Modeling Crop prices through a Burr distribution and Analysis of Correlation between Crop Prices and Yields using a Copula method

Hernan A. Tejeda

Graduate Research Assistant

North Carolina State University

email: hatejeda@ncsu.edu

Barry K. Goodwin

William Neal Reynolds Distinguished Professor

North Carolina State University

email: barry_goodwin@ncsu.edu

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Orlando, Fl, July 27-29, 2008

Modeling Crop prices through a Burr distribution and Analysis of Correlation between Crop Prices and Yields using a Copula method

Hernan A. Tejeda and Barry K. Goodwin

Abstract

The U.S. crop insurance program has major policy implications in terms of resource allocations, with government subsidies playing a major role. Efficient implementation of crop revenue insurance contracts requires accurate measures of risk for both crop prices and yields. In addition, rating methods should consider the natural hedge between prices and yields. Empirical evidence shows that crop prices tend to be positively skewed with fat tails while crop yields tend to exhibit negative skewness. This paper analysis is two-fold. It first studies crop prices using a Burr distribution, with parameters that capture skewness and kurtosis (fat tails), providing a better fit than normal or log-normal distributions currently being used. It then uses a copula method to measure the correlation between crop prices and yields - for the study of crop revenue insurance. Results indicate a smaller probability of payout than present methods being used, having direct implications on the design and rating of crop and revenue insurance contracts.

Key Words: Crop insurance, Burr XII distribution, Copula methods, indemnity payouts

---

Hernan A. Tejeda is a Graduate Research Assistant in the Departments of Economics and Agricultural and Resource Economics at North Carolina State University. Corresponding author, email: hatejeda@ncsu.edu and/or at P.O. Box 8110, Raleigh, NC, 27695

Barry K. Goodwin is William Neal Reynolds Distinguished Professor in the Departments of Economics and Agricultural and Resource at North Carolina State University.

Introduction

Crop insurance is of critical importance in the farming business to properly address production and/or revenue shortcomings for farmers and/or crop producers. Since the inception of the Federal Multiple-Peril Crop Insurance (MCPI) program in 1938 by means of the Crop Insurance Act, there has been continuous updating of federal programs to improve the application and efficiency of its use. For extensive coverage of the history of Federal MPCI through each decade until the mid 1990's, see Goodwin and Smith (1995). Another means of supporting crop producers during unanticipated devastating events has been to provide aid through federal disaster relief programs. For the theory of relation between this federal disaster aid and crop insurance, see Goodwin and Smith (1995), and for an overall historical review, see Goodwin and Vados (2007). Regarding the MCPI, the latest program change occurred in 2000, with the enactment of the Agricultural Risk Protection Act (ARPA). This change includes comprehensive sections in crop insurance coverage and agricultural assistance by increasing government subsidies, among others. For details see the Agriculture Risk Protection Act - Public Law 106–224; 2000

A major problem with the use of crop insurance has been the excessive payouts compared to the premium rates paid by farmers, see Goodwin and Smith 1995, and Goodwin 2001; generating substantial losses to the federal government. Many of these losses are a result of the premium subsidies paid by the government; hence a proper calculation of the premium rates is critical to an efficient implementation of any crop insurance program.

Currently, the primary crop revenue insurance programs in place are the Crop Revenue Coverage (CRC) and Revenue Assurance (RA) programs, which calculate premium rates considering the estimated joint distributions of yields and crop prices. These rating methods must give consideration to the correlation between prices and yields and the natural hedge that is implied by them. It is important to mention that presently all crop revenue insurance programs are under review to assess their future implementation. Basically, these programs will undergo changes such that they will all be incorporated into a single package - offering crop revenue insurance. Details for this are available through USDA. This study aims to provide a new perspective in the method of analyzing crop revenue through the joint relation between crop yields and prices.

In the present case of CRC, premium rate calculations assume crop prices follow a normal distribution[1] and for the case of RA, premium rates are calculated under the assumption that crop prices follow a log-normal distribution. For extensive details of these methods and discussion of their shortcomings see report GAO 98-111. Despite these assumptions, empirical evidence and conventional wisdom holds that prices tend to have a positively skewed distribution, with fatter (kurtosis) tails than Normal distributions. See Goodwin and Ker (2002) for an extensive review. This is one aspect of the premium rate calculation which will be addressed in this paper. Another aspect that will be addressed is how the premium rates are currently calculated for each revenue insurance program. We also compare this to a proposal for a different method considering a better depiction of the relation between crop prices and yields.

