AAEA Selected Paper
AAEA Meetings, Long Beach, California, July 27-31, 2002

Asymmetric Information in Cotton Insurance Markets: Evidence from Texas

Shiva S. Makki
The Ohio State University and Economic Research Service, USDA

Agapi L. Somwaru
Economic Research Service, USDA

ABSTRACT

In recent years, the crop insurance program has emerged as an important part of the U.S. farm policy. Farmers responded to the crop insurance program with increased participation nationwide. At issue is whether the rapid expansion of the program has worsened the asymmetric information problems in crop insurance markets. This paper investigates the presence of adverse selection in cotton insurance markets. Our results reject the conditional independence of the choice of insurance contracts and risk of loss, implying the presence of informational asymmetries between the insurer and insured in Texas cotton insurance markets. Results show that actual premium rates are significantly different from both pure and fair premium rates.

Key Words: Adverse selection, Asymmetric information, Cotton, Crop insurance, Texas

———————————————————————

The authors gratefully acknowledge helpful comments received from Joy Harwood, Demcey Johnson, Robert Dismukes, Monte Vandeveer, and Jeffrey Hopkins. The views expressed herein are those of the authors and not necessarily those of the Economic Research Service or the U.S. Department of Agriculture.

Asymmetric Information in Cotton Insurance Markets: Evidence from Texas

Shiva S. Makki and Agapi L. Somwaru

The main purpose of this paper is to analyze the choice of insurance contracts and to investigate the presence of adverse selection in a crop insurance market when multiple yield and revenue insurance products are available to farmers. The federal crop insurance program is designed to offer producers protection against production and price risks. The Agricultural Risk Protection Act of 2000, which has committed $8.2 billion to the crop insurance program, increased premium subsidies for all coverage levels in an attempt to bring more farmers into the program and, perhaps, lessen the need for ad hoc emergency assistance and counter cyclical payments.

The crop insurance program expanded rapidly after the 1996 farm legislation. The insured acres for cotton increased from about a third of planted acres in the early 1990s to over 90 percent of planted acres in 2000. In addition, the number and types of insurance products offered also increased in recent years. Such a rapid expansion of the program could worsen the asymmetric information problems, thereby affecting the actuarial soundness of the crop insurance program. Past studies are few and most of these were conducted when only yield insurance products were available to farmers and participation in crop insurance markets was quite low. This paper investigates the presence of adverse selection in Texas cotton crop insurance markets when multiple yield and revenue insurance products are available to farmers. Texas represents over 40 percent of all planted, as well as insured, acres for cotton. We apply parametric and non-parametric procedures to test for the presence of adverse selection and compare actual, pure, and fair premium rates within yield span classes to offer further evidence of adverse selection in cotton insurance markets.

Asymmetric Information in Crop Insurance Markets

Asymmetric information has long been recognized as the major cause of inequities in rating and the failure of crop insurance markets (Ray; Goodwin and Smith; Just, Calvin, and Quiggin; Makki and Somwaru). Asymmetric information manifests itself primarily as adverse selection or moral hazard. Adverse selection exists in insurance markets because of differences in inherent farm risks, arising from factors such as the farm's location or farmer's managerial abilities. In the presence of adverse selection, the insurance provider fails to accurately assess the risk of loss and, therefore, is unable to set premiums commensurate with risk. Moral hazard, on the other hand, is caused by the hidden actions of the insured which increase the risk of loss and, therefore, the likelihood of collecting indemnities.

Very few studies consider both informational problems. The literature on adverse selection and moral hazard problems was developed separately and faced different theoretical issues. In the adverse selection literature, for example, the emphasis was put on the existence and efficiency of competitive equilibrium. Theoretical and empirical studies in automobile and health insurance markets have shown that adverse selection reduces the consumption of insurance by low-risk individuals, and results in the transfer of income from low-risk to high-risk insureds and eventual market failure. The seminal works of Akerlof and Rothschild and Stiglitz derived market equilibria conditions under asymmetric information. See Makki and Somwaru for a more detailed review of theoretical and empirical studies on adverse selection in automobile and health insurance markets.

The moral hazard literature, on the other hand, emphasizes contractual arrangements and the behavioral response of the insured after the purchase of insurance. Separating adverse selection and moral hazard is difficult and requires panel data on contractual arrangements, indemnity payments, and production and management practices for individual farmers. Such data are not available at the moment, particularly for cotton. The analysis of moral hazard is left for future research. In this paper, we focus on asymmetric information caused by adverse selection problems.

Adverse Selection

Several studies have documented the presence of adverse selection and its implications for U.S. crop insurance markets. See Makki and Somwaru or Knight and Coble for a more detailed review of empirical studies on adverse selection in crop insurance markets. Past studies on adverse selection in crop insurance markets have three major drawbacks. First, most of the studies adopt parametric models that rely on restrictive functional forms and a fairly large number of exogenous variables. Second, they narrowly focus on adverse selection when only a yield insurance product was offered to farmers and participation was quite low. Third, no empirical studies explicitly test for the presence of adverse selection in the crop insurance markets. Makki and Somwaru apply both parametric and non-parametric procedures to Iowa corn and soybean insurance markets where farmers were offered a portfolio of yield and revenue insurance contracts. Their study shows, in the case of individual yield and revenue insurance products, that actual premium rates paid by high risk farmers are lower than competitive or actuarially fair rates, while in the case of area (county) yield insurance, the actual premium rates are comparable to competitive rates across all risk types. Their study, however, does not account for risk classification of crop insurance participants. Failure to account for risk classification may overstate the asymmetric information problems (Dionne, Gourieroux, and Vanasse).

