

**THE POSSIBILITY OF A PRIVATE CROP INSURANCE MARKET:  
THE THEORETICAL FOUNDATIONS**

Joseph L. Krogmeier

H. Holly Wang

May, 1998

Prepared for AAEA Selected Paper

The authors are, respectively, a self-employed farmer (former graduate student) and an assistant professor at the Department of Agricultural Economics, Washington State University.

## **Abstract**

The theoretical foundation of risk pooling in insurance has heavily depend on the independence assumption of losses, which is severely violated in crop insurance. A weaker condition, asymptotic nonpositive correlation can also lead to risk pooling and is satisfied by yield losses. Therefore, private insurance and reinsurance markets may work.

# **THE POSSIBILITY OF A PRIVATE CROP INSURANCE MARKET: THE THEORETICAL FOUNDATIONS**

## **Introduction**

In the wake of the historic 1996 farm legislation, mandating a decoupling of gradually diminishing government support from agricultural commodity price levels, the need exists for more thoroughly understanding the altered agricultural risk environment. Recent attention has focused on the potential role in farm income stabilization of federally-reinsured crop and revenue insurance programs. Given the history of actuarial problems with federal crop insurance programs (Knight and Coble; Skees, Black and Barnett; Goodwin and Smith; Wright) and the renewed interest among researchers and policy makers in modern variants of crop insurance as an income stabilization tool, it is prudent to carefully review the theoretical foundations of such insurance. The purpose of this paper is first to review these foundations; second to consider a weaker statistical condition than the usually believed condition, independency, on the random losses for effective risk pooling; and finally, in light of this condition, to briefly discuss the potential for a privatized crop insurance market.

## **Theoretical Foundations of Insurance**

The institution of insurance has evolved in modern economies as one means of risk shifting in the face of uncertainty (Arrow, chapter 5). The primary function of insurance in this regard is risk pooling. Mehr, Cammack and Rose (p.32) offer the following definition, “[i]nsurance may be defined as a device for reducing risk by combining a sufficient number of exposure units to make their individual losses collectively predictable.” In what follows, we will consider two different types of insurers’ risk, *relative* and *absolute*. We will see that the first type of risk is indeed generally decreased via combining or pooling exposure units, but the second type is not. In pursuing this discussion we will consider the statistical foundations of insurance first, drawing upon the illuminating framework provided by Cummins.

Consider the following model of an insurance pool:

$$(1) \quad S_N = \sum_{i=1}^N X_i ,$$

where  $S_N$  is the total loss (claims) of the pool in a given period of time,  $X_i$  is the loss experienced by the  $i^{\text{th}}$  exposure unit, and  $N$  is the number of exposure units in the pool. In the context of agricultural insurance, an exposure unit may be a particular farm or parcel of land. In the model given by equation (1), each individual loss is conceptualized as a random variable and the total loss experienced by the pool is random as well. In an agricultural context, the loss is most likely defined as a production shortfall from some prespecified level. If, for convenience, we assume the loss distributions of all exposure units are identical with mean  $\mu$  and variance  $\sigma^2$ , the expected total loss of the pool is:

$$(2) \quad E(S_N) = N\mu$$

and the variance of the total loss of the pool is:

$$(3) \quad \text{Var}(S_N) = N\sigma^2 + 2\sigma^2 \sum_{j=2}^N \sum_{i=1}^{j-1} \rho_{ij} ,$$

where  $\rho_{ij}$  is the correlation between the  $i^{\text{th}}$  and  $j^{\text{th}}$  exposure units.

To be of much use, additional information regarding the distribution of the  $X_i$ 's must be included in the model. Elementary discussions of risk pooling often assume the  $X_i$ 's are independently and identically distributed (i.i.d.). While this set of distributional assumptions allow application of the simplest version of a law of large numbers, it is unnecessarily strong and generally not realistic (Bühlmann). Identically distributed exposure units may generally be desirable, but this condition is by no means a necessary *statistical* condition for effective pooling. Cummins contends that a more compelling argument for homogeneity as a necessary condition for insurability involves information asymmetries between insurers and insureds. Our focus remains on the statistical foundations of insurance.

