The Design of Multiyear Crop Insurance Contracts

Ying-Erh Chen

North Carolina State University

Email: yechen@ncsu.edu

Barry K. Goodwin

North Carolina State University

Email:barry_goodwin@ncsu.edu

Introduction

Agricultural production suffers potential risks because of yield and price instabilities. These instabilities can result from various unpredictable factors, including natural disasters such as fire, drought, floods, and pest damage. Yield volatility causes price movements and also income instability for farmers. In order to help protect farmers from production, price and income risks, the Federal Crop Insurance Program provides various types of insurance.

Some insurance is based on the farm level, such as farm-level yield insurance (Multiple Peril Crop Insurance or MPCI), and farm-level revenue insurance (Crop Revenue Coverage (CRC) and Revenue Assurance (RA) coverage). The MPCI policy protects farmers against individual yield losses. Under this plan, insured farmers pay premiums based on crop yield for a specific geographical area, usually the county in which the farm is located. The Federal Crop Insurance Corporation (FCIC) provides indemnities when actual yields fall short of the farm's insured yield. In 1985, the MPCI policy changed to adopt actual production history (APH) for the insured unit in order to determine the farmers' premiums. APH is based on the average of four- to ten-years of realized yields for the insured unit.

Other types of crop insurance are based on an index or on yields at an area level such as area-level yield insurance (Group Risk Plan or GRP) and area-level revenue insurance (Group Risk Income Protection (GRIP). In GRP, each farmer chooses a coverage rate for the insured unit. Then the insurance company predicts the county yield for the current insured year based on that county's data over a span of years along an adjustment for the yield trend for every insured farmer in that area. Finally, the insurance company uses the expected county yield and coverage rate in order to calculate a yield guarantee. The most distinct characteristic of GRP is that the farmers that are insured will obtain an indemnity when the county average yield is below the

guaranteed yield regardless of the individual yield of the grower. Moreover, insured farmers under the GRP plan pay their premiums either after the crop has been harvested or when an indemnity payment is made, depending on which one is earlier.

Current MPCI and GRP are designed to mitigate monetary fluctuations resulting from yield losses for a specific year. However, yield realizations (or yield realization tendency) can vary from year to year and may depend on the correlation of yield realizations across years. Many farm programs have as their stated intention the desire of policymakers to stabilize farmers' incomes over time. If poor yield realizations can be offset by another year's better yield realizations, the actuarially fair rate, which is given by the expected loss divided by liability, is expected to decrease when current single-year MPCI and GRP contracts are extended to multiple periods. Therefore, in this proposed multiyear MPCI and GRP, insurance terms are extended to more than a year and the premium, liability and indemnity are also determined by a multiyear term. We demonstrate in this paper that, to the extent that yield and price risks are not perfectly correlated across years, significant premium savings may be possible when coverage is based upon sums or averages across years rather than on a year-by-year basis.

The proposed multiyear crop insurance plan has the following properties: (1) the plan provides lower actuarially fair premium rates than current single year plans; (2) farmers obtain a total indemnity at the end of the insured year; (3) the plan offers partial payment in the years that farmers have losses that are sufficient to cover their production costs. For example, if a farmer purchases a two-year insurance plan, the farmer will obtain total indemnity when the total yield over two years is less than the expected yield multiplied by coverage rates for two years. If the farmer experiences a serious loss in the first year, a partial payment will be paid to cover the farmer's production costs. Therefore, the farmer can have the ability to repay any loan he may

have borrowed for the first year even though total indemnity is paid in the second year. Many possible contract designs and features can be envisions within this multiyear framework. The intent of our paper is to describe and model such designs using estimates of correlation structures across years.

Our main goal is to design multiyear insurance contracts based on the joint distributions of yield across years. We will demonstrate that the actuarially fair premium rate of a multiyear insurance contract depends on the degrees of correlation of yield across years through simulation. Therefore, modeling and estimating a correlation of joint distributions for yields in a multiyear insurance design is an important task. As joint distribution for two marginal distributions may not always have a closed form expression, modeling joint distribution can become complicated.

Several approaches have been proposed in order to model the correlation between distributions without explicitly modeling the joint distribution. One commonly used approach to modeling a multivariate distribution is based on non-parametric methods such as Kernel function (Zheng et al. 2008). However, in this approach the empirical distributions are restricted to observed data and a large amount of simulations are required in order to obtain a continuous distribution (Vedenov et al. 2008).

Another school of thought uses the inverse hyperbolic sine transformation (IHST) (Johnson 1949) in order to model non-normal distributions. The IHST method was extended to model multivariate non-normal distributions that account for skewness, kurtosis, heteroskedasticity and correlation by Ramirez (Ramirez 1997). However, this approach relies on a correlation matrix that measures the dependence structure. This structure may not be straightforward to obtain in practice (Vedenov et al. 2008).

4

Recently, copula has become a popular approach in order to model joint distributions (Vedenov et al. 2008, Tejeda et al. 2008). Copula provides flexibility as a dependence function that binds marginal distributions together in order to express joint distribution without sacrificing properties of marginal distributions. In other words, when one doesn't know the form of joint distribution, one can use copula that adequately captures dependence structures of the data with reserving attractive properties of the marginal distributions. Therefore, one can use copula in order to express a multivariate distribution in terms of marginal distribution, regardless of the form of the marginal distribution. Copula methods were also adopted in our studies in order to model joint distributions of yield.