Assessment of Risk

The most critical issue for any effective insurance program is the ability to assess risk appropriately and to set premium rates commensurate with that risk. If premium rates are set to reflect average risks, the insurer will be left with an adversely selected pool composed of more high risk individuals for whom insurance is under-priced and fewer low risk individuals for whom insurance is overpriced (Rothschild and Stiglitz). Under current crop insurance programs, premium rates are primarily driven by producers' average yield, where producers with higher average yields are assessed lower premium rates on the assumption that expected losses decrease as expected yield increases. Failure to account for the distribution of yield can adversely affect the actuarial soundness of the crop insurance program. As Skees and Reed argued, it is possible that two farmers could have different expected yields and similar yield distributions, or different yield distributions and similar expected yields. Both situations can lead to inappropriate insurance rates and result in adverse selection. Inaccurate premium rates can seriously affect the viability of the crop insurance program.

In this paper, we analyze two different measures of risk. First, we use loss frequency as an indicator of risk. Loss frequency is an *ex post* observation of whether a farmer filed a claim. A farmer who filed a claim is considered to be of higher risk than the one who did not file a claim. This is the most common measure of risk used in the analysis of automobile and health insurance markets (Puelz and Snow).

Second, we develop a measure of risk that accounts for both the mean and variance of yield or revenue. For yield insurance products, the risk of loss is captured by the probability of *yield* falling below the 65% guarantee level (Y^p), while for revenue insurance products the risk of loss is captured by the probability of *revenue* falling below the 65% guarantee level (R^p). Assuming a normal distribution of yield, Y^p is estimated for *each insured unit* based on the ten years of yield history from the same unit.[1] R^p is estimated for *each insured unit* based on the ten years of yield history and the corresponding marketing year average price. Farm revenues are also assumed to be normally distributed. Hereafter, we refer to this measure of risk as the Estimated Probability of Loss (EPL).

The EPL represents asymmetric information between the insurer and the insured, in the sense that, at the time of contract choice, the insured knows more about his/her risk of loss than does the insurer. This measure of risk, constructed using past yield or revenue records provides a fairly robust measure of the individual's risk (Makki and Somwaru).

Risk Classification

Insurers often use observable characteristics to categorize individuals into different risk groups. Risk classification is useful if individuals differ by observable traits that are correlated with the risk of loss. For example, insurers in automobile insurance classify individuals based on age, gender, occupation, etc., which are correlated with risk of loss. The purpose of risk classification is to reduce the asymmetrical information between the insurer and the insured. Risk classification improves the efficiency of insurance markets and reduces cross-subsidization between risk types (Crocker and Snow). Chiappori and Salanie and Dionne, Gourieroux, and Vanasse show that risk classification reduces asymmetric information problems in automobile insurance markets. Dionne, Gourieroux, and Vanasse argue that any test for asymmetric information should be applied within each risk class. They further argue that econometric analysis should include risk categorization variables in the model.

For crop insurance, the Risk Management Agency (RMA) uses the "yield span" concept to categorize farms into different *yield classes* and charges different premium rates by yield class. The yield span classification groups farmers' yields into thirteen discrete categories (R01 to R13) based on the ratio of a farmer's yield to the average county yield. According to the yield span concept, category R01 includes farmers with the lowest average yields and is considered to represent the highest risk farmers, while category R13 includes farmers with the highest average yields and is considered to represent the least risky farmers. Yield span category R07 includes all farms whose yields are expected to be equal to the county's average yield. Yield span ranges are derived from historical county loss experience and are based on county-level average yields reported by the National Agricultural Statistical Service (NASS).

Given the importance of individual farm production history in determining yield risk, we classify farmers into two yield experience classes: those who have reported their actual production history (APH) for the last 6 years and those who did not report their yields. The latter class mostly includes units farmed by new farmers, units with newly added acres, and marginal lands. For those who do not report past yield records, RMA uses transitional yield to estimate premium rates. Transitional yield approximately represents county average yields. Since average crop yield is one of the major determinants of premium rates, we expect our analysis to offer insights into the importance of individual farm yield records for assessing the risk of loss and establishing yield span classes.

Testing for Adverse Selection

Using theoretical models, Rothschild and Stiglitz, Puelz and Snow, and others argue that insurance contracts with higher coverage levels are chosen by agents with higher probability of loss. This hypothesis suggests a positive correlation between coverage level and risk of loss. Our goal in this section is to analyze the factors that influence the choice of insurance coverage and test the correlation between the choice of coverage level and the occurrence of a loss.

An empirical test of adverse selection is to verify whether higher coverage contracts are chosen by individuals with higher probabilities of loss. In this study, we test for the presence of adverse selection using parametric and non-parametric procedures. In both procedures, we test the conditional independence of the choice of insurance coverage and the risk of loss. Rejection of independence means that there is evidence of informational asymmetries in such insurance markets (Chiappori and Salanie; Dionne, Gourieroux, and Vanasse). In other words, if the level of insurance coverage is correlated with the level of risk, then agents indeed have better knowledge of their risk.

We also test whether premium rates reflect the levels of risk associated with different contracts by comparing actual premium rates to pure as well as fair premium rates. We define and discuss the computation of these premium rates later. This particular test (comparing actual, pure, and fair premium rates) offers further evidence regarding asymmetric information caused by adverse selection problems in cotton insurance markets.

