

Mossin's Theorem given Random Initial Wealth

Professor Soon Koo [Hong](#)

Department of Business Administration

Seoul National University of Technology

Professor Keun Ock [Lew](#)

Department of Business Administration

Seoul National University of Technology

Professor Richard [MacMinn](#)

Edmondson-Miller Chair

College of Business

Illinois State University

Mossin's Theorem Given Random Initial Wealth

ABSTRACT

The purpose of this paper is to re-examine Mossin's Theorem under random initial wealth. Conditions sufficient for Mossin's Theorem depend on stochastic interdependence between risks. The correlation coefficient, however, is not adequate to investigate Mossin's Theorem in the general expected utility model. This is illustrated with two counter examples, which show that correlation is not a useful measure that leads to a definitive theorem. Alternatively, using negative and positive interdependence measures we provide necessary and sufficient conditions for a generalized Mossin Theorem to hold. In addition, Mossin's Theorem is interpreted using the notion of a mean preserving spread made popular by Rothschild and Stiglitz. Given a fair premium and interdependent stochastic conditions, we show that an individual can always obtain a final wealth distribution with less weight in its tails by selecting less than or more than full insurance.

Key Words: Mossin's Theorem, Expectation Dependence, Quadrant Dependence, Regression Dependence
Mean Preserving Spread, Mean Preserving Contraction, and Fair Insurance

I. Introduction

"A risk averse individual will choose to fully insure at an actuarially fair premium." This statement is known as the Mossin Theorem (Mossin 1968a).¹ The purpose of this paper is to reexamine the theorem under initial wealth uncertainty in a general expected utility model.

When initial wealth is random, the background risk affects the demand for insurance. Conditions necessary and sufficient for the Mossin Theorem depend on the stochastic interdependence between the risks. Intuitively, if positive correlation exists between insurable risk and background risk then Mossin's Theorem seems to be violated because a natural hedge exists. Conversely, negative correlation may look sufficient for full insurance at an actuarially fair premium since negative correlation increases total risk of final wealth. Several authors give these kinds of analyses in the mean-variance framework (Mayers and Smith Jr. 1983; Doherty 1984; Schulenburg 1986). In particular, Doherty and Schlesinger (Doherty and Schlesinger 1983a) show that Mossin's Theorem holds if the correlation between insurable and uninsurable risks is negative or zero, and that Mossin's Theorem may not be valid if the correlation is positive. They derive these results under a bivariate normal distribution within an expected utility framework.²

In the general expected utility model, however, the correlation measure is not adequate to investigate Mossin's Theorem (Doherty and Schlesinger 1983a; Schlesinger and

¹ This well-known proposition is also called the Bernoulli Principle, i.e., see Doherty, N. A. and H. Schlesinger (1983). "The Optimal Deductible for an Insurance Policy When Initial Wealth is Random." Journal of Business 56(4): 555-565, Doherty, N. A. and H. Schlesinger (1983). "Optimal Insurance in Incomplete Markets." Journal of Political Economy 91(6): 1045-1054.

² Doherty and Schlesinger (1983a) consider the case in which the random initial wealth is independent of the random loss without the bivariate normal assumption also. The results that we refer to here are those where the random initial wealth and loss are dependent; there the Doherty and Schlesinger employ the bivariate normal assumption. More recently Mossin's Theorem was investigated in Schlesinger, H. (2000). The Theory of Insurance Demand. The Handbook of Insurance. G. Dionne. Boston, Kluwer Academic Publishers: 131-151. and extended to an upper-limit insurance contract in Schlesinger, H. (2005). "Mossin's Theorem for Upper-Limit Insurance Policies." Journal of Risk and Insurance. While Mossin presented the theorem in terms of coinsurance demand, Schlesinger shows that it also holds for a choice of deductible insurance, i.e., see Schlesinger, H. (1981). "The Optimal Level of Deductibility in Insurance Contracts." Journal of Risk and Insurance 48: 465-481, Schlesinger, H. (2000). The Theory of Insurance Demand. The Handbook of Insurance. G. Dionne. Boston, Kluwer Academic Publishers: 131-151. As pointed out by Schlesinger (2006), Mossin's Theorem seems so basic it may hold when indemnity schedule depends on one choice parameter. When indemnity schedule depends on more than one choice parameter, Mossin's Theorem may be violated. Cummins, J. D. and O. Mahul (2004). "The Demand for Insurance With an Upper Limit on Coverage." Journal of Risk and Insurance 71: 253-264. show that partial deductible insurance is optimal under fair premium if the indemnity payoff is subject to an upper limit.