Pearson correlation coefficient by far is the most adopted dependence concept (Trivedi and Zimmer, 2005). Therefore, correlation of yield across years was estimated for farm- and county-level data based on Pearson correlation coefficient in preliminary empirical results. However, there are three limitations of Pearson correlation coefficient (Trivedi and Zimmer, 2005). The first limitation is that Pearson correlation coefficient represents a weakness of correlation as a measure of dependence because, in general, a zero correlation does not imply independence. The second limitation is that it is not defined for some heavy-tailed distribution whose second moments do not exist. The third limitation is that it is not invariant under a strictly increasing nonlinear transformation. These limitations motivate us to use an alternative measure of dependence: rank correlation such as Spearman's rank correlation. We used Spearman's rank correlation in our copula approach study. Different copula families have their definitions for correlation coefficients. Thus, the estimated correlation coefficients from copula may not be straightforward to interpret. Therefore, in order to easily interpret the correlation coefficients, we

transformed the dependence parameters estimated from copula to Spearman's rank correlation in our multiyear insurance contract.

A copula C is a multivariate joint distribution whose marginals are all uniformly distributed on the interval [0, 1].

$$C(u_1, \ldots, u_P) = \Pr(U_1 \leq u_1, \ldots, U_P \leq u_P)$$

Sklar (1959) demonstrated that there is always a p-dimensional copula C such that for all x in the domain of F where F is a p-dimensional distribution function with margins $F_1, \ldots, F_p$,

$$F(x_1, \ldots x_p) = C\{F_1(x_1, \ldots, F_p)\}$$

Archimedean Copula Family

An Archimedean copula with two random variables u and v is constructed as:

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$$

where

(1) φ is a continuous, strictly decreasing generator function from [0, 1] to [0,∞]

such that $\varphi(1) = 0, \varphi'(t) < 0, \varphi''(t) > 0$ for all $0 < t < 1$.

(2) φ^{-1} denotes the pseudo-inverse of φ

(3) C is the function from [0, 1] to [0, 1]

There are several generator functions in the Archimedean copula family, including the Clayton, Frank, and Gumbel Copulas. θ is the parameter of copula in the following copula function:

(a) Clayton Copula:

A Clayton Copula takes the form:

$$C(u,v;\theta) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}} \; where \; \theta \in (0,\infty)$$

(b) Frank Copula:

A Frank copula is constructed as:

$$C(u,v;\theta) = -\frac{1}{\theta}\log[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}]$$
$$where\; \theta \in (-\infty,\infty)$$

$$C(u,v;\theta) = -\frac{1}{\theta}\log[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}]$$

$$where\; \theta \neq 0 (parameter\; of\; copula)$$

(c) Gumbel Copula:

A Gumbel Copula is constructed as:

$$C(u,v;\theta) = \exp\{-[(-\log u)^{\theta} + (-\log v)^{\theta}]^{\frac{1}{\alpha}}\}$$
$$where\; \theta \in [1,\infty)$$

Normal Copula Family

Another popular copula family is the normal copula family. Farlie-Gumbel-Morgenstern (FGM), Gaussian and t-Copula functions are included in the normal copula family. Their densities are described as follows:

(a) Farlie-Gumbel-Morgenstern (FGM) copula:

A Farlie-Gumbel-Morgenstern copula is constructed as:

$$C(u,v) = uv + \theta uv(1-u)(1-v)$$

$$where \quad -1 < \theta < 1$$

(b) Gaussian Copula:

$$C(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} \exp\{-\frac{s^2 - 2R_{12}st + t^2}{2(1-R_{12}^2)}\} ds dt$$

Φ^{-1} denotes the inverse of the distribution function of the univariate standard normal distribution and R_{12} is the linear correlation coefficient of the corresponding bivariate normal distribution.

(c) t-Copula:

$$C(u,v) = \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{\Phi_v^{-1}(v)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} \exp\{1+\frac{s^2 - 2R_{12}st + t^2}{v(1-R_{12}^2)}\}^{-(v+2)/2} ds dt$$

where R_{12} the same definition as in the Gaussian copula. t_v is the t distribution with the degree of freedom v.

Types of Dispersion Structures

When implementing a three-year GRP, we need to consider the pairwise correlations between years. These pairwise correlations can be organized by using different dispersion structures. In this chapter, four types of dispersion structures were commonly used to fit the copula models: autoregressive of order 1 (Ar1), exchangeable (EX), Toeplitz (TOEP) and unstructured (UN) dispersion matrices. The corresponding dispersion matrix, or correlation matrices, when dimension is equal to 3 are as follows:

(1) Autoregressive of order 1 correlation matrix:

Ar1 imposes that the impact on crop yields for the current year diminishes in the following years.

$$\begin{pmatrix} 1 & \rho_1 & \rho_1^2 \\ \rho_1 & 1 & \rho_1 \\ \rho_1^2 & \rho_1 & 1 \end{pmatrix}$$

(2) Exchangeable dispersion structure:

EX imposes that the correlation among years does not vary over time.

$$\begin{pmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_1 \\ \rho_1 & \rho_1 & 1 \end{pmatrix}$$

(3) Toeplitz dispersion structure:

TOEP imposes that the impact on crop yields is consistent when the intervals between years are the same.

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{pmatrix}$$

(4) Unstructured dispersion structure:

UN imposes that the correlations among years vary over time.

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{pmatrix}$$

Parameter Estimation- Full Maximum Likelihood (FML) (Trivedi and Zimmer, 2007)

The FML method, which maximizes likelihood function, is a method used to estimate the parameters of the copulas and the parameters for the marginal distribution functions

simultaneously. The maximum likelihood estimation method maximizes the full likelihood function for the sample based on the multivariate data.

