PN-ABG-857
ISN 69513

Modeling the Demand for Crop Insurance

Behjat Hojjati
Nancy E. Bockstael

Reprinted from *Multiple Peril Crop Insurance: A Collection of Empirical Studies*, edited by Harry P. Mapp
Southern Cooperative Series Bulletin No. 334
Oklahoma State University, 1988

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
1776 Massachusetts Avenue, N.W.
Washington, D.C. 20036 1998 U.S.A.

Reprint No. 187

MODELING THE DEMAND FOR CROP INSURANCE

Behjat Hojjati and Nancy E. Bockstael

Since 1982 with the phasing out of the Disaster Payment Program, crop insurance has become the principal federal program for protecting producers against the adverse effects of yield variability. Farmer participation in the program, however, has historically been low. Despite a high subsidy on the premium, a large number of those who seem to need yield protection do not voluntarily participate.

In order to better understand farmers' resistance to the program, it would be useful to model, empirically, the factors affecting farmer participation. Understanding what motivates participation in crop insurance would allow policy makers to predict farmers' responses to potential changes in the program. While some preliminary research has been done in this area (e.g. Gardner and Kramer, 1986), it tends to be of an aggregate nature and is not based on a model of farmer's behavior.

The purpose of this study is to explore the factors which affect a farmer's decision to participate in the crop insurance program. A model of farmers' demand for crop insurance is constructed and estimated. Further, an attempt is make to use the resulting information about farmers' behavior to learn something about farmers' risk preference. Knowing the risk attitude of farmers can also provide valuable information for policy design.

Crop insurance is one of several risk management strategies that have been designed to reduce the variability of farm income. Specifically, the Federal Crop Insurance program is designed to aid farmers in minimizing their crop production risks. The principal type of insurance offered by the program is "All Risk" crop insurance, which provides protection against adverse weather conditions and unavoidable losses caused by plant disease, flood, fire, etc. Farmers who participate in the program must insure their entire crop within each crop insurance unit. Each pays a premium for a guarantee of a certain percentage of the expected yield. If the actual yield falls below the guaranteed

Behjat Hojjati is an Agricultural Economist with International Food Policy Research Institute, Washington, D.C. Nancy E. Bockstael is an Associate Professor of Agricultural and Resource Economics, University of Maryland, College Park, Md. The research is based on Dr. Hojjati's Ph.D. dissertation.
This paper is Scientific Article No. A-4704, Contribution No. 7700, of the Maryland Agricultural Experiment Station.

yield, an indemnity is received equal to the difference in yield times a previously agreed upon price.

The farmer's demand for crop insurance arises from uncertainty due to uncontrolled natural hazards which affect crop production. Logically, the farmer's decision of whether or not to participate in crop insurance is likely to be influenced by a comparison of expected net gain (or loss) with and without insurance. Obvious elements entering into the calculation will be the size of the insurance premium and the level of protection. Additionally, there will be other factors more difficult to assess. Principal among them is the farmer's expectation (subjective probability) that the actual yield will fall below the guaranteed yield. This consideration had special significance during the period of the current study. During this period, guaranteed yield was calculated on the basis of average historical yields in a given geographical area, not average yields on each individual's farm. As a consequence, the crop insurance program was plagued with an adverse selection problem. Individuals whose yields had historically been below the average for the area would have a high subjective probability that their yield would fall below the guaranteed yield (even in a good year). In more recent years, a farmer guarantees borrowing power by participating in the insurance program. A crop insurance policy can be used as a collateral for borrowing money.

In what follows a preliminary model is developed to describe crop insurance and crop diversification choices. This joint decision has interesting features which complicate modeling and which have not been addressed in previous research. Two of these features receive particular emphasis in the subsequent analysis. The first is the importance of risk preference in crop diversification and crop insurance decisions. The second is the juxtaposition of discrete and continuous-type decisions. The roles of these two features are discussed in the next two sections.

Once the model is developed, it is estimated for a somewhat circumscribed problem--the case where there are, at most, two potential crops which are good economic alternatives. Because of the complexity of modeling the discrete/continuous choice, this relatively easy subproblem was chosen to test the usefulness of the model. Clearly the model needs to be broadened to make it applicable to a larger percentage of the farm sector.

Finally, throughout the analysis, attention is focused in yield variability. Variability of price is ignored. This certainly flies in the face of reality.

Introducing price variability into the model would complicate it further; this step remains to be taken.

The Conceptual Model

For exposition purposes, consider the case in which soil and weather conditions limit the crops which can be planted in a county to only two (e.g. corn and soybeans). This actually makes for four different uses to which farmers can put any one unit of their land--insured corn, uninsured corn, insured soybeans, and uninsured soybeans. Each of these "uses" has its own per acre expected profit and variance of profit. The problem would be relatively simple if the farmer were forced to choose one use for all farm land. The decision problem would be a discrete one in which expected utility (as a function of the distribution of profits) from each of the four uses would be compared. A discrete choice modeling framework would be sufficient for the analysis. Alternatively, if the problem were such that the farmer always chose an interior solution--i.e. a positive number of acres allocated to each of the four alternative uses--then the decision problem would be a continuous one. One might model the choice of number of acres in each crop/insurance combination using regression analysis. Appropriate constraints across equations might be necessary, of course, if a constraint on total acreage was imposed.

Unfortunately, the decision problem is complicated to analyze because a variety of "partial" corner solutions can exist. That is, acreage may be devoted to more than one use but fewer than all uses. This type of corner solution problem arises in a number of economic decision situations, but straightforward and entirely consistent methods for modeling it have not been developed (see Bockstael, Hanemann, and Strand, 1986; Kling, 1986).

