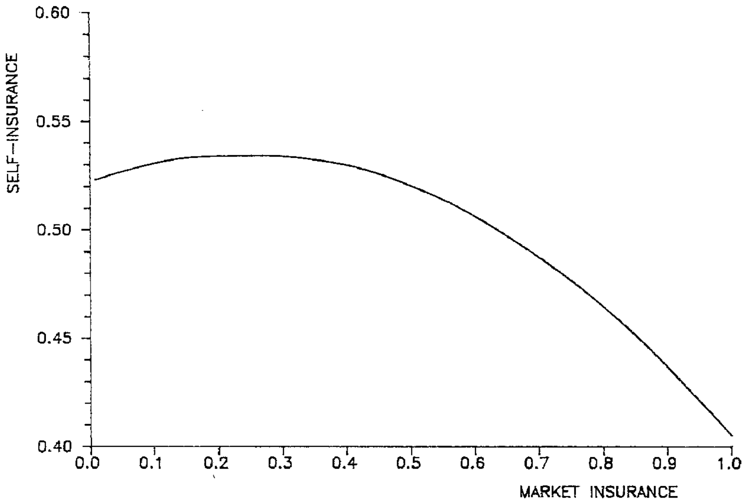

Figure 3. Relationship between market insurance and self-insurance.

6. Concluding remarks

We have examined the purchase of market insurance, self-insurance and self-protection in a world where reliability cannot be guaranteed. The model indicates that most of the results from a world in which reliability reigns do not carry over into a model with risky risk-management tools. In particular, we showed that the positive relation that usually exists between risk aversion and both market insurance and self-insurance no longer holds when these tools are not fully reliable. We also showed that, unlike in Ehrlich and Becker [1972], market insurance and self-insurance can be complements as well as substitutes in the nonreliability setting. We should, therefore, not be surprised to find that empirical evidence is often contrary to some of the existing theories, which were formulated in a world of complete reliability. While most of the criticism along these lines has focused on flaws within the expected utility paradigm itself, existing models simply do not take all of the market imperfections into account. The fact that the very tools intended to mitigate risk also introduce a new risk of their own causes the interactions to be understandably "murky." It also implies that we should probably not expect too much from the theory in terms of providing fail-proof predictions.

Clearly the issue of reliability is very real and one whose dimensions extend beyond those considered in the current paper. For example, insurance guaranty funds as well as warranties on other risk management devices might partially mitigate the nonreliability problem. However, even these safety mechanisms are not fully reliable and to the extent that the problem cannot be completely alleviated, the market is incomplete. Future research will hopefully lead to a better understanding of the extent of nonreliability and its effects on consumer behavior.

Acknowledgments

We are grateful to Neil Doherty, Louis Eeckhoudt, Henri Loubergé and an anonymous referee for reliable comments. Remaining errors are our own.

Notes

1. In the insurance literature, "self-insurance" and "self-protection" are usually referred to as "loss reduction" and "loss prevention" respectively.
2. Hiebert [1989] correctly points out that the variance of wealth is higher under nonreliability. However, general statements about risk-averse behavior do not necessarily follow from this observation. It is the lack of any second-degree stochastic dominance that leads to the ambiguities. For further discussion, see Rothschild and Stiglitz [1970].
3. Here, we are referring to what are commonly called "gross substitutes" and "gross complements." That is, we use the terminology with reference to Marshallian demand.
4. In general, this statement is not always true as insurance for a fixed risk may be a Giffen good. However, this is never the case under CARA preferences as was shown by Mossin [1968].

References

BRIYS, E. and H. SCHLESINGER, [1990]: "Risk Aversion and the Propensities for Self-Insurance and Self-Protection," *Southern Economic Journal,* 57 (October), 458–467.

COLWELL, P. and C. WU, [1988]: "Moral Hazard and Moral Imperative," *Journal of Risk and Insurance,* 55 (March), 101–117.

DIONNE, G. and L. EECKHOUDT, [1985]: "Risk Aversion, Self-Insurance and Self-Protection," *Economics Letters,* 17, 39–57.

DOHERTY, N. and H. SCHLESINGER, [1990]: "Rational Insurance Purchasing: Consideration of Contract Nonperformance," *Quarterly Journal of Economics,* 105 (February), 143–153.

EHRLICH, I. and G. BECKER, [1972]: "Market Insurance, Self-Insurance and Self-Protection," *Journal of Political Economy,* 80 (July/August), 623–647.

HIEBERT, L.D., [1989]: "Self-Insurance, Self-Protection and Increases in Risk Aversion," *Journal of Risk and Insurance,* 56 (June), 300–305.

MOSSIN, J., [1968]: "Aspects of Rational Insurance Purchasing," *Journal of Political Economy,* 76 (July/August), 553–568.

ROSS, S., [1981]: "Some Stronger Measures of Risk Aversion in the Large and in the Small with Applications," *Econometrica,* 49 (May), 621–638.

ROTHSCHILD, M. and J. STIGLITZ, [1970]: "Increasing Risk: I. A Definition," *Journal of Economic Theory,* 3, 66–84.

SCHLESINGER, H. and J.-M. GRAF v.d. SCHULENBURG, [1987]: "Risk Aversion and the Purchase of Risky Insurance," *Zeitschrift für Nationalökonomie,* 47 (April), 309–314.

TAPIERO, C., Y. KAHANE and L. JACQUE, [1987]: "Insurance Premiums and Default Risk in Mutual Insurance," *Scandinavian Actuarial Journal,* 1987, 82–97.