This paper begins by adopting a Burr type XII distribution to characterize crop prices, and compares the goodness of fit to these prices relative to Normal or Log-Normal distributions. The benefit conveyed by the Burr distribution is that it considers parameters that capture the higher moments observed in the data, hence skewness and kurtosis may be properly portrayed. See Klotz and Johnson (vol. 1, 1981.)

The second analysis considers the use of a Copula method to assess the relation between crop prices and their yields. Copulas are a convenient statistical method of measuring the correlation between variables by just considering the marginal distributions of these variables or their non-parametric marginal distributions (in the empirical case of having a large number of observations). In other words, there is no need to have previous knowledge of the degree of correlation between the data sets and their distributions, in order to calculate their joint distributional relation through a copula. As for the marginal distribution of

[1] In rigor, prices are assumed to distribute as a truncated normal (see pg. 52, GAO 98-111)

the crop yields, these are modeled through a Beta distribution - as they tend to be left or negatively skewed. See Goodwin and Ker (2002), and Gallagher (1987).

Empirical Methods

The ***Burr type XII distribution*** considered for modeling the crop prices has the following characteristics:

(3 parameters – 2 shape, 1 scale)

c.d.f. $F_{B12}(y;\alpha,\tau,\phi) = 1-\left\{1+\left(\frac{y}{\phi}\right)^{\tau}\right\}^{-\alpha}$ for y ≥ 0 and $\tau: shape1$; $\alpha: shape2$; $\phi: scale$

p.d.f. $f(y) = \left(\tau * \alpha * \left(\frac{y}{\phi}\right)^{\tau}\right)\{y^{-1} * \left[1+\left(\frac{y}{\phi}\right)^{\tau}\right]^{-(\alpha+1)}\}$

The parameters of this distribution are estimated via a Maximum likelihood method as per Watkins (1999) and Johnson (2003). Another distribution of the Burr family – the Burr type III distribution, where the variable considered is the inverse of the previous Burr type XII, was also estimated via the method of moments, see Lindsay et. al (1996). A comparison of results by simple mapping - revealed that the Burr type XII distribution characterized better the data.

The *Normal* and *log-Normal* distributions respectively; have the known characteristics below:

c.d.f. $F_N(y;\ \mu,\sigma) = \Phi(\frac{y-\mu}{\sigma})$ for y ϵ $\mathbb{R}$, $\mu: mean \in \mathbb{R}$, $\sigma: standard\ deviation > 0$

p.d.f. $f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{\frac{-(y-\mu)^2}{2\sigma^2}\}$

Log – Normal:

c.d.f. $F_{LN}(y;\ \mu,\sigma) = \Phi(\frac{\ln(y)-\mu}{\sigma})$ for $y > 0, \sigma > 0$ $\mu\ (location), \sigma(scale): both\ in \log space$

$$mean = \exp\left(\mu + \frac{\sigma^2}{2}\right); \quad variance = \exp(2\mu+\sigma^2)\left(\exp(\sigma^2)-1\right)$$

p.d.f. $f(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\{\frac{-(\ln(y)-\mu)^2}{2\sigma^2}\}$

Both Normal and log-Normal distributions are also calculated via maximum likelihood, and their likelihood results are contrasted to that obtained with the Burr type XII distribution. A test of Voung – see Voung (1989), which is a non-nested test is made here to ascertain the improvement of the Burr XII distribution over the previous two distributions.

Separately, a Beta distribution is used to model the crop yield data as it is usually negatively skewed. As mentioned above, this parametric distribution delivers appropriate results in the case of having limited amounts of data. However, for the alternative of having large amounts of observations, then non-parametric methods may be more suitable - see Goodwin and Ker (1998).

A copula method is then used to assess the correlation between these two variables, crop prices and their yields. A copula is basically a function that 'couples' together a multivariate function to their one-dimensional marginal distributions. Or in other words, a copula is a multivariate distribution function that has one-dimensional marginal functions that are uniform on the interval [0,1] - see Nelsen 1999.

Formally defining this previous concept:

Definition of a Copula (by Sklar's theorem): (see Embrechts et. al, Chapter 8, 2003).

Let H be an n-dimensional distribution function with marginals $F_1, \dots \dots \dots, F_n$.