One way of handling the problem is to model it in two pieces. First define the finite set of discrete "alternatives," each of which is a portfolio of crop/insurance "uses," and model the discrete choice among these alternative portfolios. Second, model the decision of how much acreage to allocate to each crop/insurance "use" in the chosen portfolio. The distinction between "alternatives" and "uses" is all important here. The former represents a portfolio of crop/insurance uses to which positive amounts of acreage are allocated.

An example will help in clarifying this distinction. Consider once again the case in which there are only two viable crops (e.g. corn and soybeans) available to the farmer who has a fixed amount of land. All acres may be planted in corn and insured, all acres may be planted in soybeans and insured, corn could be planted and left uninsured, soybeans could be planted and left uninsured, or any combination of the above. With two crops (e.g. corn and soybeans) and two insurance alternatives (yes and no), there are $2^4 - 1 = 15$ alternative portfolios, defined in this way. Strictly speaking, there is one more alternative, that of allowing some of the land to lie fallow. Depending on how the alternatives and the total acreage are defined, this could be an important "use" to include.

Each portfolio or "alternative" embodies between one and four crop/insurance "uses" for which positive amounts of acres are allocated. With more viable crops or more insurance choices, the number of alternative portfolios increases rapidly. In general, with n_1 crops and n_2 insurance choices, there are $2^{n_1 n_2} - 1$ discrete choices.

The farmer's decision involves a discrete choice among these alternatives or portfolios, but it also involves the continuous choice of how many acres to allocate to each crop/insurance use embodied in the chosen alternative or portfolio. For example, the discrete choice (chosen from the $2^n - 1$ alternatives available to a given farmer) might be to plant some uninsured acres in corn and some insured acres in soybeans. The continuous choice would be how many acres to plant in each. Clearly, the continuous choice is conditioned on the discrete one. What might be less obvious is that the discrete choice requires knowledge of the optimum solution to each (conditional) continuous choice. Once the form of expected utility is introduced below, this decision process can be given precise form.

Treatment of Risk

When considering the effect of risk preference on individuals' decisions, it is useful to employ the concept of expected utility. The individual's utility is a function of income (in this case, profits), but when decisions are made, this income is uncertain. Risk preferences will be reflected in the expected utility an individual associates with outcomes of differing levels of uncertainty.

In this analysis each individual's utility is assumed to be a function of expected profit and the variance of profit, so that decisions are based on a comparison of expected profit and variance of profit across alternatives. Higher moments of the distribution of profits may indeed be important to decision makers. Yet there is doubt as to how accurately individuals can assess these moments and how precisely researchers can evaluate individuals' assessments. Mean-variance analysis in empirical work has been supported by many researchers with justifications dating back to the early work of Markowitz and Freund. Additionally, Levy and Markowitz have shown that the mean-variance approach approximates situations in which the restrictive assumptions of normality and quadratic utility do not hold. The theoretical foundation of the mean-variance approach is that individuals maximize expected utility. Although critics of the expected utility theory have proposed alternative theories of behavior under risk, expected utility theory continues to be the dominant framework of empirical analysis and is employed here.

Our starting point is that utility is some (unknown) function of profit. However, profit is not fixed but has a stochastic distribution. Let us consider a Taylor series expansion of this (unknown) utility function, expanding around the mean value of profit. The Taylor series expansion of utility around the mean value of profit may be expressed as

$$(1) \quad U(\pi) = U(E\pi) + U_1(E\pi)\,(\pi\text{-}E\pi) + \frac{1}{2}[U_2(E\pi)\,(\pi\text{-}E\pi)^2] + \frac{1}{6}[U_3(E\pi)\,(\pi\text{-}E\pi)^3] + \ldots$$

where

$E\pi$ = expected value of profit, and

$U_i = \dfrac{d^iU}{d\pi^i}$.

Taking the expected value of (1) and recognizing that $E(\pi\text{-}E\pi) = 0$, expected utility can be written as

$$(2) \quad E[U(\pi)] = U(E\pi) + \frac{1}{2}[U_2(E\pi)\,M_2] + \frac{1}{6}[U_3(E\pi)\,M_3] + \ldots$$

where M_k is the k^{th} moment of profit about its mean.

Equation (2) suggests that expected utility can be expressed as a function of the expected value of profits and moments of the distribution of profits about the mean. Precisely what form the function should take or how many moments should be included in an actual estimation procedure is open to debate. As argued earlier, the higher the moment, the more difficulty the individual is likely to have in assessing it. Additionally, higher moment terms tend to add less and less to the precision of the Taylor series approximation. In the analysis which follows only the mean and variance are included linearly in the expected utility function. This is, by necessity, an arbitrary choice. Further work should assess the sensitivity of these results to the functional form and number of moments chosen.

For the model used below, expected utility has two arguments: the expected profit and variance of profit. Thus, for each of the crop/insurance alternatives described above, there will be an expression for expected utility. Individual i's expected utility associated with the j^{th} alternative is given by

$$(3) \quad E[U_j] = \bar{U}_{ij}(E\pi_{ij}, V\pi_{ij}) + \varepsilon_{ij}$$

where

$\bar{U}_{ij}$ = the systematic portion of individual i's expected utility from the j^{th} alternative,
$E\pi_{ij}$ = individual i's expected profit from the j^{th} alternative,
$V\pi_{ij}$ = individual i's variance of profit from the j^{th} alternative, and
ε_{ij} = unmeasured and unmeasurable factors which affect the expected utility of alternative j for individual i.

Note that given this simple construct, we would expect $\partial\bar{U}_{ij}/\partial E\pi_{ij}$ to be positive and $\partial\bar{U}_{ij}/\partial V\pi_{ij}$ to be negative if the individual is risk averse.