Similarly, independence is not a necessary condition for the application of more general forms of laws of large numbers and thus for potentially effective risk pooling. In this initial stage of our discussion, however, we shall begin with the simplest case, i.i.d. losses, which simplifies equation (3) since  $\rho_{ij} = 0$  for all  $i \neq j$ . We will then present the more realistic case, especially in the context of agricultural insurance, in which we assume neither identical (except for notational convenience) nor independent random losses. A particular form of statistical dependency needs, then, be assumed.

The collection of theorems known as the Weak Law of Large Numbers (WLLN) stands as one of the fundamental propositions of the theory of probability, and as such provides the statistical foundation for the critical function of risk pooling. As Bühlman (p.33) explains “[i]t is due to the convergence of the average claim toward a fixed quantity that we can expect--for large numbers--to offset the effects of chance (claim amounts) by fixed quantities (premiums).” We first consider the most basic WLLN result:

*Weak Law of Large Numbers (i.i.d. random variables).* Let  $\{X_i\}_{i=1}^N$  be a sequence of independently and identically distributed random variables, and suppose the expected value of  $X_i$  is finite for all  $i$  ( $EX_i = \mu < \infty, \forall i$ ). Then the following holds

$$(4) \quad \lim_{N \rightarrow \infty} Pr[|\bar{x}_N - \mu| < \varepsilon] = 1 \quad \forall \varepsilon > 0$$

where

$$\bar{x}_N = \frac{1}{N} \sum_{i=1}^N X_i.$$

Within the insurance context, this WLLN result states that if we assume all random losses in the insurance pool are independent and identically distributed, the average loss realized will be arbitrarily close to the true mean of the common loss distribution with probability approaching 1 as the size of the pool approaches infinity. It would then seem that if the true mean of the loss function could be accurately estimated from past loss experience; beginning each insurance period this amount (the net premium) could be collected from every insured thus ensuring the insurance pool would have sufficient funds available to

---

<sup>1</sup> A capital letter denotes a random variable and its lower case for a particular outcome.

pay any realized indemnities. This is essentially what risk pooling is all about, although as we shall see immediately below (equation (5)) it is not the whole story.

Equation (4) is the most basic result of the WLLN. The same type of convergence, however, will occur under more general conditions than i.i.d. random variables. A sufficient condition for the basic result of the law to hold is that the variance of  $\bar{X}_N$  goes to 0 as N goes to infinity (Mittelhammer, p. 262). WLLN has important implications for the behavior of an insurer's relative risk as well as for the determining of actuarially sound premiums. These implications, as well as the more general results alluded to above, will be discussed below.

Cummins stresses that the result of the WLLN does *not* imply the following

$$(5) \quad \lim_{N \rightarrow \infty} Pr\left[\left|\sum_{i=1}^N x_i - N\mu\right| < \varepsilon\right] = 1 \quad \forall \varepsilon > 0,$$

i.e., the probability that the total random loss of the pool becomes arbitrarily close to its expected value,  $N\mu$ , does not converge to 1 as the insurance pool grows large. This observation is the other side of the risk pooling story and is important to the understanding of insurers' absolute risk and the need for insurance buffer funds and risk loading of premiums.

While the WLLN is an important theoretical result, it is of more practical interest to consider insurance pools with a finite number of members. Assuming i.i.d. random losses, implying

$\bar{X}_N \sim (\mu, \sigma^2/N)$ , we can use Chebyshev's Inequality to write the following bound

$$(6) \quad Pr[|\bar{x}_N - \mu| < k\sigma/\sqrt{N}] \geq 1 - 1/k^2 \quad \forall k > 0,$$

i.e., a lower bound of  $1 - 1/k^2$  exists on the probability that the realized average loss will fall in the interval  $(\mu - k\sigma/\sqrt{N}, \mu + k\sigma/\sqrt{N})$ .

For some choice of  $k$ , which makes the lower bound on the probability given in (6) as close to 1 as we wish, the width of this "confidence" interval is determined by the standard deviation  $\sigma/\sqrt{N}$  of the distribution of

average loss. This type of reasoning leads Cummins to suggest the following measure of an insurer's relative risk (IRR)

$$(7) \quad IRR = \sigma/\sqrt{N}$$

or, simply, the standard error of the average loss. This type of insurer's risk, i.e., the risk that the realized average loss will be far from the mean of the loss distribution, decreases as the size of the insurance pool increases.