Doherty 1985; Schlesinger 2000) (Doherty and Schlesinger 1983b). For the case of random initial wealth, Mossin's Theorem may hold only under restricted conditions. (Schlesinger 2000). In this paper, two counter examples are presented to show that correlation is not a sufficient measure to lead to the theorem. In addition, we discuss the notions of dependence that have been developed, e.g., see (Lehmann 1966), (Wright 1987), and then use the notion of expectation dependence to provide necessary and sufficient conditions for Mossin's Theorem to hold in the presence of random initial wealth. We show that the rational individual would buy less than full, full or more than full insurance, as the measure of interdependence is positive, zero or negative, respectively. This result is a generalized version of Mossin's Theorem. We show that the interdependence measure is crucial because it suffices to show that, given a fair premium and positive or negative interdependent risks, an individual can obtain a final wealth distribution with less risk or equivalently with less weight in its tails by selecting either less or more than full insurance.³

The remainder of the paper is organized as follows: In section II, earlier mean-variance results are briefly summarized while in section III two counter examples are constructed to show the deficiency of the correlation measure. In section IV, we state and prove a generalized version of the Mossin Theorem and in section V we provide an interpretation of the generalized Mossin Theorem by using the notion of a mean preserving contraction. Section VI concludes.

II. A Sketch of Earlier Results

Consider a risk averse individual who is endowed with random initial wealth W and is exposed to a random loss X . The individual's wealth prospect is $Y = W - X$. We use the following notation:

| | |
|-----|---|
| W | the random initial wealth; $-\infty < w < \infty$ |
| X | the random loss; $0 \leq x < \infty$ |

³ This less risk or equivalently less weight in the tails of the distribution is the way that Rothschild and Stiglitz characterized a reduction in risk, e.g., see Rothschild, M. and J. E. Stiglitz (1970). "Increasing Risk: I. A Definition." *Journal of Economic Theory* 2: 225-43.

Mossin's Theorem Given Random Initial Wealth

| | |
|------------|---|
| $F(w, x)$ | joint distribution function for (W, X) |
| $f(w, x)$ | the joint probability density function for W and X |
| $f_w(w)$ | the marginal density function of W ; $f_w(w) = \int_0^\infty f(w, x) dx$ |
| $g(x w)$ | the conditional density function of x given w ; $g(x w) = \frac{f(w, x)}{f_w(w)}$ |
| ρ | correlation coefficient |
| a | coefficient of coinsurance, $a \geq 0$ |
| P | the premium for full insurance |
| Y | the final random wealth, $-\infty < y < \infty$; $Y = W - aP - (1 - a)X$ |

Suppose that insurance is available via a coinsurance contract. By assuming proportional loading, the insurance premium for coverage a is characterized by a $P = a(1 + \lambda)E(X)$ where P is the premium for full insurance and $\lambda \geq 0$ is a loading factor. Given the fair contract, i.e., $\lambda = 0$, the premium is equal to the expected indemnity payment by the insurer, i.e., $aP = aEX$. Now, the individual's final wealth Y is given by⁴

$$Y = W - aP - (1 - a)X \quad (1)$$

The distribution of Y depends on the joint distribution of W and X . The joint probability density function will be denoted by $f(w, x)$.

⁴ If a probability mass of nonzero loss is explicitly considered, the individual's final wealth Y will be given by $Y = W - aP - h(1 - a)X$ where h assumes the value of zero with probability q or the value of one with probability $(1 - q)$. However, all the results that we derive in this model are so intuitive that they apply trivially to the compound risk model.

1. Mean-Variance Model

An intuitive analysis of Mossin's Theorem can be given in terms of mean and variance of final wealth. Note that given an actuarially fair premium the individual's expected final wealth remains constant regardless of insurance choice but the second moment of final wealth distribution will be changed. In this case a risk averse individual will prefer a final wealth with less variance to any other wealth. This implication can be summarized as the following Lemma.