Modeling the Discrete/Continuous Choice

In this section the model of the farmer's participation decision in multiple peril crop insurance program is developed. The model incorporates both the discrete and continuous components of the decision and utilizes the expression for expected utility suggested above. The same basic model could be used with a number of different specifications of expected utility, as long as the expressions were linear in parameters. However, the more complex the expected utility function, the more difficult would be the solution of the continuous decision and the calculation of values for explanatory variables. This will become clear as we develop the model and outline the estimation procedures for our relatively simple expected utility formulation.

Recall the nature of the problem. An individual with n_1 crop possibilities and n_2 insurance possibilities can choose to allocate a fixed amount of land to any of the $2^{n_1 n_2} - 1$ sets of uses which we have called alternatives or portfolios of uses. The choice among these $2^{n_1 n_2} - 1$ alternatives will be made given an optimal allocation of land across uses. Thus if there are two crop possibilities [A and B and two insurance possibilities (yes and no)], the decision will be based on a comparison of the expected utility achievable from each of fifteen portfolios:

1) all land allocated to crop A insured,
2) all land allocated to crop B insured,
3) all land allocated to crop A uninsured,
4) all land allocated to crop B uninsured,
5) land allocated optimally between crop A insured and crop B insured,
6) land allocated optimally between crop A uninsured and crop B uninsured,
7) land allocated optimally between crop A insured and crop B uninsured, etc.

Given optimal allocations among uses, the individual will make a discrete choice among the $2^{n_1 n_2} - 1$ discrete alternatives. Of the several discrete choice models, the logit specification is used in this study because it handles multichotomous decisions quite well. A utility theoretic motivation for use of the multinomial logit model to analyze discrete choice problems has been developed by McFadden (1973). Following McFadden and using expression (3), the probability that the expected utility of the i^{th} individual from alternative j exceeds the expected utility from any other alternative can be expressed as

$$(4) \quad P_{ij} = \Pr\{[\bar{U}_{ij}(E\pi_{ij}, V\pi_{ij}) + \varepsilon_{ij}] \geq [\bar{U}_{ih}(E\pi_{ih}, V\pi_{ih}) + \varepsilon_{ih}]\}$$

for all h in the alternative set. This probability can be rewritten as

$$P_{ij} = \Pr\{\varepsilon_{ih} - \varepsilon_{iy} \leq U_{ij}(E\pi_{ij}, V\pi_{ij}) - U_{ih}(E\pi_{ih}, V\pi_{ih})\},$$

for all h, which is defined by the cumulative distribution of the stochastic variable $\eta_{ihj} = \varepsilon_{ih} - \varepsilon_{ij}$. The probability distribution of η will depend on assumptions about the distribution of the ε's. If ε is assumed distributed as a Weibull, then η is logistically distributed and the modelling approach is labelled logit analysis. (A similarly shaped sigmoid distribution function for η arises if ε is assumed normally distributed, resulting in a probit analysis). Here we assume a Weibull

distribution for the ε's (which substantially facilitates computation) as well as a $\bar{U}$ function linear in parameters. With these assumptions, the probability of individual i choosing alternative j is

$$(5) \quad P_{ij} = \frac{\exp[\beta_0 + \beta_1 E\pi_{ij} + \beta_2 V\pi_{ij}]}{\sum_h \exp[\beta_0 + \beta_1 E\pi_{ih} + \beta_2 V\pi_{ih}]}$$

where

P_{ij} = the probability that the i^{th} individual chooses the j^{th} alternative,

$\bar{U}_{ih}$ = the non-stochastic parts of the i^{th} individual utility function associated with the h^{th} alternative,
$E\pi_{ih}$ = the i^{th} individual's expected total profits from the h^{th} alternative,
$V\pi_{ih}$ = the i^{th} individual's variance of total profit from the h^{th} alternative,and
β_0, β_1, and β_2 = the unknown parameters.

Unless we specify that β_o will vary over alternatives, i.e. that certain alternatives are more likely to be chosen, ceteris paribus, then β_o cannot be recovered in the estimation process. However β_1 and β_2 can be estimated, and we do so using maximum likelihood procedures.

If it were not for the continuous dimension to the choice, the discrete choice would be easy to assess. However, the interaction of the continuous and discrete decision complicates the expression for the farmer's expected profit ($E\pi_j$) and variance of profit ($V\pi_j$) for any given alternative, because each alternative represents a set of uses where the acres are optimally allocated among uses. The actual expected profit of a portfolio of uses will be the sum of expected profits per acre from each use, weighted by the acres allocated to the use. In general, expected total profits is given by

$$(6) \quad E\pi_j = \sum_{k \in K_j} A_k E\delta_k$$

and variance of total profits is given by

$$(7) \quad V\pi_j = \sum_{k \in K_j} A_k^2 V\delta_k + \sum_{k \in K_j} \sum_{g \in K_j} A_k A_g E(\delta_k \delta_g)$$

where

$E\pi_j$ = total expected profit from the j^{th} portfolio,
$V\pi_j$ = variance of total profits associated with the j^{th} portfolio,
A_k = number of acres allocated to the k^{th} use,
K_j = the set of uses in portfolio j,
$E\delta_k$ = per acre expected profits associated with the k^{th} use,
$V\delta_k$ = variance of per acre profits associated with the k^{th} use, and
$E(\delta_k\delta_g)$ = covariance of per acre profit between the k^{th} and g^{th} uses.

Expected utility depends on the covariance between per acre profits in different uses. In general, if the profits associated with different crops are either negatively correlated or not very correlated, diversification could reduce risk. However, if there is perfect positive correlation between profits, diversification will do nothing to eliminate risk. Crop insurance becomes a more attractive risk management strategy as the correlation between crops increases.