Consider the derivation of the likelihood for a bivariate model (y_1, y_2). The marginal density functions is $f_j(y_j \mid x_j; \beta_j) = \partial F_j(y_j \mid x_j; \beta_j) / \partial y_j$

and the copula derivative is $\partial C_j((F_1(x_1; \beta_1), F_2(x_1; \beta_1); \theta) / \partial F_j$ for $j = 1, 2$

The copula density is $c(F_1, F_2) = \dfrac{d}{dy_2 dy_1} C(F_1, F_2) = C_{12}(F_1 . F_2) f_1 f_2$ where

$$C_{12}((F_1 \mid x_1, \beta_1), (F_2 \mid x_2; \beta_2); \theta) = \partial C((F_1 \mid x_1; \beta_1), (F_2 \mid x_2; \beta_2); \theta) / \partial F_1 \partial F_2$$

and the log-likelihood function is

$$L_N(y_1 \mid x_1, \beta_1), (y_2 \mid x_2; \beta_2); \theta)$$
$$= \sum_{i=1}^{N} \sum_{j=1}^{2} \ln f_{ji}(y_{ji} \mid x_{ij}; \beta_j) + \sum_{i=1}^{N} C_{12}[F_1(y_{1i} \mid x_{1i}; \beta_1),); F_2(y_{2i} \mid x_{2i}; \beta_2); \theta]$$

FML estimates are obtained by solving the score equations $\partial L_N / \partial \Omega = 0$ where $\Omega = (\beta_1, \beta_2, \theta)$. FML will be used for parameter estimation in our real data analyses.

The multiyear revenue insurance plan will be constructed as a hybrid of a single year and multiyear revenue insurance plan coverage. That is, farmers will obtain a partial payment of indemnity in each year if their actual revenue in that year falls below a certain level. Also, farmers will obtain an indemnity at the end of insured year if the average of the actual revenue for two years falls below a certain level. The partial payment of indemnity will be made to mitigate revenue loss in each year. Therefore, farmers are assured that the revenue loss would not bankrupt them because this partial payment of indemnity provides a safety net to protect farmers from the revenue loss in some degree. If farmers have an operation loan from banks, they will be able to pay the loan and keep their good credit through this partial payment of

indemnity. This will be more attractive to lenders as loans will be better collateralized. Therefore, farmers with good credit can continue to take out loans for production costs for the next period. To farmers, then, this proposed multiyear revenue insurance plan would also be more attractive than the current revenue insurance plan because of the lower actuarially fair premium rate and the partial payment of total indemnity.

Total indemnity is made when the average revenue over two years is less than the guaranteed revenue. We present the total indemnity determination of a two-year insurance plan in the following equation:

$$\frac{R_1^A + R_2^A}{2} < \beta * (\frac{E(R_1) + E(R_2)}{2})$$

Where

$R_i^A = actual\ revenue\ at\ year\ i, i = 1,2$

$E(R_i) = \mu_i\ expected\ revenue\ at\ year\ i$

$\beta = coverage\ rate, \beta \leq 1$

Partial payment of total indemnity is the difference between the threshold $(\gamma \mu_i)$ and the actual revenue (R_1^A, R_2^A), where γ is the coverage rate for the partial payment of total indemnity and μ_i is expected yield in year i. The ratio γ can be calculated in the following:

$$\gamma = \frac{estimated\ variable\ cost}{predicted\ price}$$

Simulation-Actuarially Fair Premium Rate and Correlation across Years

In this section, we will show the simulation results for the actuarially fair premium rates based on different correlation of yields among different years for the two-year and three-year yield crop

insurance plans. For simplicity, we only simulated yield distributions for yield crop insurance plans. However, the results can be applied to revenue insurance plans as well. Several assumptions were made in the simulations. These assumptions are described as follows.

(a) The yield distribution for each year follows a Beta distribution with the shape parameters

$\alpha = 2, \beta = 1.5$ and data ranging from 0 to 200.

(b) The correlation coefficient of yield distributions between the first year and second year varies between -1 and 1. The correlation coefficient of yield distributions among the first, second and third years is between -0.5 and 1.

(c) The farmer selects the 70% coverage rate and the expected yield is 160.

Indemnity is paid if the averaged yield over two or three years is below the guaranteed yield, which is 0.7×160=112. Then the actuarially fair rates are calculated based on the loss function. Results are shown in Tables 1. In Table 1 , $rate_{12}$ and $rate_{123}$ are the actuarially fair premium rates of two-year and three-year yield crop insurance plans, respectively; rho is the Spearman correlation coefficient of yield distributions among the first, second and third year.

**Table 1. Simulation Result for Two-and-Three Year Yield Crop Insurance Plan.**

Observation	$rate_{12}$	$rate_{123}$	rho	Observation	$rate_{12}$	$rate_{123}$	rho
1	0	N/A	-1	11	0.11	0.087	0
2	0.03	N/A	-0.9	12	0.116	0.097	0.1
3	0.044	N/A	-0.8	13	0.122	0.097	0.2
4	0.055	N/A	-0.7	14	0.129	0.114	0.3
5	0.066	N/A	-0.6	15	0.133	0.121	0.4
6	0.074	0.007	-0.5	16	0.14	0.13	0.5
7	0.082	0.035	-0.4	17	0.144	0.137	0.6
8	0.091	0.052	-0.3	18	0.149	0.144	0.7
9	0.097	0.066	-0.2	19	0.155	0.152	0.8
10	0.104	0.076	-0.1	20	0.162	0.158	0.9
				21	0.166	0.165	1

Note: The rate for a single year plan is 0.165, which is the same as rates for multi-year plans

when rho=1.