Conditional Independence

Let Y be the endogenous variable under study, θ be the decision variable, and X a matrix of exogenous variables. The decision variable provides no additional information on Y if and only if the prediction of $f(Y)$ based on X and θ jointly coincides with its prediction based on X alone. This inference can be written in terms of conditional probabilities as:

(1) $$f(Y|X, \theta) = f(Y|X),$$

Where $f(./.,.)$ denotes a conditional probability density function (see Gourieroux and Monfort, pp. 470-75). In our application, Y is the loss frequency, θ is the choice of coverage level, and X is a set of exogenous variables, including those that affect the decision variable and that define risk classes. Equation (1) means that the choice of coverage level, θ, provides no useful information in predicting the loss frequency, Y. In other words, if equation (1) holds, then we can conclude that θ and Y are independent, conditional on X. Equation (1) can be written in equivalent form as:

(2) $$f(\theta|X,Y) = f(\theta|X),$$

where the loss frequency does not provide useful information in predicting the conditional coverage choice (Dionne, Gourieroux, and Vanasse). This characterization may also be interpreted as the description of what the insured would decide if he had more knowledge on future risk than the insurer. If indeed the insured individual has more information on the risk of loss than the insurer, then the equality will not hold in equation 2 (Dionne, Gourieroux, and Vanasse). In this application, we shall use such a characterization.

Econometric Model

Conditional dependence of coverage level and the risk of loss is analyzed in an econometric framework, where "conditional" means conditional on all variables observed by the insurer. That is, several other factors, including risk aversion, type of insurance product, and cost of insurance, also influence the choice of an insurance coverage level, and the models that test independence of the coverage level and the risk of loss should account those factors. An econometric model can be used to control for factors that might affect the choice of the coverage level. The econometric model is specified in general form as:

(3) $$f(\theta|X,Y) = f(\theta; Y, E(Y|X), X),$$

where E(Y|X) is an the expected value of Y computed from initial information. By introducing the estimated expectation of loss frequency (E(Y|X)) in equation (3), we account for any omitted non-linear effects. That is, equation (3) avoids the difficulty of distinguishing the informational content of a decision variable and any omitted non-linear effect of the initial exogenous variables (see Dionne, Gourieroux, and Vanasse for more details). If the estimated coefficient of Y is no more significant when E(Y|X) is introduced into the model, this means that there is no informational asymmetries in the insurance market.

We use the relationship in equation (3) to specify a censored TOBIT model for analyzing the choice of coverage levels as follows:

$$(4) \qquad \theta = \begin{cases} \theta_L & \text{if } \theta^* \leq \theta_L \\ \theta^* = \alpha y + \lambda E(y/x) + x\beta + e_i & \text{if } \theta_L < \theta^* < \theta_U \\ \theta_U & \text{if } \theta^* \geq \theta_U \end{cases}$$

where θ is the chosen coverage level, which ranges from a minimum of 50 percent (θ_L) to a maximum of 75 percent (θ_U) offered at 5 percent increments. α, λ, and β are regression coefficients (see Long, pp. 212 for more on censored TOBIT models). The set of relevant explanatory variables in equation (4) include loss frequency, estimated expectation of loss frequency, estimated probability of loss (EPL), yield span categories, availability of farm yield records, type of insurance contract, farm practice, unit structure, farm size, elected price, producer cost of insurance, and premium subsidies.

The loss frequency and EPL represent the underlying risks associated with the insurance contracts. Statistically significant coefficients for loss frequency and EPL imply that low-risk and high-risk farmers purchase different coverage levels, while non-significant coefficients imply that the chosen coverage level is independent of risk of loss. In some ways, these two measures of risk embody asymmetric information between the insured and the insurer. That is, at the time of insurance purchase, farmers know more about the distribution of yield on his/her farm than does the insurer.

Yield span and availability of farm yield records account for the risk classification used by RMA. For econometric analysis, the 13 yield span classes are represented by discrete numbers, with one representing lowest yield class and 13 representing highest yield class. Yield span classification and farm yield records are key factors in computing premium rates (GAO;

Skees and Reed) and are likely to be important determinants of insurance coverage choice. At issue is whether the yield span classification helps in reducing asymmetric information problems in crop insurance markets.

Availability of farm yield records is critical for assessing the risk of loss. At issue is whether those farmers who did not have yield records represent higher risk than those who reported yields. In this model, we use a dummy variable to represent the availability or non-availability of farm yield records. In particular, the dummy variable takes the value of one for those that have reported yields for each of the last 6 years and zero for those who did not report yields in any of the years or only a few of the years.

Type of insurance contract refers to whether farmers purchased yield or revenue insurance. Controlling for the type of insurance is important because the levels of risk protection provided by yield and revenue insurance contracts may be different even if they offer same level of coverage. Type of insurance is likely to influence farmers' decisions regarding coverage levels. In our econometric model, the type of insurance is represented by a dummy variable, with zero for choice of yield insurance product and one for choice of revenue insurance product. We exclude CAT from the analysis because the level of coverage is fixed at 50 percent of the expected yield and it is a completely subsidized product (except for a small fixed fee). Since the dependent variable in our econometric model is the coverage level, CAT contracts are unlikely to provide rich information about farmers' behavior.

Farm practice -- i.e., whether or not a farm is irrigated -- is included because irrigation has the potential to reduce yield risks and may provide farmers with incentive to select lower coverage levels. For the econometric analysis, the farm practice variable is set equal to one for irrigated farms and zero for non-irrigated farms.

Farm size is used to represent a producer's risk aversion behavior because farm size is often positively correlated with wealth. We expect the choice of insurance coverage level to be related to farm size in a manner consistent with the decreasing risk aversion hypothesis. The logic that producers with large farm sizes are wealthier and, most likely, less risk averse. One expects that large-farm operators would be more likely to choose lower coverage contracts as they are able to diversify their risks and manage variations in income within their operations better than would farmers with smaller operations. Farm size is computed by aggregating all the insured units by the same farm operator.