Then there exists an n-copula C such that for all **x** in $\mathbb{R}^n$

$$H(x_1, \dots \dots \dots, x_n) = C(F_1(x_1), \dots \dots \dots, F_n(x_n)).$$

For $F_1, \dots \dots \dots., F_n$ being all continuous, then C is unique. Conversely, if C is an n-copula and $F_1, \dots \dots \dots., F_n$ are (cumulative) distribution functions, then the function H defined above is an n-dimensional (cumulative) distribution function with marginals $F_1, \dots \dots \dots., F_n$.

i.e. for a univariate distribution function F, the generalized inverse of F, is:

$F^{-1}(t) = \inf\{x \in \mathbb{R} \mid F(x) \geq t\}$ for all t in [0,1]. Then for any $\boldsymbol{u}$ in $[0,1]^n$,

$C(u_1, \ldots \ldots \ldots \ldots \ldots \ldots, u_n) = H\left(F_1^{-1}(u_1), \ldots \ldots \ldots \ldots \ldots .. F_n^{-1}(u_n)\right)$

Elliptical copulas (Normal or t-student) are restricted to radial symmetry and don't have a closed form.

The Normal Copula distribution is (see Freez and Valdez 1998):

$C(u, v) = H[\Phi^{-1}(u), \Phi^{-1}(v)]$ for $\theta\ (copula\ parameter) \epsilon\ [-1,1]$

Yet a different type of copulas is also available. These are the Archimedean type copulas, and they may be preferred in our analysis because they have a closed form and also capture asymmetric correlation between the tails of the marginal distributions (i.e. different dependence at one end of the tail than at the other end). See Embrechts et al. (Chapter 8. 2003).

<u>Archimedean Copulas</u>: - Not Derived Directly by applying Sklar's theorem to multivariate distributions. See Embrechts et. al (Chapter 8. 2003).

Let X and Y be continuous random variables, with joint bivariate distribution H and marginal distribution functions F and G, respectively.

Consider a strictly increasing continuous function $\lambda : [0,1] \rightarrow [0,1]$ such that $\lambda(0) = 0$ and $\lambda(1) = 1$, and suppose that (see Nelsen 1999):

$\lambda(\Pr\{X \leq x, Y \leq y\}) = \lambda(\Pr\{X \leq x\})\lambda(\Pr\{Y \leq y\})$; in particular for all $0 \leq x, y \leq 1$.

i.e. $\lambda(H(x, y)) = \lambda(F(x))\lambda(G(y))$

If we let $\phi(t)$ = - $\log\lambda(t)$ for $0 < t \leq 1$ (i.e. $\phi(t)$ is a convex decreasing function s.t. $\phi(1) = 0$), then the previous equation becomes:

$\phi\big(H(x, y)\big) = \phi(F(x)) + \phi(G(y))$; $H(x, y)$ is the joint distribution of (x, y) as mentioned before.

Arranging for copulas, the distribution becomes:

$$H_{\phi}(x, y) = \phi^{-1}[\phi(x) + \phi(y)].$$

and is generally referred to as an "Archimedian" Copula. (Note that $H_{\phi}(x, y) = C_{\phi}(x, y)$ - both referred to as Archimedian Copula, being ϕ a convex decreasing function).

Three typical Archimedian Copulas considered are:

i. Clayton family: where $\phi(t) = \left(t^{-\theta} - 1\right)/\theta$; for $\theta \epsilon [-1, \infty]\backslash\{0\}$, then

$H_{\phi}(x, y) = Max[(x^{-\theta} + y^{-\theta} - 1)^{-\frac{1}{\theta}}, 0]$; yet becomes:

$$H_{\phi}(x, y) = (x^{-\theta} + y^{-\theta} - 1)^{-\frac{1}{\theta}} \quad \text{for } \theta > 0$$

ii. Frank family: where $\phi(t) = -\ln\left((e^{-\theta t} -- 1)/(e^{-\theta} - 1)\right)$; for $\theta \epsilon \mathbb{R} \setminus \{0\}$, then:

$$H_{\phi}(x, y) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta x} - 1)(e^{-\theta y} - 1)}{e^{-\theta} - 1}\right)$$

Such that by Frechet-bounds:

$$\text{Lim}_{\theta \to -\infty} H_{\phi} = \max(x + y - 1, 0)\,; \quad \text{Lim}_{\theta \to 0} H_{\phi} = x * y; \quad \text{Lim}_{\theta \to \infty} H_{\phi} = \min(x, y)$$

iii. Gumbel family: where $\phi(t) = (-lnt)^{\theta}$; for $\theta \geq 1$, then

$$H_{\emptyset}(x, y) = \exp\left(-[(-lnx)^{\theta} + (-lny)^{\theta}]^{\frac{1}{\theta}}\right)$$