Now we can see the difficulty with direct application of discrete choice analysis. The $E\pi_j$ and $V\pi_j$ in (5) are quite complicated and depend on optimal values of acreage allocations to each use (the A_k's in expressions (6) and (7)). Thus the farmer is assumed to compare expected utility of portfolios where the expected utility of the j^{th} portfolio is given by

$$(8) \qquad E[U_j] = \max_{A_1,\ldots,A_{K_j}} [\beta_0 + \beta_1 (\sum_{k\in K_j} A_k E\delta_k) + \beta_2 (\sum_{k\in K_j} A_k^2 V\delta_k + \sum_{k\in K_j}\sum_{g\in K_j} A_k A_g E(\delta_k\delta_g)) + \varepsilon_j]$$

where $\sum_{k\in K_j} A_k = \bar{A}$ and $\bar{A}$ is the total number of acres available to the individual.[1]

In order to estimate the discrete choice model, it is necessary to obtain values for $E\pi_j$ and $V\pi_j$, but these depend on the solution for the optimal acreage allocation for each possible portfolio of uses. The form of the functions for the optimal acreage allocation depends on the number of non-zero uses embodied in the portfolio. When the portfolio involves only one use, the constraint implies $\bar{A}$ goes to that use. When the portfolio involves two uses (e.g. uninsured soybeans (SU) and insured corn (CI)), then maximizing (8) subject to the constraint implies

$$(9) \qquad A_{CI} = \frac{\beta_1 [E\delta_{CI} - E\delta_{SU}] + \beta_2 [\bar{A}\ (V\delta_{CI}) - E(\delta_{CI}\delta_{SU})]}{\beta_2 [2(V\delta_{CI} + V\delta_{SU}) - 2E(\delta_{SU}\delta_{CI})]}$$

and

$$A_{SU} = \bar{A} - A_{CI}$$

[1]Prior to 1982 the guaranteed yields were calculated with respect to average area yield. Beginning 1982, an Individual Yield Coverage Program (IYCP) was designed to offer farmers who produce higher yields more protection. However, for 1982, less than one percent of the potential farmers selected IYCP. Since this study is based on data for 1982, any changes which have been made since 1982 do not affect the development and the result of the model in the later sections.

where

A_{CI} = total acres devoted to insured corn, and
A_{SU} = total acres devoted to uninsured soybeans.

The optimal acreage allocation functions will take different forms for different numbers of uses in the portfolio. Even though the arguments in the acreage allotment functions are defined differently for different portfolios, the parameters are, however, identical. Thus, the acreage allotment functions can be pooled across observations and non-linear maximum likelihood estimation procedures used to estimate the common parameters, β_1 and β_2.

Each of the above estimation procedures (the discrete choice model and the continuous model) provides a means of estimating the unknown parameters, β_1 and β_2. Ideally, the two-choice problems should be estimated simultaneously to maximize efficiency. However, techniques were not readily available to accomplish this, and an iterative procedure was employed which nonetheless used information from both types of decisions. This procedure is outlined in the estimation section.

The Data and Calculation of Relevant Variables

The theoretical model developed above is micro-economic in nature in the sense that it is derived from postulates of individual behavior. Hence, for the purpose of estimation, the ideal data are microdata, i.e. data on individual farms. However, microdata are not readily available. By necessity, aggregate data are used in this study to demonstrate the model. The data are on a county basis and include information on yield, guaranteed yield, premium, market price, indemnity price, potential acres, and acres insured. In terms of the model, each county is being treated as if it were a single farm. While greater disaggregation is preferable in the abstract, aggregation to the county level does mitigate some of the problems which arise because of the nature of guaranteed yield calculations and because of unmeasurable differences in individual farm situations.

The data consist of a cross-sectional sample of 140 counties in the states of Georgia, Illinois, Indiana, Kentucky, Maryland, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Virginia, and West Virginia. The sample was restricted to counties which for the most part have insured no more than two different crops over the last decade. The two-crop limit served to reduce the number of portfolios available, yet allowed demonstration of the nature of the choice model.

According to FCIC personnel, the overwhelming majority of farmers who participate in the crop insurance program choose the highest coverage level and highest guarantee price. This is borne out by the observed sample of farmer participation decisions. For the purpose of estimation, we consider only two insurance options: a) not insuring and b) insuring with the highest coverage level of insurance protection and highest price protection. The insurance decision is to insure at 75 percent or not at all.

Even with two crops and two insurance alternatives, there are fifteen possible portfolios. While somewhat restrictive, this limited problem still allows us to accomplish the objective of demonstrating a model which captures both the discrete and continuous nature of the farmer's decision under uncertainty.

For each county and each crop, yearly data on yield per acre from 1967 to 1982 were taken from USDA Crop Reporting Service bulletins. Data on price received by farmers for each crop were also obtained from the same source. Data on potential acres, acres insured, and indemnities were obtained from the Federal Crop Insurance Office in Kansas City, Missouri. Among the total sample 43 percent had yields below the guarantee level and received indemnity payments in 1982. The loss ratio (indemnity/premium) ranged from zero to 15.15 across counties in 1982.

The model developed above requires an approximation of an individual's perception of expected profit and variance of profit associated with each use ($E\delta_k$, $V\delta_k$) as well as covariances between uses [$E(\delta_k\delta_g)$]. The expected profit and variance of profit, in turn, depend on the expected yield and variance of yield of each use in the portfolio. Since farmers' perceptions of the distribution of yields are likely to be affected by their past crop production experiences, county-level data on yield per harvested acre from 1967 to 1982 were used in calculating proxies for expected yield and variance of yield. It should be noted here that yield per harvested acre (the only data readily available) is not the same as yield per planted acre, and it is the later that the FCIC insures. Consequently, subjective probabilities calculated on the basis of yield per harvested acre will be optimistic (biased upward). Marginal and joint density functions of yield based on this historical data were determined for each use in each county in order to calculate the expressions for the expected profits and variance of profits for each portfolio.