Unit structure refers to the way insurable units are defined. There are two common types of insured units: basic units and optional units. Basic units consist of all acreage in a county held by the insured under identical ownership. Optional units are subdivided basic units. The criteria for optional subdivision are based on location and production practices. Allowing insured farmers to have multiple, separately insured units has been criticized as a source of increased losses due to a reduction in the portfolio effect of spatial diversity and an increased potential for fraudulent reporting of losses through shifting of reported production from a unit on which a loss is being claimed to one on which there is no loss (Knight). Farmers who insure a basic unit get discounts of up to 10% on their premium rates. For the econometric analysis, unit structure is set equal to 0 for those who insured a basic unit and 1 for those who insured optional units.

Elected price, which is a certain fixed proportion of the expected price, is used to compute indemnities. In case of revenue insurance, elected price also serves as a guaranteed price. The procedure for estimating expected prices varies by insurance plan. For example, APH contracts use USDA's projected farm-level price for the crop year, while CRC uses the planting-time quote on the harvest-time futures price. This variable indicates the importance of price guarantee in the choice of coverage levels.

The cost of insurance, captured by the premium per dollar of liability, is calculated as producer paid premium (excluding subsidy) divided by total liability. Liability represents the maximum potential indemnity or value of the insurance contract if a producer loses the entire crop. Subsidy indicates the proportion of total premium subsidized by the government. Producer cost of insurance and government subsidy are important variables that influence the choice of coverage level. In this study, we refrain from analyzing the effects of premium subsidy on adverse selection in crop insurance markets (see Goodwin and Vandeveer for analysis of the effects of federal subsidies on crop insurance markets).

Non-Parametric Tests

We apply non-parametric tests to complement the econometric tests of correlation between the choice of coverage level and the risk of loss. The advantage of non-parametric methods over econometric methods is that non-parametric methods do not require specification of a functional form. In addition, non-parametric methods account for more complicated non-linear relationships between the level of coverage and the risk of loss (Chiappori and Salanie). We apply these non-parametric tests within each yield span category to recognize RMA's risk classification and to avoid the potential bias caused by differences in premium rates between categories.

Two non-parametric tests, the Kruskal-Wallis Π^2 and the Kolmogorov-Smirnov, are performed to investigate whether the choice of insurance coverage level and the risk of loss are independent. The Kruskal-Wallis test statistic is given by:

(5) $$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{T_i^2}{n_i} - 3(N+1)$$

where N is the sample size, T_i is the ranks assigned to the ith group and n_i is the number of observation in the i^{th} group. The test statistic *H* approximately follows a chi-squared distribution with k-1 degrees of freedom, k is number of groups or populations. See Conover (p. 229) for more details on the Kruskal-Wallis test. The Kolmogorov-Smirnov test statistic is given by:

$$K = \sqrt{M} \sup_x \left| F_M(x) - F(x) \right| \tag{6}$$

where $F_M(x)$ is the empirical cumulative density function (cdf) and F(x) is the cdf of a $\Pi^2(1)$. Under conditional independence, the test statistic K converges to a distribution that is tabulated in Massey and also presented in standard statistics textbooks (Conover, p.462). When we cannot reject conditional independence, we conclude that there is no asymmetric information. Rejection of independence implies asymmetric information between the insured and the insurer, but this may occur due to either adverse selection or moral hazard (Chiappori and Salanie).

Comparing Actual, Pure, and Fair Premiums

In this section, we compare the actual premium rate to pure and fair premium rates to provide further evidence of asymmetric information in cotton insurance markets. The actual premium is the total premium reported in the insurance records. Total premium includes the portion of the premium paid by the government. The actual premium rate is obtained by dividing the total premium by liability.

The pure premium rate is the ratio of indemnities paid to policy holders divided by the insurance provided to those policy holders (liability). This measure is called the pure premium because it represents the premium rate that should have been charged to exactly offset the losses. The pure premium rate is also referred to as the loss-cost ratio in the crop insurance industry.

The fair premium rate, on the other hand, is the expected indemnity per dollar of liability. In an efficient insurance market, the fair premium is equal to the expected indemnity (Rothschild and Stiglitz; Puelz and Snow). We calculate the expected indemnity for each insurance contract separately. For example, expected indemnity, E(I), for a typical APH contract is estimated as follows:

(7) $$E(I) = \frac{1}{T}\sum_{t=1}^{T} MAX[0,(\theta Y^{e} - Y^{t})P^{g}]$$

where 2 is the chosen coverage level, y^e is the expected yield, y^t is the realized yield in year t, P^g is the guaranteed price (or elected price), and *T* is the number of periods for which yield records were available. We use 10 years (1990-99) of reported yields for each insured unit. The expected indemnity payment for a typical CRC contract are estimated as follows:

(8) $$E(I) = \frac{1}{T}\sum_{t=1}^{T} MAX[0,(\theta Y^{e} \max(P^{g}, P^{t}) - Y^{t} P^{t})]$$

where 2, y^e, y^t, and P^g are as defined earlier. P^t is the market price in year t.

If there are no informational asymmetries in the market, the difference between actual and pure (as well as fair) premium rates should not be significantly different from zero. Under asymmetric information, however, one would expect differences to exist between actual and pure (or fair) premium rates. These differences result because an accurate determination of individual farmers' risk is either not possible or prohibitively expensive. Since adverse selection reflects the inability of the insurer to rate the risk appropriately, significant differences between actual and pure premium rates, and actual and fair premium rates indicate the presence of adverse selection in crop insurance markets.

We use non-parametric tests to demonstrate the differences, if any, between the actual and pure premium rates and between actual and fair premium rates. The two non-parametric tests performed are the Kruskal-Wallis Π^2 test and the Kolmogorov-Smirnov test as described above.