From Table 1, several conclusions can be made: (1) the actuarially fair premium rates for two-year and three-year yield crop insurance plans are at the minimum (i.e. 0) when the correlation coefficients are close to -1 and -.5, respectively; (3) the actuarially fair premium rates for two-year and three-year yield crop insurance plans are at the maximum (.166 and .165, respectively) when the correlation coefficient is 1, which is the same as the actuarially fair premium rate of a one-year yield crop insurance plan; and (4) the actuarially fair premium rates of the two-year and three-year yield crop insurance plan with rho<1 are lower than the actuarially fair premium rate for a single year. We can use the following graph to present this simulation result for the relationship between correlation across years and actuarially fair premium rate.
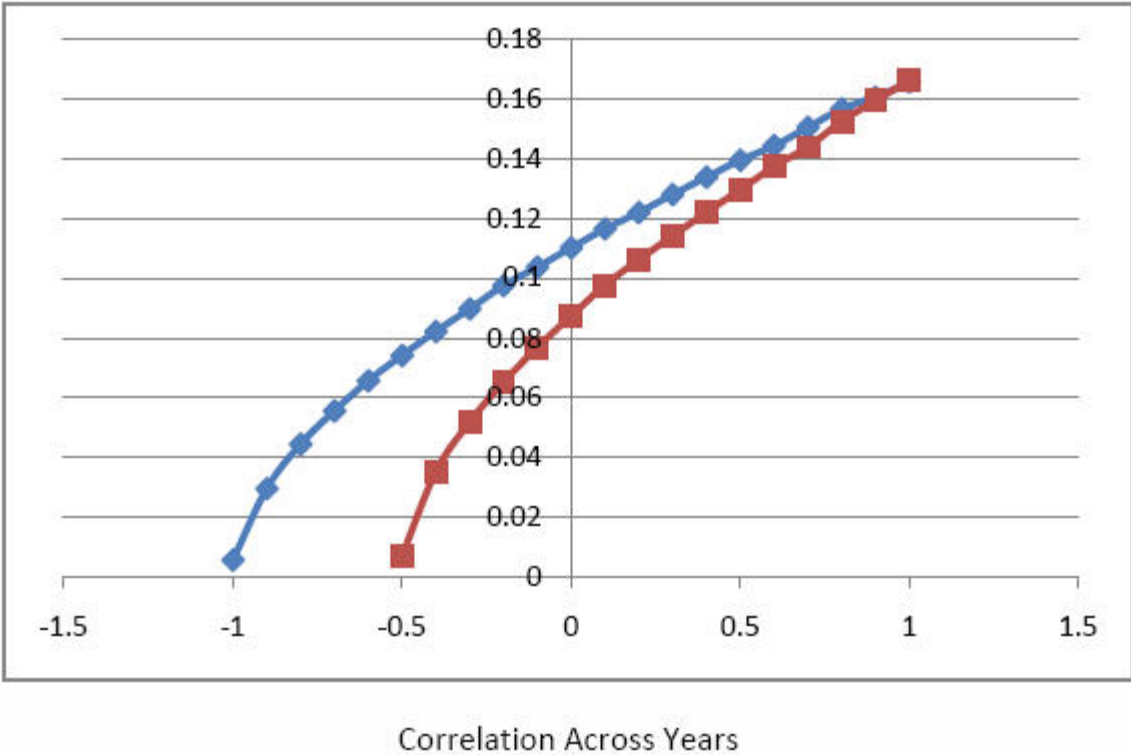
Figure 1. Relationship between Actuarially Fair Premium Rate and Correlation across Years

Preliminary Empirical Results for Multiyear Yield Crop Insurance Plan

The simulation results suggested that when the Spearman's rank correlation of yields across years is not perfectly positive correlated, the actuarially fair rate can be lower in the multiyear yield insurance plan than the single-year plan. In this section, we used farm and county-level data for Iowa corn to investigate the presence and level of correlation of yield across years from Pearson correlation coefficient estimates. Though we mentioned that Pearson correlation coefficient may be a weak correlation as a measure of dependence, Pearson correlation coefficient results estimated from real data can be viewed as an indicator if multiyear insurance plan can possibly offer lower premium in real case because it's a simple calculation. As long as Pearson correlation coefficient from real data is not close to 1, we have good reason to believe that the implementation of multiyear insurance plan can offer lower actuarially fair premium rate in practice. We will see that Pearson correlation coefficients estimates suggest that correlation of yield across years is not significantly correlated in most cases on the farm and county-level data. Therefore, this motivates us to obtain better dependence parameter estimates than Pearson correlation coefficient from the copula approach and estimate the actuarially fair premium rates of multiyear insurance plans based on the copula models.

Data

The farm-level data used in this study were collected from the Risk Management Agency (RMA) of the USDA. We examined corn yield data in Iowa and this farm-level data set consists of yearly yield quantity, insured acres of yield measured, the years of yield measured, county and state location, the type of insurance plan and practice for each insured farm. This farm-level data

set consists of 99 counties containing 42722 observations of Iowa corn. The beginning years and ending years of farm data vary from farm to farm. The beginning years range from 1986 to 1989, and the ending years range from 1995 to 1997. Data which are not continuous for some farms were excluded from our analysis due to the difficulty of measuring correlation of yields between two years. A "transitory yield" is assigned to a farmer whose yield history is not sufficient to calculate actuarially fair premium. Transitory yields were excluded in our study.

Since each yield observation in this farm data is continuous for 10 years in the same location, we don't have to consider the rotation practice issue in our analysis. All farms in this data set insured their crops on at least one occasion during the time period examined and the locations (county and state) for each insured farm were reserved in the data. This allowed us to aggregate farm-level yields under different insurance plans and practices for each county, and estimate yield correlation across years (Pearson correlation coefficient) on the county- and state-level. The aggregated farm-level data on the county- and state-level allowed us to (1) investigate the insured yield correlation on the county- and state-level across years; (2) compare insured yield correlation on the county level across years with yield (including uninsured and insured farms) correlation across years from county-level data; (3) investigate the level of spatial correlation of yield correlation across years among counties.