In this study it is assumed that yields are distributed approximately normally.[2] Since marginal distributions are assumed to be approximately normal, the joint distribution is treated as bivariate normal. Additionally, in the calculation of expected profit and variance of profit, the insurance premium is the only variable cost subtracted, because data were not available on operating costs.

The exact procedure for calculating the expected values, variances, and covariances of per acre profits for different uses is presented in the Appendix. These formulas were applied to historical data to generate the necessary explanatory variables. It is interesting to note that when these calculations are made from actual data, uses which reflect non-participation in crop insurance tend to exhibit higher expected profits and higher variance of profits. Crop insurance appears to reduce risk but lower mean profits, suggesting that the decision to participate will depend heavily on risk preferences.

Estimation Results

Since the farmer's participation decision and acreage allocation decision are generated by the same utility maximization problem, the logit model (expression (5)) and the continuous choice (expressions such as (9)) contain the same parameters. If these models were estimated independently, two sets of estimates of the same ß's would be obtained. Alternatively, if simply the continuous choice problem was estimated conditioned on the actual discrete choice, information would be lost and the coefficient estimates would not be efficient. The practical difficulties inherent in simultaneous estimation argue for a two-stage iterative estimation procedure in this preliminary investigation. That is, the estimated coefficients from one segment of the model are used as initialization parameters in the estimation of the other segment, and vice versa, until coefficient estimates converge. Lee and Trost and Heckman suggest a similar procedure.

In this analysis we must begin with the continuous model. The acreage allocation functions are pooled across observations and non-linear maximum likelihood methods used to estimate the common parameters. Then the estimates obtained from this procedure are employed as initial values in the discrete choice model. The parameter estimates from this stage are then used

---

[2] The Shapiro-Wilk statistic was used to test for normality. The null hypothesis that each county yield is distributed normally was not rejected.

to initialize the continuous choice model, and so on. Unfortunately, this does not ultimately guarantee identical parameter estimates in both the continuous and discrete stage. Only simultaneous estimation would do that. However, when convergence criteria are met in both estimations, the final sets of estimated parameters are quite close. The reported results in this section are the estimates after convergence.

The results of the last round estimation from the system of acreage allocation equations (the continuous choice) are presented in Table 1. A chi-squared test on the likelihood ratio statistic is used to test the hypothesis that all parameters are equal to zero.

The calculated statistic is significant at the 1 percent level of significance. Thus, the null hypothesis that both coefficients are equal to zero is rejected with 99 percent confidence. The estimated coefficients are also individually significantly different from zero at the 95 percent confidence level (indicated by the t-statistics) and have the expected sign.

The discrete choice of farmer's participation behavior was estimated using the logit discrete choice model (expression (5)). The results of the estimation are given in Table 2.

Two statistical tests are used in the analysis of the multinomial logit. One is the likelihood ratio statistic. The likelihood ratio statistic at two degrees of freedom is 204.4703. The distribution with two degrees of freedom at the 99 percent confidence level is 9.21034. Thus, the null hypothesis that all coefficients are jointly equal to zero is rejected at the 0.01 level of significance.

The other test statistic which is useful in describing the explanatory power of the multinomial logit is

$$(10) \quad \rho^2 = \left[1 - \frac{L(\beta_0)}{L(\hat{\beta})}\right]^{n/2}$$

The value of ρ^2 is .99 and indicates that the model explains the major portion of the variation.

Another "goodness of fit" measure sometimes employed with discrete choice models is the "percent correctly predicted." This statistic depends on a particular definition of "correct prediction." The observation is considered

Table 1. Estimation Results (Continuous Choice)[a]

Variable	Coefficient	t-statistic	Test	Statistic	Degrees of Freedom
Expected profit	.46967	1.84			
			$-2 \log \left[\frac{L(\beta_0)}{L(\hat{\beta})}\right]$	2089.62	2
Variance of profit	-.8330	-5.4			

[a] In the estimation of the continuous model, the number of observations (580) increases by the number of chosen alternatives. For example, if in a county four alternatives are chosen (i.e. plant corn, plant soybeans, plant corn and insure, plant soybeans and insure) four observations are considered, and so on. Number of observations = 2380.

Table 2. Estimation Results (Discrete Choice)

Variable	Coefficient	t-statistic	Test	Statistic	Degrees of Freedom
Expected profit	.5038	1.98			
			$-2 \log \left[\frac{L(\beta_0)}{L(\hat{\beta})}\right]$	204.4703	2
Variance of profit	-.5829	-8.61			

Percent predicted correctly = 38.1; p2 = .99; observations = 140.

correctly predicted if the alternative with the highest predicted probability of being chosen is actually the observed choice. By this standard, if the multinomial logit predicts that, for a given observation, the utility of alternative j is the highest of all available alternatives' probabilities and the individual is observed to choose the j^{th} alternative, then the prediction for this observation is considered correct. The percentage of cases "correctly predicted" by this criteria is 38.1 percent. This is rather encouraging. With fifteen alternatives, the odds of predicting correctly with no information is 7 percent. It is also encouraging because the measure underestimates the value of a model. The "percent predicted correctly" statistic gives no weight (as regression techniques implicitly do) to being close. If the actual choice is the one predicted to be the second best by the model, the observation is still considered incorrectly predicted.

The estimated parameters of the logit analysis are both significantly different from zero at the 95 percent confidence level and have the expected sign. The positive coefficient on the expected profit indicates that as the expected profit of alternative j increases relative to other alternatives, the probability of choosing alternative j increases relative to other alternatives. On the other hand, the negative coefficient on variance indicates that farmers prefer alternatives with less variation, other things being equal. In other words, farmers are risk averse.