Evidence of Adverse Selection

Data used in this study are from USDA's RMA, which maintains records of all individual farmers who purchase federally-backed crop-yield or revenue insurance. The data pertain to cotton farmers in Texas for the 1999 crop year. We selected a sample of about 50,000 unit-level insurance contracts. For each contract, we have several variables that describe the characteristics of the insurance, such as choice of insurance plans and coverage levels, premium and subsidy rates, elected price, and indemnity payments. We also have variables that help assess the risk associated with the contract – past yield records, yield span, farm practice, and liability. Other variables, such as unit option and reporting agency, are related to contract design and implementation.

In this section, we present the results of parametric and non-parametric procedures to test for the conditional independence of risk of loss and the choice of insurance contract. We also present the results of non-parametric procedures to test the differences among actual, pure, and fair premium rates within each yield span category.

Econometric Results

We analyze the choice of coverage level using a censored TOBIT model described in equation 4. We estimate two models: model 1 excludes the expected value of loss frequency, while model 2 includes the expected value of loss frequency.

The expected values of loss frequency, E(Y/X), were computed using a two-step procedure. *First*, we estimate loss frequency as a function of all the available information (X), using a negative binomial regression model: $Y = f(X)$. *Second*, using the estimated negative binomial regression function, we compute the expected value of loss frequency or E(Y/X). The expected value of loss frequency is then incorporated into model 2 to account for the nonlinear effects of variables on loss frequency. Our main goal is to test for correlation between the choice of coverage level and the occurrence of loss. The occurrence of loss is measured by the loss frequency and the estimated probability of loss (EPL). Table 1 presents the estimated results for Texas.

Econometric results suggest that the loss frequency and the choice of coverage level are correlated or that the null hypothesis of conditional independence is rejected, implying the presence of informational asymmetries in the cotton insurance market. The coefficient for loss frequency is positive and significant in both models, indicating that those who experience high losses choose higher coverage levels (Table 1).

Our results also reveal a strong relationship between EPL and choice of coverage level in both models, even after controlling for various factors including risk classification, type of insurance product, farming practices, farm size, and cost of insurance (Table 1). The positive and significant coefficients for the estimated probability of loss indicate that those farms that have a higher probability of yield or revenue falling below the guarantee level are more likely to choose higher coverage contracts. This implies that farmers choose coverage levels depending on their expected risk of loss, with high risk farmers purchasing higher coverage levels and vice versa. This positive correlation between the choice of coverage level and the EPL implies the presence of asymmetric information in the Texas cotton insurance markets.

Results show that the yield span -- the ratio of expected yield to the county-level average -- is a significant variable in the choice of coverage level (see Table 1). The positive relationship between coverage level and yield span implies that those farms with expected yield greater than the county average are more likely to buy higher coverage levels. This is probably because the premium rates and yield span categories are inversely proportional, implying lower premium rates for farms with higher expected yields. The standard RMA assertion is that, on average, the probability of a loss is greater for producers whose yields are below the county average and vice versa. Our finding, however, is not consistent with this assertion. Our analysis indicates that high risk farmers are more likely to buy higher coverage contracts. We, therefore, argue that average yields or yield spans do not appropriately represent the actual risk of loss.

Our results indicate that past yield records are an important determinant of the choice of insurance coverage level. Our results show that those farms that reported at least 6 years of yield records are likely to purchase lower coverage levels compared to those who did not have past yield histories (Table 1). This is an interesting finding given the argument that those who lack APH yields pose greater insurance risk. Table 1 also indicate that unit structure significantly affects the choice of insurance coverage levels. Results show that farmers who opted for basic units are likely to buy higher coverage levels. This is probably because of the 10 percent premium discount farmers receive for choosing the basic unit option.

The estimated relationship between farm size and choice of coverage level is negative and significant (Table 1). The negative coefficient implies that small farm operators are more likely to purchase higher coverage levels and vice versa. Large farm operators may use other risk management strategies such as diversification and options contracts, which are often substitutes for crop insurance.

A negative coefficient for farm practice indicates that farmers who irrigate their land prefer lower coverage compared with non-irrigated farms (Table 1). This is consistent with our hypothesis that irrigated farms may face lower yield risks compared to non-irrigated farms.

Cost of insurance is also a critical factor that influences farmers' choice of insurance coverage levels. Producers who buy crop insurance do not consider the total premium rate as an indicator of cost. Instead, they commonly consider the subsidized premium rate as the true indicator of the cost of insurance. This is the reason for using producer premium and subsidy rates separately in the econometric analysis. Our analysis indicates that both premium per dollar of liability and subsidy rates are important in choosing the level of insurance coverage (Table 1). Premium rates are positively correlated with coverage level because farmers are offered a menu of coverage-premium combinations, where higher coverage levels are associated with higher premium rates. The subsidy associated with different coverage levels can also influence the choice of coverage level, since the federal subsidy reduces the actual cost of crop insurance. The subsidy variable has a negative sign because, by design, the rate of subsidy declines as the level of coverage increases.

Non-Parametric Test Results

The Kruskal-Wallis and the Kolmogorov-Smirnov tests are used to test the hypothesis of independence between the choice of insurance coverage and the risk of loss within each yield span class. Since RMA utilizes yield spans to group farmers into different yield classes and to price insurance contracts, failure to account for such a risk classification may generate a bias that tends to overestimate the level of asymmetric information between the insured and the insurer. Therefore, we apply these tests to each yield span class separately. Table 2 presents the results of non parametric tests.