To define the concept of absolute risk, multiply the probability inequality in (6) by  $N$ , yielding

$$(8) \quad Pr\left[\left|\sum_{i=1}^N x_i - N\mu\right| < k\sigma\sqrt{N}\right] \geq 1 - 1/k^2, \quad \forall k > 0.$$

In this case, we see that the width of the interval which contains the realized total loss  $s_N$ , i.e.,  $(N\mu - k\sigma\sqrt{N}, N\mu + k\sigma\sqrt{N})$ , with a probability as close to 1 as we wish (depending upon our choice of  $k$ ) depends upon  $\sigma\sqrt{N}$ . As a consequence of (8), Cummins suggests the following definition of an insurer's absolute risk (IAR)

$$(9) \quad IAR = \sigma\sqrt{N}$$

i.e., the standard deviation of total loss of the pool. From (8), it is apparent that the probable deviation of the total loss of the pool from its expected value becomes increasingly large as the insurance pool increases, a condition captured by the concept of absolute risk.

It is imprecise to claim, common in the insurance literature, that risk becomes negligible in large insurance pools. From our discussion above, we see that risk is reduced in large pools in the relative sense but not in the absolute sense. This distinction between types of risk is important to the concept of a buffer fund as well as to the premium setting process of the insurance firm.

In the insurance literature, the event of an aggregate loss occurring (a realization of  $S_N$ ) which is so large as to deplete the insurance fund is captured by the concept of *ruin*. It has been suggested in the literature that a possible objective criterion for the management of an insurance pool is to minimize the probability of ruin in a given time period or perhaps maximize returns subject to maintaining a specified probability of ruin (Bühlmann). The premium surplus above the expected value of the aggregate loss ( $N\mu$  under homogeneity) required to maintain a particular probability of ruin is referred to as the *buffer fund*. Central limit theory, while limited in its practical relevance given the highly skewed loss distributions typically encountered in insurance problems, is useful in illustrating the concept of a buffer fund (Cummins).

Assuming i.i.d. random losses and that the approximation implied by the Lindberg-Levy Central Limit Theorem (Mittelhammer, p.270) is sufficiently precise, we can write

$$(10) \quad Pr\left[\left(\sum_{i=1}^N x_i - N\mu\right)/\sigma\sqrt{N} < y_\alpha\right] = Pr[(s_N - N\mu) < y_\alpha\sigma\sqrt{N}] = 1 - \alpha ,$$

where  $y_\alpha$  is a real number such that  $\phi(y_\alpha) = 1 - \alpha$ , and  $\phi$  is the standard normal distribution function. Thus to avoid ruin with probability  $1 - \alpha$ , the insurance fund must have a liquid buffer fund of the size  $y_\alpha\sigma\sqrt{N}$ .

The buffer fund is therefore a function of absolute risk and grows with the size of the pool. However, the buffer fund required per exposure unit,  $y_\alpha\sigma/\sqrt{N}$ , is a function of relative risk and declines with the size of the pool. Another way to view this is that since the size of the buffer fund is proportional to the *square root* of  $N$ ; as the size of the pool grows, the buffer fund amount allocated to each policy (buffer load) decreases.

Premium rate setting is approached in various ways in the insurance literature, but the discussion generally begins with the concept of “pure” or “net” premium (Hogg and Klugman; Borch; Goodwin and Smith). The net premium is simply the expected indemnity per exposure unit. The gross premium, the amount paid by the insured per exposure unit in order to be eligible for coverage, is larger than the net

premium by an amount referred to as the loading factor. We can examine the components of the loading factor by decomposing the gross premium in the following manner

$$(11) \quad P = P_N + A + L,$$

where  $P$  is the gross premium,  $P_N$  is the net premium,  $A$  is an administrative cost load, and  $L$  is the buffer load. An additional loading component to offset the risk of inaccurately estimating  $P_N$  might also be included in some instances. Assuming identically distributed losses and that there is no deductible or cap on the maximum indemnity, then it is clear that the expected indemnity equals the expected loss, i.e.,  $P_N = \mu$ .