Other than analyzing the farm-level data for Iowa corn, we also analyzed the county-level data. County corn yield data from 1928 to 2007 in Iowa were obtained from the Risk Management Agency (RMA) of the USDA. The average county yield for 80 years in Iowa was calculated, and the top ten production counties in Iowa were chosen for analysis.

Production costs were collected for Iowa from the State Extension Services. We used variable costs and futures to calculate coverage rates for partial payment. Table 2 shows the values of the

variable costs in different states, futures prices and futures periods for different crops. Variable

costs include fertilizer, seed, pesticide, dryer haul, machinery fuel, machinery repairs, hauling,

and interest on pre-harvest variable costs in our study. Iowa State Extension Service provides

variable costs at different yield levels and only one of the yield levels was chosen in our study.

We used futures prices from the Chicago Board of Trade for corn. Then the coverage rates of

partial payment, which is also shown in Table 2, were calculated based on the variable cost

divided by the futures.

Table 2. Coverage Rate of Partial Payment for Iowa Corn

Year of budget	State	Commodity	Practices	Period of futures	Variable cost based on the yield level	Coverage rate of Partial Payment
2009	Iowa	Corn	Corn following corn	Dec/2010	145bu./ac	2.99/4.45[*]=.67

Note: The denominator is the futures from Chicago Board of Trade.

Remove Time Trend for County Data

With the improvement of technology, crop yields increase over time. Thus, we need to remove

the time trend for the yield data collected over different years. First we fit the time trend model

for crop yields over years. The relationship between crop yield y_t and time could be represented

as

$$y_t = X_t \beta + e_t \qquad\qquad ...eq(1)$$

where X_t represents the linear or nonlinear function of time.

If crop yields increase over time linearly, we can use the following trend model to fit the

relationship between crop yield and time:

$$y = \beta_0 + \beta_1 t + e_t$$
$$\text{where } X_t = (1,t) \text{ and } \beta = (\beta_0, \beta_1) \qquad \qquad ...eq(2)$$

If crop yields grow quadratically, we can use the following trend model to fit the relationship between crop yields and time:

$$y = \beta_0 + \beta_1 t + \beta_2 t^2 + e_t$$
$$\text{where } X = (1,t,t^2) \text{ and } \beta = (\beta_0, \beta_1, \beta_2) \qquad \qquad ...eq(3)$$

Regression analysis suggested that corn yields of county data in our analysis grow quadratically (i.e. the null hypothesis of $\beta_2 = 0$ is rejected), and thus eq (3) was adopted to remove the time trend. After we fit the time trend model for crop yield over time, we obtained trend-predicted yield data $(\hat{y}_t)$ and deviation from the trend (e_t). Since lots of empirical studies support the idea that deviations from the time tend to be proportional to crop yield, we use the following equation to normalize crop yields over time (Miranda and Glauber, 1997):

$$\tilde{y}_t = \hat{y}_t (1 + \frac{e_t}{\hat{y}_t}) \qquad \qquad \dots eq(4)$$

where

$\hat{y}_t$ is detrended yield data of last observation

$\tilde{y}_t$ is normalized yield data over time

Empirical Results

In this section, we will present the following: (1) Pearson correlation coefficient of corn yields of county data for each county between two consecutive years; and (2) Pearson correlation coefficients of the corn yields for each county between two consecutive years, after we aggregated the farm data to a county level; (3) Pearson correlation coefficient of the corn yield

between two consecutive years, after we aggregated the farm data to a state level. In our results for p-values, p-values < 0.0001 were replaced by zero.

Top ten production counties of Iowa corn were chosen. After detrending the normalized county yield data, Pearson correlation coefficients were calculated and the significance of the correlations between y_1 and y_2, which are yields between two years, were tested. Pearson correlation coefficients and p-values are summarized in Table 5. In Iowa, Pearson correlation coefficients are positive except in Scott County. No p-value in these counties is significant when the significant level is equal to 0.01 or 0.025. In other words, there is no significant correlation of yield across yeas in these 10 counties in Iowa. From these Pearson correlation coefficients estimated from county data in Iowa, we can conclude that yield across years are not significantly correlated when significant level is equal to .01 and .025.

**Table 5**. Pearson Correlation coefficient results for Iowa corn (county-level data)

Iowa			
Cedar County	0.012 (0.091)	Hamilton County	0.156 (0.168)
Scott County	-0.014 (0.089)	Wright County	0.215 (0.056)
Ground County	0.156 (0.179)	Webster County	0.197 (0.081)
Marshall County	0.137 (0.225)	Humboldt County	0.15 (0.184)
Hardin County	0.124 (0.274)	Boone County	0.161 (0.155)

Pearson Correlation Coefficient for Farm Data

Pearson correlation coefficient and p-value for Iowa corn at the state level based on the aggregated farm data is -0.362 and 0.337, respectively. The statistics indicate the correlation across years and the significance of correlation. Iowa corn has negative correlation coefficient across years and this may suggest that moral hazard and adverse selection may result in unstable corn yield in Iowa. P-value of Iowa corn is not significantly correlated at level 0.05. The correlation results of Iowa corn can provide a guideline for government agency to prioritize states to implement the multiyear insurance plans. For example, the actuarially fair premium rate for Iowa corn can be much lower based on the multiyear insurance plan than single year plan due to the negative correlation.

Next, we show Pearson correlation coefficients for corn at the county level based on the farm-level data through the following correlation maps. The following figure provides Pearson correlation coefficient and p-value results for farm data in graphical form at the county level. We used four colors in the maps to show four levels of Pearson correlation coefficients. Moreover, we used histograms to show p-value results for most counties.