Lemma 1: Assume that the hypotheses of the mean-variance model hold. Then, given an actuarially fair premium, the individual will purchase less than full, full or more than full insurance, as the covariance is positive, zero or negative, respectively, i.e.,

- i. $a < 1$ if $\text{Cov}(X, W) > 0$
- ii. $a = 1$ if $\text{Cov}(X, W) = 0$
- iii. $a > 1$ if $\text{Cov}(X, W) < 0$

Proof: The proof is straightforward. If an insurance contract can be made without cost, i.e., $P = EX$, then the individual's expected wealth remains constant regardless of a . This can be easily confirmed as follows

$$EY = EW - aP - (1 - a)EX = EW - P = EW - EX .$$

Now, let the variance of final wealth Y for given a be denoted by $V(a)$. Then, $V(a)$ is

$$V(a) = \text{Var}(W) + (1 - a)^2 \text{Var}(X) - 2(1 - a)\text{Cov}(W, X) .$$

In the present mean-variance model, since $V''(a) = 2\text{Var}(X) > 0$, the conditions for less than full, full, and more than full insurance are $V'(1) > 0$, $V'(1) = 0$, and $V'(1) < 0$ respectively. Differentiating $V(a)$ with respect to a and evaluating at $a = 1$ yields $V'(1) = 2\text{Cov}(X, W)$. Hence, the individual will buy less than full, full or more than full insurance, as the covariance is positive, zero or negative, respectively. **QED**

The Lemma is directly derived from our risk model. Equivalently, necessary and sufficient conditions for optimal insurance coverage can be derived for any suitably

modified mean-variance model (Mayers and Smith Jr. 1983; Doherty 1984; Schulenburg 1986).

2. Expected Utility Model under a Normal Distribution

Corresponding results can be shown in the expected utility model if normality is assumed. Consider a risk averse individual with the utility function u which satisfies $u' > 0$ and $u'' < 0$. The individual chooses the insurance coverage a by maximizing expected utility such that

$$U(a) \equiv Eu(Y) = \int_{-\infty}^{\infty} \int_0^{\infty} u(w - aP - (1-a)x) f(w, x) dx dw \quad (2)$$

The first and second order conditions for a maximum are

$$U'(a) = \int_{-\infty}^{\infty} \int_0^{\infty} u'(y)(x - P) f(w, x) dx dw = 0 \quad (3)$$

$$U''(a) = \int_{-\infty}^{\infty} \int_0^{\infty} u''(y)(x - P)^2 f(w, x) dx dw < 0, \quad (4)$$

respectively. The second order condition in (4) is satisfied since $u'' < 0$. In this case, if the joint distribution of W and X is bivariate normal, then we can confirm the results of the lemma one in the present expected utility model. To see this, utilizing the theorem developed by (Rubinstein 1976)⁵ we can rewrite $U'(1)$ as

$$\begin{aligned} U'(1) &= E(u'(W - P)(X - P)) \\ &= Eu'(W - P)(EX - P) + Cov(u'(W - P), X) \\ &= Cov(u'(W - P), X) \\ &= Eu''(W - P) Cov(W, X) \end{aligned} \quad (5)$$

⁵ If W and X are normally distributed, Y will also be normal. Hence, $Cov(u'(Y), X) = Eu''(Y) Cov(Y, X)$.

Due to the concavity of U the conditions for less than full, full, and more than full insurance are $U'(1) < 0$, $U'(1) = 0$, and $U'(1) > 0$, respectively. Thus, we have optimal less than, full, or more than full insurance as $\text{Cov}(W, X) > 0$, $\text{Cov}(W, X) = 0$ or $\text{Cov}(W, X) < 0$, respectively. This result is equivalent to the lemma one. Doherty and Schlesinger (1983a) derive similar results using a deductible contract. In essence, they show that, if the distribution of W and X is bivariate normal and negative correlation exists between W and X then the individual buys full insurance at an actuarially fair price.

III. Counter Examples

It has frequently been convenient to characterize return and risk by the mean and variance of the final wealth distribution. This approach still suggests significant insights about optimal choice. In the general expected utility model, however, the correlation measure is not adequate to investigate the demand for insurance. For example, in the Doherty-Schlesinger (1983a) model if the normality assumption is dropped, optimal full coverage may not be desirable even though negative correlation exists. Moreover, under positive correlation optimal full or more than full insurance is possible. This is illustrated in the following two numerical examples.

Counter Example 1: Optimal Partial Insurance under Negative Correlation?