The amount by which the gross premium exceeds the net premium ( $A + L$ ) is essentially a “risk premium”, and as such is an extremely important determinant of the demand for insurance. If  $A + L$  were equal to zero, leaving  $P = P_N$  then a risk neutral insured would be indifferent between paying the gross premium or facing the risk of loss. Since in practical applications  $A + L$  is some positive amount, and hence  $P > P_N$  it is generally assumed that the purchase of insurance is explained by risk aversion (Goodwin and Smith).

If we assume economies of scale in the administration function, then the administrative cost per exposure unit,  $A$ , will decline as the size of the insurance pool grows. Furthermore under the conditions above (i.i.d. random losses), the buffer load per exposure unit,  $L = y_\alpha \sigma / \sqrt{N}$ , will decline as the size of the pool increases. Thus for a sufficiently large insurance pool, the amount  $A + L$  will not be large, i.e., the risk premium a risk averse insured must pay to obtain coverage will be small. Under these conditions we would expect the insurance market to be viable.

From our discussion heretofore, it is apparent that the critical ingredient for risk pooling is that the variance of the average loss diminishes as the size of the insurance pool increases. Under this basic condition relative risk decreases as the size of the pool increases thus providing a statistical basis for predicting future

losses. Furthermore under this condition, the buffer load also declines as the pool grows larger, thus ensuring (assuming economies of scale in administration) that the gross premium will exceed the actuarially fair net premium by a relatively small amount. At this point we might ask ourselves what minimal conditions are required of the random losses to ensure  $var(\bar{X}_N) \rightarrow 0$  as  $N \rightarrow \infty$ ? Certainly i.i.d. random losses assures this; but, do weaker (more inclusive) conditions exist? The answer is yes. The section below examines one such condition, asymptotic nonpositive correlation, which has some intuitive appeal in the agricultural context.

### A Generalization: Asymptotic Nonpositive Correlation

In what follows we will explicitly relax the assumption of independence. Given the reliance of crop yields on large scale weather patterns, the assumption of independently distributed losses is particularly tenuous in the present context. In place of the independence assumption we will assume asymptotic nonpositive correlation (a.n.c.). Mittelhammer (p. 266) defines an a.n.c. random sequence as follows:

The sequence of random variables  $\{X_i\}_{i=1}^N$ , where  $var(X_i) = \sigma_i^2 < \infty \forall i$ , is said to be *asymptotically nonpositively correlated* if there exists a sequence of constants  $\{a_t\}_{t=1}^\infty$  such that  $a_t \in [0,1] \forall t$ ,  $\sum_{t=1}^\infty a_t < \infty$ , and

$$(12) \quad cov(X_i, X_{i+t}) \leq a_t \sigma_i \sigma_{i+t} \forall t > 0 .$$

The two conditions,  $a_t \in [0,1] \forall t$  and  $\sum_{t=1}^\infty a_t < \infty$ , imply that  $a_n \rightarrow 0$  as  $n \rightarrow \infty$ . The  $a_t$ 's represent upper bounds to the correlations between  $X_i$  and  $X_{i+t}$ . The definition implies that  $X_i$  and  $X_{i+t}$  cannot be positively correlated when  $t \rightarrow \infty$ . In the context of crop insurance, the natural way to order exposure units in a sequence is spatially so that  $t$  represents an ordinal measure of physical distance between exposure units.

The concept of a.n.c. seems quite sensible when applied to crop losses in an agricultural context. Under this assumption, for example, the correlation between the losses of two fields 500 miles apart will generally be less positive than the correlation between two adjacent fields. The definition places no lower bound on the amount of negative correlation between the losses of any two exposure units; although these are likely rare and thus largely irrelevant.