In the case of APH contracts, the hypothesis of independence between the choice of insurance coverage level and the loss frequency is rejected in 8 of 13 yield span classes using the Kruskal-Wallis test and in all yield span classes using the Kolmogorov-Smirnov test (Table 2). Both tests also reject the independence of the choice of coverage and the EPL within every single yield span class.

In the case of CRC, however, the results are less clear. For example, the Kruskal-Wallis test fails to reject independence of the choice of coverage level and loss frequency in 8 of 13 yield span classes (Table 2). The Kolmogorov-Smirnov test, however, rejects independence between the choice of coverage level and loss frequency in all yield span classes. The Kruskal-Wallis test rejects independence of the choice of coverage level and EPL in yield span classes R04 and higher.

These results imply that the choice of insurance contract and the risk of loss are correlated, or that the null hypothesis of conditional independence is rejected, even after controlling for various risk classification variables used in crop insurance. This finding implies the presence of informational asymmetries between the insurer and the insured in the crop insurance markets analyzed in this study.

Comparing actual, pure, and fair premium rates

In an efficient (full information) insurance market the actual, pure, and fair premium rates must be approximately equal. We test for the difference between actual and fair premium rates as well as the difference between the actual and pure premium rates across different risk classes using non-parametric tests.

Table 3 presents the results of comparing actual premium rates to fair and pure premium rates using non-parametric tests. The tests are applied to APH and CRC contracts separately at the 65% coverage level within each yield span category. The 65% coverage level is chosen because this is the level of coverage most commonly purchased by farmers. The two non-parametric tests, Kruskal-Wallis and Kolmogorov-Smirnov, reject the hypothesis that actual and fair rates are not different for both APH and CRC within most yield span classes (Table 3). That is, the computed values of the test statistics for APH and CRC exceed that of the critical values at the 5 percent level of significance. Results also reject the hypothesis that the actual and pure premium rates are not different for both APH and CRC contracts. Both the Kruskal-Wallis and Kolmogorov-Smirnov tests yield consistent results. These results suggest that the insurance rates for cotton in Texas do not reflect the risk of loss appropriately in most yield span categories.

Two factors suggest that the observed informational asymmetries are more likely due to adverse selection than moral hazard. First, crop insurance contracts do not insure the full loss. Since crop insurance protects only up to 75 percent of expected yield or revenue, it seems unlikely that farmers would indulge in activities that would decrease yield or revenue by more than 25 percent. This is similar to the arguments made for deductibles to reduce moral hazard problems in automobile and health insurance markets. Second, the risk of loss in agriculture is more due to weather and other natural factors. Insured individuals are less likely to affect the distribution of losses caused by such exogenous factors; hence the observed differences between the actual and the pure premium rates are most likely due to adverse selection.

Conclusions

In this study, we assess the risk of loss and test for the presence of asymmetric information in U.S. crop insurance markets for cotton. The data pertain to cotton producers for the crop year 1999 in Texas. The data gathered for the study provide the first opportunity to test for asymmetric information in cotton insurance markets.

We explicitly test for asymmetric information using parametric and non-parametric procedures. Both procedures reject conditional independence between the choice of coverage levels and the risk of loss, implying the presence of informational asymmetries between the insurer and insured in Texas cotton insurance markets. Our findings provide useful insights into asymmetric information problems and methods adopted to assess the risk of loss in crop insurance markets. Our results are consistent with Makki and Somwaru who show the presence of asymmetric information in the Iowa corn crop insurance markets.

Another contribution of this paper, as compared to the earlier literature, is that we test for asymmetric information within yield span classes. Such testing within a homogenous group reduces the likelihood of mistakenly concluding that asymmetric information exists. Since premium rates vary across yield span classes, ignoring these classifications may produce misleading results. Results show that the choice of coverage level and the risk of loss are correlated within each yield span class, implying the presence of asymmetric information in the insurance markets analyzed in this study.

Comparing the actual, pure, and fair premium rates across different yield spans offers further evidence of the presence of asymmetric information. In a full information market, all three premium rates must be approximately equal. Non-parametric tests show that actual premium rates are different from both pure and fair premium rates, implying informational asymmetries in the crop insurance market for cotton. We find that actual premium rates are higher than pure or fair rates for cotton.

Whether the problem of asymmetric information is intrinsic to crop insurance or whether it is a consequence of program design is beyond the scope of this paper. However, we do recognize that government regulations might have contributed to the presence of asymmetric information and, more particularly adverse selection, in crop insurance markets. For example, as a public program, the government does not allow the insurer to condition the contract on observable traits, e.g., demographic factors, that would improve risk assessment of individual farms. Also, under current reinsurance arrangements, insurance companies transfer most of the risk to the government, and therefore have little incentive to monitor problems generated by asymmetric information.

The analysis in this study is limited to cotton in Texas. An extension of the study to include other crops and States would be useful. Furthermore, the data used in this study represent only a single year, 1999, which may not be a representative. An extension of this analysis to include more years would provide a more robust set of results. In addition, a panel data may facilitate tests to discriminate between adverse selection and moral hazard.

Crop insurance programs are subjected to government regulations that prevent insurers from conditioning the contract on observable factors such as farm size, education, or gender of the farm operator. These factors may be strongly correlated with adverse selection and/or moral hazard problems. However, adverse selection or moral hazard problems that arise because an observable risk indicator is ignored do not constitute market failure. A study to understand the role of government regulations and market failures would be useful for both policy makers and private insurers.

End Note:

1. We recognize that several studies, including Buccola; Moss and Shonkwiler; Nelson and Preckel; and Taylor, reject the normality assumption. However, there seems to be no consensus among these studies regarding skewness of the distributions. More recently, Just and Weninger fail to reject normality tests for yield distribution of Kansas farm-level wheat, corn, and sorghum yield data. In another study, Just, Calvin, and Quiggin assume a normal distribution for corn yield histories.