As shown in condition (iii) of the lemma, negative correlation leads to more than full insurance in the mean-variance model. This, however, may not be true in the general expected utility model. Let $W = Z + 3$ and $X = -12Z^2 + 10Z + 2$, where Z is a uniformly distributed random variable on $[0, 1]$. Then, $EW = 7/2$, $\sigma_W^2 = 1/12$, $X \in [0, 49/12]$, $EX = P = 3$, $\sigma_X^2 = 17/15$, $\text{Cov}(W, X) = -1/6$, and $\rho_{WX} \approx -0.5443$. Suppose that an individual has a strictly concave utility function $u(Y) = \sqrt{Y}$. With a fair premium $P = 3$ it follows that

$$\begin{aligned}
 U'(1) &= E(u'(W - P)(X - P)) \\
 &= \int_0^1 \left(-6z^{\frac{3}{2}} + 5z^{\frac{1}{2}} - \frac{1}{2}z^{-\frac{1}{2}} \right) dz \\
 &= -\frac{1}{15}
 \end{aligned}$$

Thus, even though X and W are negatively correlated, more than full insurance is not desirable. Only less than full coverage is optimal and this contradicts condition (iii) of lemma one.

Counter Example 2: Optimal Full Insurance under Positive Correlation?

If X and W are positively correlated then they tend to move together and so an increased loss may be associated with increased wealth. This positive correlation corresponds to a natural hedge. Hence, we may guess that partial insurance coverage with a fair premium is optimal. This is shown in condition (i) of the lemma. Once again, however, this result does not hold in the general model.

A counter-example similar to that of counter-example one can be constructed. Let $W = Z + 13/12$, $X = 12Z^2 - 10Z + 25/12$. and let Z and u be defined as in the previous counter-example. It follows that $E(W) = 19/12$, $\sigma_W^2 = 1/12$, $X \in [0, 49/12]$, $E(X) = P = 13/12$, $\sigma_X^2 = 17/15$, $\text{Cov}(W, X) = 1/6$, and $\rho_{WX} \approx 0.5443$. Direct calculation shows $U'(1) = 1/15 > 0$. Thus, given a positive correlation between X and W , full or more than full insurance is desirable. Hence, this shows once again that the correlation coefficient is not a satisfactory measure to generate the Mossin Theorem.

IV. Generalized Mossin Theorem

When initial wealth is random it is evident that the stochastic interdependence between two random variables can play a critical role in determining optimal insurance coverage. The counter examples in the previous section, however, suffice to show that the correlation coefficient while suggestive does not provide the appropriate means to generalize the Mossin theorem. The notion of interdependence needs to be strengthened.

Several notions of dependence exist in the literature, e.g., see (Lehmann 1966; Samuelson 1967; Brumelle 1974; MacMinn 1984; Wright 1987; Aboudi and Thon 1995). The three most frequently noted notions of dependence are quadrant, regression and expectation dependence. Suppose X and Y are random variables with joint distribution function F and marginal distributions F_X and F_Y .

Definition 1: X is negatively quadrant dependent (NQD) with respect to Y if

$$F(x,y) \leq F_X(x)F_Y(y) \quad (6)$$

for all x, y . The dependence is strict if the inequality holds for at least some pair (x, y) .

It may be observed that the covariance of (X, Y) is negative if X is negative quadrant dependent with respect to Y , i.e., see lemma three in (Lehmann 1966). Note that the definition of negative quadrant dependence could also have been stated as X is NQD with respect to Y if

$$F_X(x|Y \leq y) \leq F_X(x) \quad (7)$$

for all x, y . Let \mathcal{F}_Q denote the family of distributions F satisfying (6) and (7); let \mathcal{G}_Q denote the family of distributions satisfying the converse of (6) and (7) so that it represents the totality of positively quadrant dependent distributions (PQD). The NQD notion expresses the fact that knowing Y is small decreases the probability that X is small.

The notion of negative dependence might be expressed differently as

$$F_X(x|Y \leq y) \leq F_X(x|Y \leq y') \quad (8)$$

for all $y < y'$ and all x . The next notion of negative dependence, however, is typically based on a stronger condition.

Definition 2: X is negatively regression dependent (NRD) with respect to Y if

$$F_X(x|Y = y) \text{ is increasing in } y \quad (9)$$

The family of distributions F of (X, Y) for which (9) holds will be denoted by \mathcal{F}_R and the family of distributions for which the conditional distribution in (9) is a decreasing function of y will be denoted by \mathcal{G}_R ; the latter set contains all the distributions exhibiting positive regression dependence (PRD). Lehmann has shown that (9) implies (8) and (8) implies (7), i.e., see lemma four in (Lehmann 1966); the reverse implications do not generally hold. Hence, $\mathcal{F}_Q \supset \mathcal{F}_R$.