For Iowa corn in Figure 2, Pearson correlation coefficient increases from the western to eastern Iowa. From Figure 3 the average annual precipitation increases from the western to eastern Iowa. Figure 2 and 3 suggest that (1) precipitation may be a key factor to influence yield stability of Iowa corn; (2) Iowa corn is more stable in eastern Iowa than west and thus Iowa corn may prefer to grow under average annual precipitation between 34-38 inches based on the precipitation map.
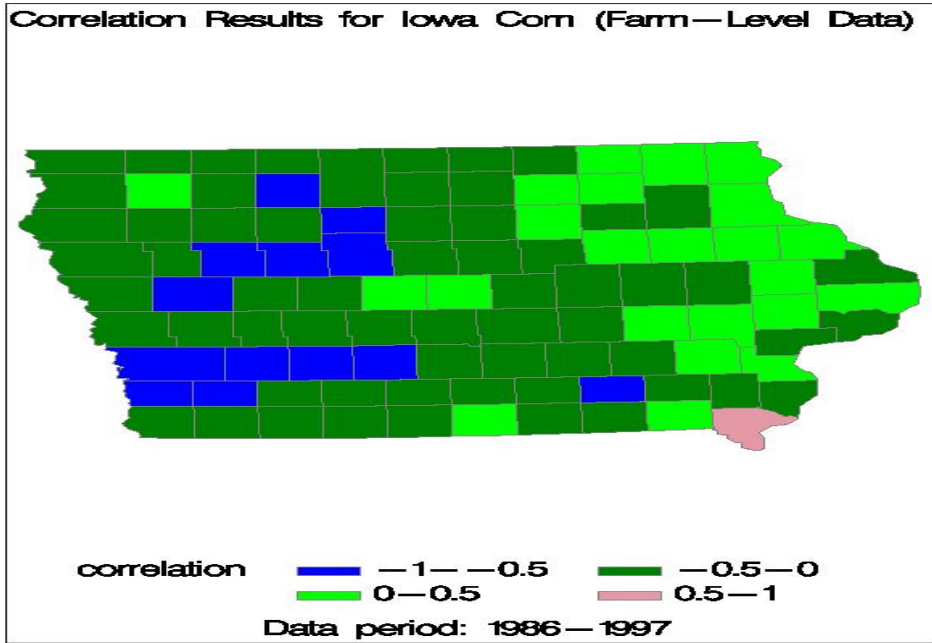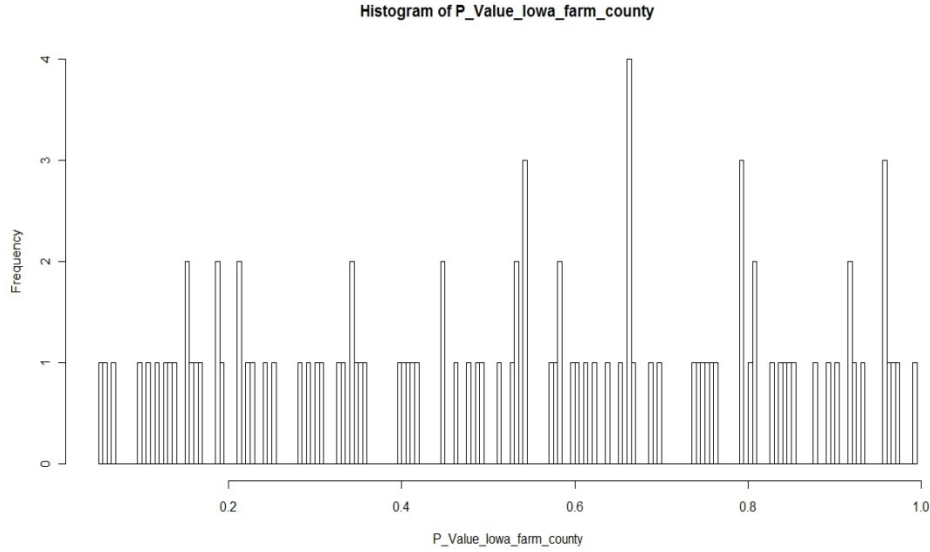
Figure 2. Correlation Result for Iowa Corn (Farm-Level Data)
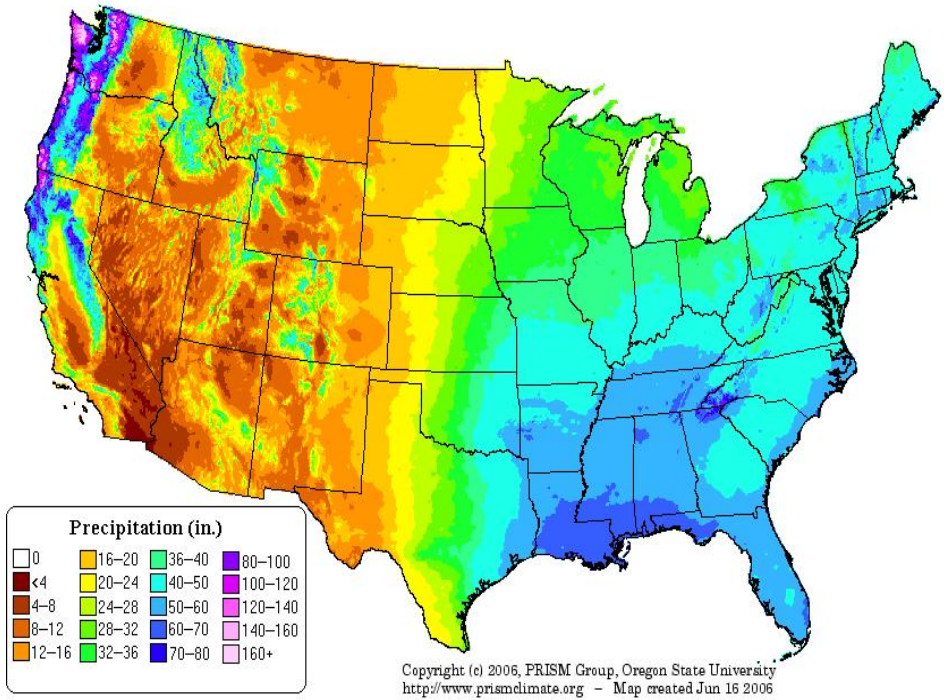
Figure 3. Annual Precipitation Map in the United State from 1971 to 2000