Assuming identically distributed losses for notational convenience, the variance of the average loss in the pool may be written

$$(13) \quad \text{Var}(\bar{X}_N) = \frac{1}{N^2} (N\sigma^2 + 2 \sum_{i=1}^{N-1} \sum_{t=1}^{N-i} \text{cov}(X_i, X_{i+t})) ,$$

where the  $X_i$  are ordered spatially and  $t$  represents an ordinal measure of physical distance. If we assume that losses are a.n.c. then using the relationship given in equation (12) we can place the following bound on the variance

$$(14) \quad \begin{aligned} \text{Var}(\bar{X}_N) &\leq \frac{1}{N^2} (N\sigma^2 + 2\sigma^2 \sum_{i=1}^{N-1} \sum_{t=1}^{N-i} a_t) \\ &= \frac{\sigma^2}{N} + \frac{2\sigma^2}{N^2} \sum_{t=1}^{N-1} (N-t) a_t \\ &\leq \frac{\sigma^2}{N} + \frac{2\sigma^2(N-1)}{N^2} \sum_{t=1}^{N-1} a_t \end{aligned} .$$

If we take the limit as  $N \rightarrow \infty$  of both sides of (14), recognizing that  $\lim_{N \rightarrow \infty} \sum_{t=1}^N a_t$  is bounded by our definition of a.n.c., then we see that  $\text{var}(\bar{X}_N) \rightarrow 0$  as  $N \rightarrow \infty$ . As mentioned above, that the variance of  $\bar{X}$  collapses as  $N$  goes to infinity is a sufficient condition for a WLLN result to obtain (see Mittelhammer, Theorem 5.22). If we relax the notationally convenient assumption of identically distributed losses, then under these conditions we would still obtain the following WLLN result

$$(15) \quad \lim_{N \rightarrow \infty} \text{Pr}[|\bar{x}_N - \bar{\mu}_N| < \epsilon] = 1 \quad \forall \epsilon > 0$$

which is identical to the one in equation (4) except insofar as  $\mu$  has been replaced by  $\bar{\mu}_N = E(\bar{X}_N)$ , the mean of the associated sequence of expected values.

Using arguments similar to those used above, we see that under conditions of a.n.c. an insurer's relative risk (bounded from above by the square root of the bound given in (14)) declines as the size of the

insurance pool grows. Similarly, the insurer's absolute risk will increase as the size of the pool expands; but the buffer load needed to counter this risk will decrease with pool size.

### **The Potential for a Private Crop Insurance Market**

Now it is usually believed that private market for crop insurance will not work because the indemnity is not independent. A leading article by Miranda and Glauber concludes that the major reason of crop insurance market failure is systematic risk, the dependency (positive correlation) of farm yield loss. Their measure of systematic risk is defined by the ratio of the coefficient of variation of total indemnities paid to the coefficient of variation *if* indemnity payments were independent as:

$$(16) \quad \frac{R_C}{R_I} = \frac{\sqrt{\text{Var}(\sum x_i)}}{E(\sum x_i)} \bigg/ \frac{\sqrt{\sum \text{Var}(x_i)}}{E(\sum x_i)} = \sqrt{\frac{\text{Var}(\sum x_i)}{\sum \text{Var}(x_i)}}$$

The high ratio calculated from the crop insurance industry leads them to this conclusion.

It should be noted that first, this measure of systematic risk is in the sense of absolute risk; second, the systemic risk is bounded as the number of insureds increases; and third even if the systemic risk is still much higher than 1 after pooling a large number of insureds, the total risk faced by the insurer should decrease under a.n.c which can be seen by the coefficient of variation itself (in either absolute or relative sense):

$$(17) \quad R_C = \frac{\sqrt{\text{Var}(\sum x_i)}}{E(\sum x_i)} \leq \frac{\sigma}{\mu} \sqrt{\frac{1}{N} + \frac{2(N-1)\sum a_i}{N^2}}.$$

We have demonstrated above that the critical statistical condition determining the viability of an insurance pool is the behavior of relative risk as the insurance pool grows and not the behavior of absolute risk. From their work, although a positive contribution to the literature, we cannot conclude that a privatized crop insurance market is doomed to fail.

The critical issue, in terms of the viability of risk pooling, is the rate at which relative risk diminishes as the pool size grows as compared to the ideal case of i.i.d. random losses. The speed of convergence of relative risk to zero will in turn be determined by how quickly the sequence of bounds on the correlations between losses, i.e.,  $\{a_i\}_{i=1}^N$ , converges to zero. This is essentially an empirical question and may have important implications for the actuarial soundness of crop insurance programs and for the optimal trading of risk in a reinsurance market.