In qualitative terms, the results suggest, as expected, that any changes in the Federal Crop Insurance policy toward enhancing expected profit and reducing variance of profits should increase the rate of participation in the crop insurance program. The model indirectly shows the effect increased participation in crop insurance might have on crop diversification. If policies were undertaken to make crop insurance participation more appealing, the resulting increases in participation might lead to changes in the crop mix in areas where diversification was used as a risk managing tool.

Unlike regression coefficients, the estimated coefficients of a discrete choice model are difficult to interpret intuitively. From (5) we see that the predicted probability of choice j is given by

$$(11) \quad \hat{P}_j = \frac{\exp\,(\hat{\beta}_1 E\pi_j + \hat{\beta}_2 V\pi_j)}{\sum_k \exp\,(\hat{\beta}_1 E\pi_k + \hat{\beta}_2 V\pi_k)} \,,$$

so that the predicted ratio of the odds of choosing j relative to some other choice m are

$$(12)\quad \frac{\hat{P}_j}{\hat{P}_m} = \frac{\exp\,(\hat{\beta}_1 E\pi_j + \hat{\beta}_2 V\pi_j)}{\exp\,(\hat{\beta}_1 E\pi_m + \hat{\beta}_2 V\pi_m)} \;.$$

The estimated coefficients relate the odds of choosing one alternative over another to the difference in the values of explanatory variables associated with the two alternatives. Our estimated model indicates that a one-dollar change in the difference between expected profits from alternative j and expected profits from alternative m would produce a change in the log of the odds of choosing alternative j of approximately .5. That is

$$(13)\quad \frac{\partial \ln(\hat{P}_j/\hat{P}_m)}{\partial(E\pi_j - E\pi_m)} = \hat{\beta}_1$$

Put another way, a one-percent change in the difference in expected profits between two alternatives produces a percentage change in the odds of choosing those two alternatives equal to about .5 times the difference in expected profits,

$$(14)\quad \ln\,(\hat{P}_j/\hat{P}_m) = \hat{\beta}_1\,(E\pi_j - E\pi_m) + \hat{\beta}_2\,(V\pi_j - V\pi_m)$$

so that

$$(15)\quad \frac{\partial \ln(\hat{P}_i/\hat{P}_m)}{\partial \ln(E\pi_i - E\pi_m)} = \hat{\beta}_1\,(E\pi_i - E\pi_m) \approx .5\,(E\pi_i - E\pi_m)$$

Likewise, a one-percent change in the difference in variance of profits between two alternatives produces a percentage change in the odds of choosing those two alternatives equal to about -.58 times the difference in variance of profits.

The multinomial logit model can be used to predict the probabilities of choosing different alternatives given hypothetical changes in the economic decision environment. Predicted probabilities could be obtained by substituting into expression (11) the estimated parameters and values for the expected profits and variance of profits incorporating the hypothetical change. There are a number of ways in which these predicted probabilities can be used; the one chosen here (and frequently used in discrete choice analysis) is to interpret these predicted probabilities as predicted percentages of the aggregate.

It is in this role of prediction that the discrete choice model is particularly useful. One can postulate a policy which would alter the calculations of expected profit and variance of profit, as defined in the Appendix, and then predict the ultimate decisions - the percentage of acres which would be allocated to different uses. Of particular interest would be that percentage which would be insured given some hypothetical change in the crop insurance program.

There are many ways in which the crop insurance program could be made more attractive. The policy makers may be interested in predicting the impact of a major change in the program such as in premium rate, guaranteed yield coverage, etc. In a survey conducted by the General Accounting Office, farmers have cited inadequacy in coverage level as a reason for low participation. Consequently, we consider a hypothetical change in coverage level to demonstrate the predictive power of the model. In order to predict farmers' participation in crop insurance, expected profit and variance of profit were calculated by changing the maximum guaranteed yield from 75 percent to 85 percent. Over the counties in the sample, 4,038,909 acres were not insured in 1982. The model predicted that should the guarantee level be increased to 85 percent of average yield, acres insured would increase by 468,124, or about 12 percent of previously uninsured acreage.

Final Comments

The purpose of this paper was to explore a rather different model of crop insurance demand. The most crucial question in crop insurance policy is: Why do the vast majority of farmers choose <u>not</u> to participate in the program? One way to address this issue is to attempt to model, and therefore describe, the process by which farmers make their participation decisions. This preliminary paper tried to do that, taking into account some key aspects of the nature of the decision.

The first consideration, and the one which dictated the design of the model, is that the participation decision is interrelated with other decisions. The farmer's decision about crop insurance participation is intricately linked to crop choices. This aspect of the problem changes the decision from a simple dichotomous discrete choice (participate or not) to a more complex discrete continuous choice problem. Problems like this are scattered throughout economics but never seem to be handled very satisfactorily. In the analysis, an iterative solution to the interrelated discrete and continuous decisions is employed.

The second consideration is that of uncertainty. Since the issue is one of insurance, attitudes towards uncertainty obviously must play a role in decision making. Calculations of means and variances of profits using historical data suggest that crop insurance generally provides an alternative with a lower average profit but also a lower variance in profit, suggesting the importance of risk preference in participation decisions. Uncertainty has been a popular topic in economic research, yet no consensus has developed as to its appropriate treatment in empirical work. The approach used in this analysis is a simple mean-variance one, chosen to minimize the complications introduced into an already complicated model.

The paper emphasizes the complexity inherent in modeling the crop insurance participation decision when its interrelationship with crop decisions and its dependency on risk attitudes are both taken into account. It also provides an example of the useful sorts of predictions such discrete/continuous models can generate. Yet, the results are limited by the scope of the observations, the arbitrary choice of expected utility functions, the assumption of non-stochastic price, the less-than-optimal iterative estimation procedure, and the limited one-year planning horizon inherent in the decision framework. If these shortcomings were addressed, the general approach could provide especially useful information for policy makers.