Figure 4. Histogram of P-Values of Iowa at County Level (Crop: Corn)

A Two-and-Three-Year Insurance Contract Demonstration-Adair in Iowa

We used eq(3) and eq(4) to detrend and normalized corn yield data for Adair in Iowa and the parameter estimates for detrended data are shown in Table 6. After we obtained the detrended normalized data, we used Goodness-of-Fit Tests for Normal and Beta Distributions to see which distribution supports the detrended normalized data. From Table 7, the tests are rejected at the 0.05 significance level so that the detrended normalized data doesn't have a normal distribution. From Table 7, the tests cannot be rejected at the significance levels (1% and 5%), so the detrended normalized data may follow a Beta Distribution.

To model the dependence structure of joint distribution for the two-year insurance plan, we used the "Copula" package (Yan, 2006) provided in R to model the copula distributions. We used several copula functions (Normal, Clayton and Flank Copula) with Beta marginals to estimate parameters of Beta distributions and copula parameter simultaneously; we then selected the

copula which fits the data best according to their maximum likelihood values and Akaike's information criterion (AIC) for two- and three- year insurance plans, respectively. In the following Tables, α_1, α_2, α_3, β_1, β_2 and β_3 are estimated shape parameters for Beta Distributions; and $\rho, \rho_1, \rho_2, \rho_3$ are Spearman's correlation coefficients converted from the copula parameters using the Adair County yield data in Iowa.

Table 6. Parameter Estimates for Detrended Data

Variable	Parameter Estimate	Standard Error
Intercept	32.48	5.482
t	0.53	0.312
t^2	0.013	0.003

Table 7. Goodness-of-Fit Test for Normal and Beta Distribution

Goodness-of-Fit Test for Normal Distribution		
Test	Statistic	p-value
Kolmogorov-Smirov	0.117	<0.01
Cramer-von Mises	0.204	<0.005
Anderson-Darlilng	1.285	<0.005
Chi-Square	12.411	0.015
Goodness-of-Fit Test for Beta Distribution		
Chi-Square	6,472	0.167

For the three-year plan, we used the Beta distributions as marginal distributions to fit to Normal, Clayton and Frank copula models with the autoregressive of order 1 correlation structure, Exchangeable dispersion structure, Toeplitz dispersion structure and Unstructured dispersion structure. We only present the Frank Copula fitting results with Beta marginals and the UN correlation structure under the three-year GRP because this fitting has the lowest AIC values among copula models with different correlation structures.

Table 8. Estimation results of Copula model

Estimation Result for Frank Copula with Beta Marginals under Two-Year GRP		
Variable	Parameter estimate	Standard error
α_1	7.781	1.234
β_1	3.596	0.548
α_2	7.8	1.237
β_2	3.614	0.551
Spearman's correlation(ρ)	0.467	0.711
Maximized likelihood	100.043	
Estimation results for Frank Copula with Beta marginals and UN Correlation structure under three-year GRP		
α_1	7.795	1.24
β_1	3.667	0.56
α_2	7.684	1.231
β_2	3.58	0.549
α_3	7.723	1.233
β_3	3.613	0.553
ρ_1	0.029	0.113
ρ_2	0.175	0.109
ρ_3	0.002	0.113
AIC=314.98		

Based on the results above, we simulated a joint yield distribution with two marginal beta distributions based on the correlation we obtained from the copula function. Then we estimated the actuarially fair premium rates for two-year crop insurance from the joint yield distributions. We also used the same process to estimate actuarially fair premium rates for three-year insurance contract. These results are shown in Table 9.

After we obtained the coverage rates for partial payments from Table 2, we demonstrated how partial payment can help farmers repay their debt when they have yield loss in either one of the two years or both years. The results are shown in Table 9. The data period is from 1926 to 2005 and the Spearman correlation coefficient between two years is 0.072 estimated from the Frank Copula. We then simulated a joint yield distribution with two marginal beta distributions with 0.072 correlation between years. Please see the parameters of beta distributions estimated by the Frank Copula in Table 9. The estimated actuarially fair premium rate based on the simulated model is 0.003 for a two-year insurance contract. We assumed the expected yield is 160 bushels per acre and the coverage rate for the total indemnity is 70%. In Table 2, the coverage rate for the partial payment is 67% of Iowa corn. We assume the realized yield is 90 and 135 bushels per acre in the first and second insured years, respectively, and the insured price is $2.5 per bushel. The average farm size of Iowa corn is 242 acres (2009, Goodwin). Since yield in the first insured year is below the guaranteed yield, farmers are eligible to obtain the partial payment with $13,310, as shown in Table 9. We also assumed farmers have a loan equal to the total variable cost. Therefore debt is the product of the total cost per acre times the farm size, which is equal to $109831.7. We assumed farms are engaged in the futures market. Therefore, the expected revenues will be the product of futures price times yield per acre and then times farm size, which

is equal to $96961. Therefore farmers are able to repay loan based on the partial payment and expected revenue.