Finally, expectation dependence was defined and introduced by (Wright 1987). Actually (Brumelle 1974) and (MacMinn 1984) used notions of expectation dependence without the explicit definitions. As in regression dependence, a weaker or stronger form may be used to define the concept. Wright introduced the weaker notion as in the following definition.

Definition 3: X is negatively expectation dependent (NED) with respect to Y if

$$E(X|Y \leq y) \geq EX \tag{10}$$

for all y .

The family of distributions F of (X, Y) for which (10) holds will be denoted by \mathcal{F}_E and the family of distributions for which the condition in (10) is reversed will be denoted by \mathcal{G}_E ; the latter set contains all the distributions exhibiting positive expectation dependence (PED). The next theorem is one implication of NED.

Theorem 1: $\text{Cov}(X, g(Y)) > 0$ for every decreasing real valued function g if and only if X is NED with respect to Y .⁶

Observe that (7) yields

$$\begin{aligned} \int_{-\infty}^{\infty} (F_X(x) - F_X(x|Y \leq y)) dx &= \int_{-\infty}^{\infty} x dF_X(x|Y \leq y) - \int_{-\infty}^{\infty} x dF_X(x) \\ &= E(X|Y \leq y) - EX \\ &\geq 0 \end{aligned} \tag{11}$$

⁶ This is similar to theorem 3.1 in Wright, R. (1987). "Expectation Dependence of Random Variables, With an Application in Portfolio Theory." *Theory and Decision* 22(2): 111-124. The proof is similar as well.

and so NQD implies NED in its weaker form. Further we now have NRD implies NQD which in turn implies NED. The reverse implications do not generally hold. Hence, $\mathcal{F}_E \supset \mathcal{F}_Q \supset \mathcal{F}_R$. To see that the reverse implications do not hold, consider the following example. Let the joint distribution of (W, X) be given by the following 3×3 table

$$\begin{pmatrix} (w_1, x_1) & (w_1, x_2) & (w_1, x_3) \\ (w_2, x_1) & (w_2, x_2) & (w_2, x_3) \\ (w_3, x_1) & (w_3, x_2) & (w_3, x_3) \end{pmatrix} = \begin{pmatrix} 0.15 & 0 & 0.05 \\ 0 & 0.25 & 0 \\ 0.55 & 0 & 0 \end{pmatrix} \quad (12)$$

where $w_1 < w_2 < w_3$ and $(x_1, x_2, x_3) = (1, 2, 6)$. Then it follows by direct calculation that this X is NED with respect to W . It also follows by direct calculation that X is neither NQD nor NRD with respect to W .

There is a stronger definition of NED that is similar to the definition of NRD and this version of expectation dependence was used in (Brumelle 1974; MacMinn 1984). Brumelle showed that NRD implies the stronger NED in definition four.

Definition 4: X is strongly negatively expectation dependent with respect to Y if

$$E(X|Y = y) \text{ is decreasing in } y \quad (13)$$

Either version of expectation dependence may be used to provide a more robust version of Mossin's theorem. Using the weaker form of expectation dependence, the following generalization of the Mossin theorem provides the necessary and sufficient conditions for less than full, full, or more than full insurance given a fair premium.⁷

Theorem 2a (Generalized Mossin Theorem): Given a fair premium, the risk averse individual will purchase less than full or more than full insurance if and only if X is positively or negatively expectation dependent with on W , respectively.

Proof: Consider the sign of $U'(1)$. For $a = 1$, $Y = W - P$ and

⁷ The appendix contains a simple and direct proof of the generalized Mossin theorem for the stronger version of expectation dependence.