References

Anderson, J. R., J. L. Dillon and J. B. Hardaker. *Agricultural Decision Analysis.* Ames: Iowa State University Press, 1977.

Arrow, K. J. *Essay in the Theory of Risk Bearing.* Chicago: Markham Publishing Co., 1971.

Bockstael, Nancy, W. Michael Hanemann and Ivar Strand. *Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models.* Report presented to the EPA under Cooperative Agreement CR-81143-01-1, Vol. II, 1986.

Boskin, M. J. "A Conditional Model of Occupational Choice." *J. Pol. Econ.* 82(1974):389-98.

Day, R. "Probability Distributions of Field Yields." *J. Farm Econ.* 47(1965):713-41.

Dillon, J. L. *The Analysis of Response of Crop and Livestock Production.* New York: Pergammon Press, 1977.

Domencich, T. A. and D. McFadden. *Urban Travel Demand: A Behavioral Analysis.* New York: North Holland Publishing Co., 1975.

Duncan, George M. "Formulation and Statistical Analysis of the Mixed, Continuous/ Discrete Dependent Variable Model in Classical Production Theory." *Econometrica* 48(1980):839-52.

Hanemann, W. M. "Discrete/Continuous Models of Consumer Demand." *Econometrica* 52(1984):541-61.

Hausman, J. S. and D. A. Wise. "A Conditional Probit Model for Quantitative Choice." *Econometrica* 46(1978):403-26.

Heckman, J. J. "Sample Selected Bias as a Specification Error." *Econometrica* 47(1979):153-61.

Kling, Catherine Louise. "Measuring the Recreational Benefits of Environmental Amenities Using Multiple Site Models: An Evaluation of Techniques." Ph.D. dissertation, University of Maryland, 1986.

Lee, L. F. and R. P. Trost. "Estimation of Some Limited Dependent Variable Models with Applications to Housing Demand." *Journal of Econometrics* 8(1978):357-82.

Levy, H. and H. Markowitz. "Approximating Expected Utility by a Function of Mean and Variance." *Amer. Econ. Rev.* 69(1979):308-17.

Maddala, G. S. *Econometrics.* New York: Wiley Publishing Co., 1977.

Maddala, G. S. *Limited Dependent and Qualitative Variables in Econometrics.* New York: Cambridge University Press, 1983.

Markowitz, M. H. *Portfolio Selection: Efficient Diversification of Investments.* New York: John Wiley and Sons, 1959.

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers of Econometrics*, ed. P. Zarembha. New York: Academic Press, 1973.

McFadden, D. "The Mathematical Theory of Demand Models," in *Behavioral Travel Demand Models*, pp. 305-26. Lexington, MA: D. C. Heath and Company, 1975.

Newbery, David M. G. and Joseph E. Stiglitz. *The Theory of Commodity Price Stabilization*. Oxford: Clarendon Press, 1981.

U.S. Congress. "Review of Federal Crop Insurance Program." Hearings Before the Subcommittee on Conservation, Credit and Rural Development of the Committee on Agriculture, 98th Congress, 1st Session, April 20, 1983.

U.S. Congress, Senate Select Committee on Investigation of Crop Insurance. "Investigation of Crop Insurance." Hearings, 67th Congress, 4th Session, April 24-27, 1953.

U.S. Department of Agriculture. "An Inside Look at All-Risk Crop Insurance." FCIC. March 1980.

Appendix

Calculation of Expected Per Acre Profits and Variance of Per Acre Profits for Each Use

To give numerical values to expected profits and variance of profits for portfolios, one must calculate expected profits per acre and variance of per acre profits associated with different uses. Let us first consider the per acre profits for an uninsured crop where we suppose that prices and operating cost are non-stochastic but yields are stochastic. Then per acre profits for the uninsured crop X (δ_{xu}) may be expressed as

$$(A1)\quad \delta_{xu} = p_x X - C_x$$

where

p_x = market price for crop X
X = yield per acre for crop X
C_x = per acre cost associated with crop X

Thus, the farmer's per acre expected profit can be written as

$$(A2)\quad E[\delta_{xu}] = p_x E(X) - C_x$$

where

$E[\delta_{xu}]$ = per acre expected profit associated with an uninsured crop
E(X) = expected value of yield (crop X)

The variance of profits per acre from an uninsured crop is

$$(A3)\quad var[\delta_{xu}] = p_x^2\ var(X)$$

where

$var[\delta_{xu}]$ = variance of profit per acre
var(X) = variance of yield

Since price is (unrealistically) assumed to be nonstochastic, variance of profit depends directly on the variance of yield.

The expressions for the expected value and variance of per acre profits are more complex when the use involves planting an insured crop. The per acre profit associated with this particular use, if yield is above the guaranteed level, is expressed as

$$(A4)\quad \delta_{xIA} = p_x X - (C_x + R_x) \qquad \text{if } x > \alpha\bar{X}$$

where

δ_{xIA} = per acre profit associated with an insured crop (X) when yield falls above the insured level ($\bar{X}$)

$\alpha\bar{X}$ = guaranteed yield coverage

C_x = per unit cost of production associated with crop X

R_x = per acre cost of premium associated with crop X

If the farmer participates in the Federal Crop Insurance program and his yield falls below the insured level, his per acre profit is

$$\text{(A5)}\quad \delta_{xIB} = p_x X + (\alpha\bar{X} - X)\bar{p}_x - (C_x + R_x) \qquad \text{if } X \leq \alpha\bar{X}$$

where

δ_{xIB} = per acre profit associated with insured crop (X) when yield falls below the insured level