Table 1 - Censored TOBIT on coverage level, Texas, Cotton

Variable	Model 1: Conditional on Loss Frequency		Model 2: Conditional on Loss Frequency and Expected Loss Frequency	
	Coefficient	**Z-Ratio**	**Coefficient**	**Z-Ratio**
Intercept	-0.3809	-0.18	-5.7676	-2.75
Loss Frequency (Y)[1]	0.0910	2.60	0.1303	3.80
E(Y/X)[2]			8.8403	19.00
EPL[3]	0.3021	42.95	0.6010	35.09
Yield Span	0.0901	11.63	0.0802	10.56
Yield Experience	-0.5003	-13.31	-1.5106	-23.44
Unit Structure	-0.3427	-10.51	-0.3451	-7.17
Farm Size	-0.0097	-4.72	-0.0206	-9.86
Farm Practice	-0.1406	-2.44	-0.2303	-4.06
Insurance Plan	-0.0284	-0.25	-1.0562	-8.59
Elected Price	89.87	27.68	87.10	27.38
Producer Premium	2.7040	313.63	2.7172	321.27
Premium Subsidy	-1.6667	-264.20	-1.6749	-270.82
Adjusted R-Square		0.7196		0.7210

1. Y is the loss frequency;
2. E(Y/X) is the estimated expected loss frequency;
3. EPL is estimated probability of loss.

Table 2 - Testing the conditional independence of the level of insurance coverage and the risk of loss within the risk category: Non-parametric test results, Texas, Cotton

Insurance Contract	**Yield Span Class[1]**	**Conditional Independence between**			
		Coverage level and Loss Frequency		**Coverage level and EPL[2]**	
		Kruskal-Wallis Test	Kolmogorov-Smirnov Test	Kruskal-Wallis Test	Kolmogorov-Smirnov Test
APH	R01 (1293)	27.35	2.73	383.70	9.75
	R02 (1098)	4.07	1.05	227.90	7.22
	R03 (1645)	0.51	0.77	354.84	9.11
	R04 (2567)	0.30	0.56	328.30	8.28
	R05 (3393)	5.23	1.24	477.72	9.30
	R06 (4024)	5.87	0.85	630.64	10.97
	R07 (4105)	4.40	1.50	621.09	11.60
	R08 (4093)	16.89	2.55	624.20	11.00
	R09 (3787)	2.32	1.64	637.62	11.43
	R10 (3406)	0.58	0.70	505.59	10.07
	R11 (3303)	2.62	0.81	414.54	9.32
	R12 (2806)	13.13	1.90	446.06	9.60
	R13 (10604)	59.96	3.40	1668.09	17.86
Critical Value @0.05 Significance		**3.84**	**1.36/√N**	**18.31**	**1.36/√N**
CRC	R01 (52)	6.35	0.56	4.92	0.72
	R02 (73)	0.32	0.17	11.58	0.98
	R03 (147)	0.03	0.21	15.85	1.28
	R04 (218)	0.76	0.27	20.37	1.40
	R05 (358)	0.03	0.09	68.92	2.16
	R06 (397)	0.50	0.25	74.83	2.64
	R07 (359)	6.35	0.77	75.24	2.90
	R08 (313)	5.64	1.01	79.50	3.28
	R09 (329)	3.15	0.44	65.75	2.51
	R10 (249)	4.43	0.59	71.24	2.72
	R11 (167)	0.003	0.50	34.80	2.09
	R12 (134)	0.73	0.48	33.03	1.86
	R13 (403)	6.61	1.04	51.48	1.70
Critical Value @0.05 Significance		**3.84**	**1.36/√N**	**18.31**	**1.36/√N**

1. Values in the parentheses indicate the number of contracts (N).
2. EPL is the estimated probability of loss defined as the probability of yield or revenue falling below the guarantee level.

Table 3 - Testing for the difference between different premium rates: Non-parametric test results, Texas, Cotton.

Insurance Contract	Risk Class	Conditional Independence between			
		Actual and Fair Premium Rates		Actual and Pure Rates[2]	
		Kruskal-Wallis Test	Kolmogorov-Smirnov Test	Kruskal-Wallis Test	Kolmogorov-Smirnov Test
APH/65	**Yield Span[1]**				
	R01 (1293)	6.69	1.89	87.58	4.58
	R02 (1098)	25.02	2.96	134.14	5.38
	R03 (1645)	15.44	3.02	274.14	7.98
	R04 (2567)	47.65	4.52	381.38	9.22
	R05 (3393)	51.42	4.02	504.52	10.38
	R06 (4024)	122.33	5.77	479.11	9.99
	R07 (4105)	116.36	5.88	430.85	9.78
	R08 (4093)	186.29	6.52	413.92	9.84
	R09 (3787)	210.33	7.05	356.01	9.17
	R10 (3406)	195.07	6.54	210.06	6.99
	R11 (3303)	177.70	6.85	226.82	7.20
	R12 (2806)	127.27	6.40	227.40	7.10
	R13 (10604)	756.93	13.42	655.78	11.83
Critical Value @0.05 Significance		**16.92**	**1.36/√N**	**16.92**	**1.36/√N**
CRC/65					
	R01 (52)	16.14	2.31	6.51	1.22
	R02 (73)	0.55	0.71	1.02	0.87
	R03 (147)	17.41	2.94	8.75	1.03
	R04 (218)	30.45	2.81	56.92	2.43
	R05 (358)	8.00	2.09	123.82	3.51
	R06 (397)	25.30	2.85	117.96	4.18
	R07 (359)	37.93	3.01	84.97	3.70
	R08 (313)	14.55	2.32	59.00	3.38
	R09 (329)	35.60	2.87	51.25	2.92
	R10 (249)	17.08	2.85	56.63	3.06
	R11 (167)	16.62	2.76	60.27	3.36
	R12 (134)	11.98	2.37	36.63	2.77
	R13 (403)	21.52	3.10	67.24	4.13
Critical Value @0.05 Significance		**14.07**	**1.36/√N**	**14.07**	**1.36/√N**