Mossin's Theorem Given Random Initial Wealth

$$\begin{aligned}
 U'(1) &= \int_{-\infty}^{\infty} \int_0^{\infty} u'(w-P)(x-P) f(w,x) dx dw \\
 &= E(u'(W-P))(EX-P) + \text{Cov}(u'(W-P), X) \\
 &= \text{Cov}(u'(W-P), X)
 \end{aligned} \tag{14}$$

Therefore $U'(1) < 0$ if and only if $\text{Cov}(u'(W-P), X) < 0$ or equivalently, by theorem one, if and only if X is positively expectation dependent with respect to W . Hence, positive expectation dependence is necessary and sufficient to yield less than full insurance optimal. Similarly, $U'(1) > 0$ if and only if $\text{Cov}(u'(W-P), X) > 0$ or equivalently, by theorem one, if and only if X is negatively expectation dependent with respect to W . Hence, it follows that negative expectation dependence is necessary and sufficient to yield more than full insurance optimal. **QED**

It is important to note that the expectation dependence used in the generalized Mossin theorem is stronger than a condition on the covariance of W and X ; indeed if the distribution of (W, X) is in \mathcal{F}_Q then it exhibits negative quadrant dependence and so yields negative covariance but \mathcal{F}_Q is a subset of the family of negative expectation dependent distributions \mathcal{F}_E . We have shown that negative quadrant dependence implies negative covariance and negative expectation dependence but the converse is not generally true.

It should also be noted that (Aboudi and Thon 1995) have demonstrated that if negative regression dependence holds and the insurance premium is fair then the risk averse individual prefers high to low coverage. The result here shows that negative expectation dependence and a fair insurance premium provides the risk averse individual with an incentive to purchase more than full insurance if it is available and so the two results are the same but the conditions required for the same conclusion are different. Negative expectation dependence is a more general condition in the sense that negative regression dependence implies negative expectation dependence but we may have a distribution that exhibits negative expectation dependence but not negative

regression dependence.⁸ Hence, the generalized Mossin theorem here is indeed a generalization of the results in the literature.

V. Mean Preserving Contraction Interpretation

The optimality of less than full or more than full coverage with a fair premium can be explained by appealing to the Rothschild-Stiglitz notion of increasing risk. Rothschild and Stiglitz (1970) showed in part that all risk averse investors preferred a risk Y to another risk Z if the latter risk is a mean preserving spread of the former or equivalently if Z has more weight in the tails of its distribution than Y . To follow this line of reasoning, let Y_F and Y_L denote two risks where Y_F represents final wealth given full insurance coverage and Y_L represents final wealth given less than full insurance coverage. Then

$$Y_F = W - P$$

$$Y_L = W - aP - (1 - a)X$$

Given a fair premium, moving from full to less than full coverage does not change the mean, i.e., $EY_F = EY_L$. We know from the theorem two that $Eu(Y_L) > Eu(Y_F)$ if X is PED with respect to W . The next theorem shows that this result follows because reducing insurance coverage reduces weight in the tails of the final wealth distribution; equivalently reducing insurance coverage yields a mean preserving contraction.

Theorem 3 (Contraction Theorem): Given a fair premium, the individual holding full insurance can obtain a final wealth distribution with less weight in both tails by

⁸ For example, if we set $(w_1, w_2, w_3) = (7, 8, 9)$ and $(x_1, x_2, x_3) = (1, 2, 6)$ for the given distribution table in (12) then we have $P = EX = 1.5$, and for $u(Y) = \sqrt{Y}$ it follows that

$$U'(1) = E(u'(W - P)(X - P)) = \frac{1}{2} E\left(\frac{X - P}{\sqrt{W - P}}\right) = \frac{0.01257}{2} > 0.$$

Recall, from the example, that this distribution does not exhibit NQD or NRD. Therefore this illustration confirms our argument that expectation dependence, rather than regression dependence, is the suitable measure to define the condition for Mossin Theorem.

- a) reducing the insurance to less than full coverage if X is strongly PED with respect to W ,
- b) increasing the insurance to more than full coverage if X is strongly NED with respect to W .

Proof: First, consider (a) and note that if $E(X | W)$ is increasing in w , then there exists a wealth v such that EX is greater than, equal to or less than $E(X | W)$ as w is greater than, equal to or less than v , respectively. Let $P\{Y \leq y\}$ denote the probability distribution of final wealth Y . We want to show that the distribution of Y has less weight in both tails when insurance coverage is reduced from full coverage. Hence we must show that (i) $-\partial P\{Y > y\}/\partial a = \partial P\{Y \leq y\}/\partial a < 0$ for $a = 1$ and sufficiently large y and (ii) $-\partial P\{Y \leq y\}/\partial a < 0$ for $a = 1$ and sufficiently small y .