$\alpha\bar{X}$ = guaranteed yield coverage

$\bar{p}_x$ = per acre indemnity price associated with crop X

Based on the expressions (A4) and (A5), the per acre expected profit for the participating farmer can be written as

$$\text{(A6)}\quad E[\delta_{xI}] = p_x \int_{\alpha\bar{X}}^{+\infty} X\, f(X)\, dX - \int_{\alpha\bar{X}}^{+\infty} (C_x + R_x)\, f(X)\, dX + p_x \int_{-\infty}^{\alpha\bar{X}} X\, f(X)\, dX$$

$$- \int_{-\infty}^{\alpha\bar{X}} (C_x + R_x)\, f(X)\, dX + \bar{p}_x \int_{-\infty}^{\alpha\bar{X}} (\alpha\bar{X} - X)\, f(X)\, dX$$

or

$$\text{(A7)}\quad E[\delta_{xI}] = p_x E(X) + \bar{p}_x \int_{-\infty}^{\alpha\bar{X}} (\alpha\bar{X} - X)\, f(X)\, dX - (C_x + R_x)$$

where

$E[\delta_{xI}]$ = per acre expected profit associated with planting crop X and insuring

$f(X)$ = the density function of yield

In this case the expected profit is not determined solely by the expected yield. It also depends on expected return from insurance.

The per acre variance of profit associated with this crop/insurance choice can be defined as

$$\text{(A8)}\quad \mathrm{var}[\delta_{xI}] = \Bigg\{ p_x^2\,\mathrm{var}(X) + \bar{p}_x^2 \int_{-\infty}^{\alpha\bar{X}} (\alpha\bar{X} - X)^2 f(X)\,dX + 2\bar{p}_x p_x \int_{-\infty}^{\alpha\bar{X}} X(\alpha\bar{X} - X)\, f(X)\,dX - \bar{p}_x^2 \left[\int_{-\infty}^{\alpha\bar{X}} (\alpha\bar{X} - X)\, f(X)\,dX \right]^2 - 2\bar{p}_x p_x E(X) \int_{-\infty}^{\alpha\bar{X}} (\alpha\bar{X} - X)\, f(X)\,dX \Bigg\}$$

where

$\mathrm{var}[\delta_{xI}]$ = variance of profit per acre of crop with insurance

The above expression indicates that the variance of profit depends on the variance of yield and the expected value of yield.

The covariance of profit between uses i and j is given by $E[\delta_i\delta_j] - [E\delta_i E\delta_j]$, which takes a different form depending on whether i and/or j are insured uses. If both represent insured crop uses, then

$$\text{(A9)}\quad \mathrm{Cov}_{ij} = \Bigg\{ \int_{\alpha\bar{X}_i}^{+\infty} \int_{\alpha\bar{X}_j}^{+\infty} [p_{xj}X_j - (C_{xj} + R_{xj})]\,[p_{xi}X_i - (C_{xi} + R_{xi})]\, h(X_jX_i)\, dX_j dX_i + \int_{-\infty}^{\alpha\bar{X}_i} \int_{-\infty}^{\alpha\bar{X}_j} [p_{xj}X_j + \bar{p}_{xj}(\alpha\bar{X}_j - X_j) - (C_{xj} + R_{xj})]\, [p_{xi}X_i + \bar{p}_{xi}(\alpha\bar{X}_i - X_i) - (C_{xi} + R_{xi})]\, h(X_jX_i)\, dX_j dX_i +$$

$$\int_{\alpha\bar{X}_j}^{-\infty} \int_{-\infty}^{\alpha\bar{X}_j} [p_{xj}X_j + \bar{p}_{xj} (\alpha\bar{X}_j - X_j) - (C_{xj} + R_{xj})] [p_{xi}X_i - (C_{xi} + R_{xi})]$$

$$h(X_jX_i)\, dX_jdX_i + \int_{-\infty}^{\alpha\bar{X}_i} \int_{\alpha\bar{X}_j}^{-\infty} [p_{xj}X_j - (C_{xj} - R_{xj})] [p_{xi}X_i + \bar{p}_{xi}$$

$$\left.(\alpha\bar{X}_i - X_i) - (C_{xi} + R_{xi})]\, h(X_jX_i)\, dX_jdX_i \right\} - E[\delta_i]E[\delta_j]$$

where

$h(X_jX_i)$ = the joint distribution density function between yields.

The covariance between profits depends on the covariance between yields. Other things being equal, the more profit associated with alternative crops tends to move together, the less do variations in profits cancel out; hence, the greater the variability of total profit.

The above rather elaborate expression can be contrasted with the covariance when both crops are uninsured. This is given by

(A10) $Cov_{i\,j} = \{p_{xi}p_{xj}E(X_iX_j) - C_{xj}p_{xi}E(X_i) - C_{xi}p_{xj}E(X_j) + C_{xi}C_{xj}\} - \{[p_{xj}E(X_j) - Cxj]\, [p_{xi}E(X_i) - C_{xi}]\}$

or

(A11) $Cov_{i\,j} = p_{xi}p_{xj}CovX_iX_j - 2p_{xj}C_{xi}E(X_j) - 2p_{xi}C_{xj}E(X_i)$

Finally, when one crop is insured (X_i) and the other (X_j) is not, their covariance is given by

(A12) $$Cov_{i\,j} = \int_{-\infty}^{+\infty} \int_{\alpha\bar{X}_j}^{+\infty} [p_{xi}X_i - (C_{xi} + R_{xi})]\, [p_{xj}X_j - C_{xj}]\, h(X_jX_i)\, dX_jdX_i$$

$$+ \int_{-\infty}^{+\infty} \int_{-\infty}^{\alpha\bar{X}_j} \{[p_{xi}X_i + \bar{p}_{xi}\, (\alpha\bar{X}_i - X_i) - (C_{xi} + R_{xi})]\, [p_{xj}X_j - C_{xj}]$$

$$h(X_jX_i)\, dX_jdX_i\} - E[\delta_i]E[\delta_j]$$