1. Values in parentheses indicate the number of contracts (N).
2. Actual rate is the premium per dollar of liability.
3. Actual producer rate is the producer premium per dollar of liability.
4. Pure premium represents indemnity paid out per dollar of liability, also known as the loss-cost ratio.
5. Fair premium refers to expected indemnity per dollar of liability.

References

Akerlof, G. A. "The Market for Lemons: Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84(1970):488-500.

Buccola, S. T. "Testing for Nonnormality in Farm Net Returns." *American Journal of Agricultural Economics,* 68(1986):334-43.

Chiappori, P. and B. Salanie. "Testing for Asymmetric Information in Insurance Markets." *Journal of Political Economy*, 108(2000):56-78.

Coble, K. H., T. O. Knight, R. D. Pope, and J. R. Williams. "An Expected Indemnity Approach to the Measurement of Moral Hazard in Crop Insurance." *American Journal of Agricultural Economics* 79(February 1997):216-226.

Conover, W. J. *Practical Nonparametric Statistics*, 2nd Edition, New York: John Wiley and Sons, 1980.

Crocker, K. J. and A. Snow. "The Efficiency Effects of Categorical Discrimination in the Insurance Industry." *Journal of Political Economy*, 94(1986):321-344.

Dionne, G., C. Gourieroux, and C. Vanasse. "Testing for Evidence of Adverse Selection in Automobile Insurance Markets: A comment." *Journal of Political Economy*, 109(2001):444-453.

General Accounting Office (GAO). *Crop Insurance: Further Actions could Strengthen Program's Financial Soundness*, United States General Accounting Office, Washington, DC., 1999

Goodwin, B. K. "An Empirical Analysis of the Demand for Multiple Peril Crop Insurance." *American Journal of Agricultural Economics* 75(May 1993):425-434.

Goodwin, B. K. and V. H. Smith. *The Economics of Crop Insurance and Disaster Aid*, American Enterprise Institute Press: Washington, D.C., 1995

Goodwin, B. K. and M. Vandeveer. "An Empirical Analysis of Acreage Distortions and Participation in the Federal Crop Insurance Program." Washington D.C: U. S. Department of Agriculture, 2000.

Gourieroux, C. and A. Monfort. *Statistics and Econometric Models*. London: Cambridge University Press, 1995.

Harwood, J., D. Heifner, K. Coble, J. Perry, A. Somwaru. Managing Risk in Farming: Concepts, Research, and Analysis, Agricultural Economic Report Number 774, Economic Research Service, U.S. Department of Agriculture, 1999.

Just, R. E., L. Calvin, and J. Quiggin. "Adverse Selection in Crop Insurance: Actuarial and Asymmetric Information Incentives." *American Journal of Agricultural Economics*, 81(1999):834-849.

Just, R. E. and Q. Weninger. "Are Crop Yields Normally Distributed?" *American Journal of Agricultural Economics*, 81(1999):287-304.

Knight, T. O. *Examination of Appropriate Unit Discount Adjustment by Crop and Region*. Research report submitted to RMA, 1998.

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.

Long, J. S. *Regression Models for Categorical and Limited Dependent Variables*. New Delhi: Sage Publications, 1997.

Makki, S. S. and A. Somwaru. Asymmetric Information in the Market for Yield and Revenue Insurance Products. *Technical Bulletin no. 1892*, U.S. Department of Agriculture, April 2001.

Massey, Jr., F. J. "The Kolmogorov-Sminonv Test for Goodness of Fit." *Journal of American Statistical Association* 46(1951):68-78.

Moss, C. B. and J. S. Shonkwiler. "Estimating Yield Distributions with a Stochastic Trend and Nonnormal Errors." *American Journal of Agricultural Economics* 75(1993):1056-62.

Nelson, C. H. and P. V. Preckel. "The Conditional Beta Distribution as a Stochastic Production Function." *American Journal of Agricultural Economics* 71(1989):370-78.

Puelz, R. and A. Snow. "Evidence on Adverse Selection: Equilibrium Signaling and Cross-Subsidization in the Insurance Market." *Journal of Political Economy* 102(1994):236-257.

Quiggin, J., G. Karagiannis, and J. Stanton. "Crop Insurance and Crop Production: An Empirical Study of Moral Hazard and Adverse Selection." *Australian Journal of Agricultural Economics* 37(1993):95-113.

Ray, P. K. *A Manual on Crop Insurance for Developing Countries*. Rome: FAO, 1974.

Risk Management Agency. Crop Insurance for Cotton: Premium Rates and Related Matters. WWW.FCIC.USDA.GOV, U.S. Department of Agriculture, 1999.

Rothschild, M. and J. E. Stiglitz. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics* 90(1976):629-649.

Skees, J. R. and M. R. Reed. "Rate Making for Farm-Level Crop Insurance: Implications for Adverse Selection." *American Journal of Agricultural Economics* 68(1986):653-659.

Taylor, C. R. "Two Practical Procedures for Estimating Multivariate Nonnormal Probability Density Functions." *American Journal of Agricultural Economics* 72(1990):210-17.