The probability distribution can be written as

$$\begin{aligned} P\{Y \leq y\} &= P\{W - aP - (1-a)X \leq y\} \\ &= \int_0^\infty \int_0^{h(a,x)} f(w,x) dw dx \end{aligned} \tag{15}$$

where

$$h(a,x) = y + aP + (1-a)x \tag{16}$$

By Leibniz formula, the derivative with respect to coverage is

$$\begin{aligned} \frac{\partial}{\partial a} P\{Y \leq y\} &= \frac{\partial}{\partial a} \left[\int_0^\infty \int_0^{h(a,x)} f(w,x) dw dx \right] \\ &= \int_0^\infty \left(\frac{\partial h}{\partial a} f(h(a,x), x) \right) dx \\ &= \int_0^\infty ((P-x) f(y - aP - (1-a)x, x)) dx \end{aligned} \tag{17}$$

Evaluating the above derivative at $a = 1$ yields

$$\frac{\partial}{\partial a} P\{Y \leq y\} \Big|_{a=1} = \int_0^\infty (P-x) f(y + P, x) dx \tag{18}$$

To show that (i) holds let y be such that $y + P = w > v$. Then

$$\begin{aligned}
 -\frac{\partial}{\partial a} \mathbb{P}\{Y > y\} \Big|_{a=1} &= -\frac{\partial}{\partial a} (1 - \mathbb{P}\{Y \leq y\}) \Big|_{a=1} \\
 &= \int_0^\infty (P - x) f(w, x) dx \\
 &= \int_0^\infty (P - x) f_w(w) g(x|w) dx \\
 &= f_w(w) (P - E(X|w)) \\
 &< 0
 \end{aligned} \tag{19}$$

The inequality in (19) holds because $w < v$ yields $P = EX < E(X | w)$. To show (ii) holds, let y be such that $y + P = w < v$. Then

$$\begin{aligned}
 -\frac{\partial}{\partial a} \mathbb{P}\{Y \leq y\} \Big|_{a=1} &= -\int_0^\infty (P - x) f(w, x) dx \\
 &= -\int_0^\infty (P - x) f_w(w) g(x|w) dx \\
 &= -f_w(w) (P - E(X|w)) \\
 &< 0
 \end{aligned} \tag{20}$$

The inequality in (20) holds because $P = EX > E(X | w)$. Hence, the weight in both tails of the final wealth distribution is reduced by reducing coverage and condition (a) holds. The same procedure suffices to show that condition (b) holds. **QED**

Given the interdependence conditions, a marginal decrease or increase from full insurance generates a mean preserving contraction in risk.

VI. Summary and Conclusions

This paper provides a generalization of Mossin's theorem to a setting with random initial wealth. It is a necessary generalization for the study of the demand for insurance in a financial market setting. Uninsurable initial wealth risk affects the

optimal coverage of insurable risk; random initial wealth provides some new insight into the role that stochastic interdependence between risks plays in determining the optimal insurance coverage. Somewhat counter-intuitively the correlation coefficient is not a satisfactory measure to investigate the demand for insurance. If the correlation measure is replaced by an expectation dependence condition then unambiguous results can be derived. In essence, we show that, if positive, zero, or negative expectation dependence exists, then the rational individual will purchase less than full, full or more than full insurance coverage, respectively. From a traditional viewpoint, optimal coverage that is less than or more than full insurance coverage with a fair premium may seem strange for a risk averse individual. These results, however, are hardly surprising since moving from full coverage generates a mean preserving contraction if initial wealth and losses are interdependent. We demonstrate the contraction by assuming full insurance and stochastic conditions and then showing that by increasing or decreasing the amount of insurance the risk averse individual is able to reduce the weight in both tails of the final wealth distribution and consequently decrease the risk of the final wealth distribution. More work remains to be done. Some of the standard results on insurance demand need to be modified within a more robust financial market framework.

Appendix

The following is a simple alternative proof of the generalized Mossin theorem with the stronger form of expectation dependence. It should be clear here that strong expectation dependence can only be used as a sufficient condition.

Theorem 2b: Given a fair premium, the individual will purchase less than full or more than full insurance if X is strongly positively or negatively expectation dependent with respect to W , respectively.

Proof: Recall that X is strongly PED with respect to W if $E(X | W = w)$ is increasing in w ; similarly X is strongly NED with respect to W if $E(X | W = w)$ is decreasing in w . Hence, evaluate the sign of $U'(1)$.