Comparing to the case of property insurance, in which losses are also correlated due to natural hazards such as hurricanes<sup>2</sup>, we might expect the existence of a private agricultural reinsurance market in which risks are traded at the national and/or international level at which yield losses for various crops and regions are hardly correlated.

Since yield losses can be assumed independent (or near independent) across time, yield risks might also be pooled over time through cumulating premium surplus, and/or multi year contracts. Suppose the coefficient of variation in (17) is for one year, and then the coefficient of variation for T years will be smaller as:

$$(18) \quad R_C^T = \frac{\sqrt{\text{Var}(T \sum x_i)}}{E(T \sum x_i)} = \frac{R_C}{\sqrt{T}}$$

Trading in an options or futures market will further facilitate the reinsurance market, such a futures market for crop yields have been available in Chicago Board of Trade since recently.

## Conclusion

We have reviewed the statistical foundations of insurance in this article, and demonstrated that a necessary condition for effective risk pooling is not the loss independency, but a weaker condition,

---

<sup>2</sup>The authors are aware of the turbulence experienced by the property insurance market particularly in recent years (see Lewis and Murdock), and don't intend to argue that the effect of government intervention in this instance is necessarily negative.

asymptotically nonpositive correlation. The later condition may be satisfied in case of agricultural production, which implies the possibility of private agricultural insurance markets. Our contention is not that the effect of government intervention is necessarily negative, but that the possibility of a private agricultural insurance and reinsurance market is not precluded by the statistical nature of agricultural production, and this possibility should not be dismissed out of hand.

## REFERENCES

- Arrow, K.J. *Essays in the Theory of Risk-Bearing*. New York: North-Holland, 1974.
- Borch, K. *The Mathematical Theory of Insurance: An Annotated Selection of Papers on Insurance published 1960-1972*. Lexington, MA: D.C. Heath, 1974.
- Bühlmann, H. *Mathematical Methods in Risk Theory*. New York: Springer-Verlag, 1970.
- Cummins, D.J. "Statistical and Financial Models of Insurance Pricing and the Insurance Firm." *The Journal of Risk and Insurance*, 58(1991): 261-302.
- Goodwin, B.K., and V.H. Smith. *The Economics of Crop Insurance and Disaster Aid*. Washington, D.C.: The AEI Press, 1995.
- Hogg, R.V., and S.A. Klugman. *Loss Distributions*. New York: John Wiley, ????
- Knight, T.O., and K.H. Coble. "Survey of U.S. Multiple Peril Crop Insurance Literature Since 1980." *Review of Agricultural Economics*, 19(1997): 128-156.
- Lewis, C.M. and K.C. Murdock. "The Role of Government Contracts in Discretionary Reinsurance Markets for Natural Disasters." *The Journal of Risk and Insurance*, 63(1996): 567-597.
- Marshall, J.M. "Insurance Theory: Reserves versus Mutuality." *Economic Inquiry*, 12(1974): 476-492.
- Mehr, R.I., E. Cammack, and T. Rose. *Principles of Insurance*, 8<sup>th</sup> ed. Homewood, Illinois: Richard D. Irwin, Inc., 1985.
- Miranda, M.J. and J.W. Glauber. "Systematic Risk, Reinsurance, and the Failure of Crop Insurance Markets." *American Journal of Agricultural Economics*, 79 (1997): 206-215.
- Mittelhammer, R.C. *Mathematical Statistics for Economics and Business*. New York: Springer-Verlag, 1996.
- Skees, J.R., J.R. Black, and B.J. Barnett. "Designing and Rating an Area Yield Crop Insurance Contract." *American Journal of Agricultural Economics*, 79(May 1997):430-438.
- Wright, B.D. "Public Insurance of Private Risks: Theory and Evidence from Agriculture." In *Government Risk-Bearing: Proceedings of a Conference Held at the Federal Reserve Bank of Cleveland, May 1991*, ed. M.S. Sniderman, pp. 45-65. Boston: Kluwer Academic Publishers, 1993.