An example of the three-year contract is also shown in Table 9. The Spearman correlation coefficients estimated from the Frank Copula are 0.0028, 0.0087 and 0.0020 between years one and two, years one and three, and years two and three, respectively. The estimated actuarially fair rate is 0.0009 based on the simulated model. The assumptions we made for the two-year insurance contract hold for the three-year insurance contract. These assumptions include the amount of expected yield, coverage rates of total indemnity and partial payment, the amount of yield in the first and second year, the insured price and farm size. We further assumed the realized yield in the third year is 100 bushels per acre. Since the amount of yield in the first and third insured years is below the guaranteed yield, farmers will obtain partial payment in these two insured years with $13,310 and $7,260, respectively. In the third insured year, though the total yield for the three insured years is below the guaranteed yield, farmers will not obtain indemnity in the last insured year. That is because the partial payment made in the first and third year is higher than the total indemnity, which is $6,655 as shown in Table 9.

Table 9. Adair County in Iowa (corn) for Two-and-Three-Year Insurance Contract

	Two-Year Insurance Contract	three-year insurance contract
Spearman's rank correlation	.072	0.0028(1&2), 0.0087(1&3), 0.0020(2&3)
Parameter of Beta Distribution	$(\alpha_1,\beta_1)=(7.78,3.59)$ $(\alpha_2,\beta_2)=(7.8, 3.61)$	$(\alpha_1,\beta_1)=(7.79,3.66)$ $(\alpha_2,\beta_2)=(7.68, 3.58)$ $(\alpha_3,\beta_3)=(7.72,3.63)$
Actuarially fair rate	0.003	.0009
Expected yield	160[1]	160
Coverage rate of guaranteed yield for two-(and three) year insurance contract	70%	70%
Guaranteed yield for two(three) years	224(=2*70%*160)	336(=3*70%*160)
Coverage rate of partial payment	67% (from Table 1)	67% (from Table 1)
Guaranteed yield in each year	112(=67%*160)	112(=67%*160)
Assumed yield in year 1 & 2 (&3)	90, 135	90,135, 100
Partial payment	13310(=2.5*(112-90)*242) in year 1	(1) 13310 (=2.5*(112-90)*242) in year 1 (2) 7260 (=2.5*(112-100)*242) in year 3
Total indemnity	N/A(90+135>224)	N/A 6655[2]=2.5*(336-90-135-100)*242
Short-term Debt	109831.7(=3.13*145*242)	
Expected revenues in year 1	96921(=4.45*90*242)	

Note 1: Yield is on a per acre basis.

Note 2: total indemnity will not be paid in the third insured year because $6655 is lower than the sum of $13310 and 7260.

Conclusion

In summary, we proposed multiyear insurance plans that provide lower premium rates and can be attractive for farmers. We used simulations to demonstrate that actuarially fair premium rate for the multiyear rates were lower than single-year plans when correlation of yield between years is less than 1. We provided comprehensive estimates of Pearson correlation coefficients for yields between two years at the state and county level based on the farm-level data. We also investigate correlation patterns at county level in Iowa. We can also see correlation patterns vary from county to county, mostly due to geographic locations and weather patterns. Based on the histograms of p-values in Figure 4, the correlation is not significant. Our simulation results suggested that the multiyear insurance plan has advantages over current plans when the correlation coefficient is less than 1. Hence, the results of correlation provide solid evidence that the proposed multiyear insurance plan will have lower actuarially fair premium rate than current single year plans in practice. Moreover, our estimates can provide a guideline for government agency to decide and prioritize counties to implement the multiyear plans.

Other than estimating correlations of yields between two consecutive years using the sample Pearson correlation coefficient, we also estimated the correlations using the copula method. The copula method implicitly models the marginal distributions and thus may provide better estimate of the correlation than Pearson correlation coefficient.

We also showed a multiyear insurance contract design example (Adair County in Iowa) which demonstrated the implementation in details. We showed in the example that how farmers can obtain partial payment each year and total indemnity at the end of the insurance term. We anticipate that the proposed plans will be of interest to both government agencies and farmers.

Reference

Dmitry Vedenov. "Application of Copula Estimation of Joint Copula Yield Distributions." Selected Paper prepared at the American Agricultural Economics Association Annual Meeting, 2008.

Hernan A. Tejeda, Barry K. Goodwin. "Modeling Crop prices through a Burr distribution and Analysis of Correlation between Crop Prices and Yields Using a Copula method." Selected Paper prepared at the American Agricultural Economics Association Annual Meeting, 2008.

Johnson, N.L. "Systems of Frequency Curves Gener- ated by Methods of Translation." Biometrikaff. 36(1949):149-76

Mario J. Miranda and Joseph W. Glauber. "Systemic Risk, Reinsurance, and the Failure of Crop Insurance Markets." American Journal of Agricultural Economics 79 (February 97):206-15.

P. K. Trivedi and D. M. Zimmer. Copula Modeling: An Introduction for Practitioners, Foundations and Trends R in Econometrics, vol 1, no 1, pp 1–111, 2005.

Qiujie Zheng, H. Holly Wang, Qinghua Shi. "Estimating Multivariate Yield Distributions Using Nonparametric Methods." Selected Paper prepared for at the American Agricultural Economics Association Annual Meeting, 2008.

Ramirez, O.A. "Estimation and Use of a Multivariate Parametric Model for Simulating Heteroskedastic, Correlated, Nonnormal Random Variables: The Case of Corn Belt Corn, Soybean, and Wheat Yields." American Journal of Agricultural Economics, 79, (February, 1997): 191-205.

Yun Yan."Enjoy the Joy of Copulas: With a Package copula." Journal of Statistical Software 21(4) 2007.