Mossin's Theorem Given Random Initial Wealth

$$\begin{aligned}
 U'(1) &= \int_{-\infty}^{\infty} \int_0^{\infty} u'(w-P)(x-P) f(w,x) dx dw \\
 &= \int_{-\infty}^{\infty} u'(w-P) \left[\int_0^{\infty} (x-P) f(x|w) dx \right] f_w(w) dw \\
 &= \int_{-\infty}^{\infty} u'(w-P) [E(X|w) - P] f_w(w) dw & (21) \\
 &= Eu'(W-P) (EX - P) + Cov(u'(W-P), E(X|W)) \\
 &= Cov(u'(W-P), E(X|W))
 \end{aligned}$$

The second equality follows by noting that $f(w, x) = f_w(w) f_x(x|w)$ where $f_w(w)$ is the marginal density of W and $f_x(x|w)$ is the conditional density of X . Therefore, if $E(X|W)$ is increasing in w then $U'(1) < 0$ and so $a < 1$ is optimal. Similarly if $E(X|W)$ is constant, then $U'(1) = 0$ and $a = 1$ is indeed optimal. Finally, if $E(X|W)$ is decreasing in w then $U'(1) > 0$ and so $a > 1$ is optimal. **QED**

References

- Aboudi, R. and D. Thon (1995). "Second-Degree Stochastic Dominance Decisions and Random Initial Wealth with Applications to the Economics of insurance." Journal of Risk and Insurance **62**(1): 30-49.
- Brumelle, S. L. (1974). "When Does Diversification Between Two Investments Pay." Journal of Financial and Quantitative Analysis: 473-83.
- Cummins, J. D. and O. Mahul (2004). "The Demand for Insurance With an Upper Limit on Coverage." Journal of Risk and Insurance **71**: 253-264.
- Doherty, N. A. (1984). "Efficient Insurance Buying Strategies for Portfolios of Risky Assets." Journal of Risk and Insurance(51): 205-224.

Mossin's Theorem Given Random Initial Wealth

- Doherty, N. A. and H. Schlesinger (1983). "The Optimal Deductible for an Insurance Policy When Initial Wealth is Random." Journal of Business **56**(4): 555-565.
- Doherty, N. A. and H. Schlesinger (1983). "Optimal Insurance in Incomplete Markets." Journal of Political Economy **91**(6): 1045-1054.
- Lehmann, E. L. (1966). "Some Concepts of Dependence." Annals of Mathematical Statistics **37**: 1137-1153.
- MacMinn, R. D. (1984). "A General Diversification Theorem: A Note." Journal of Finance **39**(2): 541-50.
- Mayers, D. and C. W. Smith Jr. (1983). "The Interdependence of Individual Portfolio Decisions and the Demand for Insurance." Journal of Political Economy **91**(2): 304-311.
- Mossin, J. (1968). "Aspects of Rational Insurance Purchasing." Journal of Political Economy **76**(4): 553-568.
- Rothschild, M. and J. E. Stiglitz (1970). "Increasing Risk: I. A Definition." Journal of Economic Theory **2**: 225-43.
- Rubinstein, M. (1976). "The Valuation of Income Streams and the Pricing of Options." The Bell Journal of Economics and Management Science **7**(Autumn): 407-425.
- Samuelson, P. A. (1967). "General Proof That Diversification Pays." Journal of Financial and Quantitative Analysis **2**(2): 1-13.
- Schlesinger, H. (1981). "The Optimal Level of Deductibility in Insurance Contracts." Journal of Risk and Insurance **48**: 465-481.
- Schlesinger, H. (2000). The Theory of Insurance Demand. The Handbook of Insurance. G. Dionne. Boston, Kluwer Academic Publishers: 131-151.
- Schlesinger, H. (2005). "Mossin's Theorem for Upper-Limit Insurance Policies." Journal of Risk and Insurance.
- Schlesinger, H. and N. A. Doherty (1985). "Incomplete Markets for Insurance: An Overview." Journal of Risk and Insurance **52**(3): 402-423.

Mossin's Theorem Given Random Initial Wealth

Schulenburg, J. M., Graf v. d (1986). "Optimal Insurance Purchasing in the Presence of Compulsory Insurance and Insurable Risks." Geneva Papers on Risk and Insurance: 5-16.

Wright, R. (1987). "Expectation Dependence of Random Variables, With an Application in Portfolio Theory." Theory and Decision **22**(2): 111-124.