Lower tail dependence is captured by the Clayton family for $\theta > 0$ - which due to its limited parameter space, results in this copula only capturing positive correlation for such lower tail dependence. see Freez and Valdez (1998). Upper tail dependence is captured or reflected in the Gumbel family for $\theta > 1$, which once again due to its parameter space, holds only for positive codependence. Nonetheless, negative dependence may be obtained in both previous copulas by initially pre-multiplying either series by -1. That is the pair (-X, Y) or (X, -Y) may be modeled as a joint distribution. In addition, all three models include the special case for independent marginal distributions between x and y, which is at $\theta = 0$. i.e, for $\theta = 0$, all three Copula families become: see Genest & Rivest, (1993):

$$H_{\phi} = xy$$

The Frank family is the only type of Archimedian copula that holds for radial symmetry. see Embrechts et. al (Chpt 8, 2003). Yet it permits regularly both positive and negative correlations. Hence this Frank family copula will be used in our modeling, along with an elliptical (regular) normal copula for comparison.

The Kendall's Tau coefficient is used as a dependence, association or correlation measure between the marginal distributions in a copula. This is a rank coefficient that doesn't depend on the specification of the marginal distributions, but only on the copula used. The coefficient's population version is the probability of concordance (i.e. positive relation) minus the probability of discordance (i.e. inverse relation). See Nelsen, (1999)

$$\tau_{xy} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

For the Normal (Elliptic) Copula: see Freez and Valdez (1998)

$\tau_{xy} = \frac{2}{\pi} \arcsin \theta$ with $\theta \epsilon [-1, 1]$: the parameter of the Copula.

For the Frank (Archimedian) family: see Genest (1987):

Frank family: $\tau_{\alpha} = 1 + \frac{4}{\theta}\{D_1(\theta) - 1\}$;

with $\alpha = e^{-\theta}$ or $\theta = -log\alpha;$ and $D_1(\theta) = \frac{1}{\theta}\int_0^{\theta} \frac{t}{e^t - 1} dt$ $with$ $\theta \epsilon \mathbb{R} \setminus \{0\}$,

Data

Modeled prices consist of monthly averages of daily February future prices for corn and soybean - for delivery in December and November, respectively. This is data from the Chicago Board of Trade and the observations are from 1959 to 2007 obtained through CRB. Crop yields data are observations from a regular corn and soybean producing county in Iowa – Kossuth, obtained from the NASS of the USDA. These yields are calculated over the acres planted, and not harvested, so as to obtain a realistic view of the ex-ante conditions during planting. Only the corn crop had all the recorded years (1927 - 2007) for acres planted, so a proxy obtained from these planted acres for corn, in combination with the acres harvested for soybean - was used to estimate the missing records of planted acres for soybean (1927-1969). The yearly crop data, from 1927 through 2007, has been de-trended by following a regular procedure consisting in regressing the yields through two time regressors – one linear and one squared (better fit than plain linear), such that each observation is afterwards transformed relative to the predicted value of the latest data observation (2007). i.e. $\tilde{y}_t = \hat{y}_T * (1 + \frac{e_t}{\hat{y}_t})$; with t = 1927,……2007; T = 2007

The reason for the error term 'adjustment' is that yield changes (or trend deviations) occur at higher yields, as data tends to show - see Goodwin and Mahul (2004), Goodwin and Ker (1998), Gallagher (1987). The data for crop prices was de-trended in a similar form.

In the case of RA insurance, ex-ante crop prices are considered in the same manner (i.e. February future prices with delivery in December and November for corn and soybean, respectively); however only data from mid 1980's onward is used.

Summary measurements for the data are in Table 1.

Results

The Burr distribution provided a better overall fit for the crop prices when compared to the Normal distribution and the log-Normal distribution. By using a method of Maximum Likelihood estimation - see Watkins 1999 & Johnson 2003, the following was obtained and contrasted to Normal, and Log-Normal distributions:

Burr Distribution (Standard deviation in parenthesis):

Parameters:	Corn	Soybean
$\tau\ (shape1) =$	7.466	21.938
	(5.3637)	(4.5814)
$\alpha(shape2) =$	1.578	0.105
	(1.0827)	(0.01998)
$\phi\ (scale) =$	252.452	329.224
	(46.883)	(12.355)

Normal Distribution:

Parameters:	Corn	Soybean
$\mu\ (mean) =$	235.44	519.98
$\sigma^2(variance) =$	2,364.64	14,013.83

Log-Normal Distribution:

Parameters:	Corn	Soybean
$\mu\ (location) =$	1.69308	1.8285
$\sigma(scale) =$	0.0388	0.0364

The Vuong test is used to compare these non-nested models, by calculating the log of their likelihood ratio. The log ratio calculated is:

$$m_i = \log\left(\frac{f_1(y_i|\mathbf{x_i})}{f_2(y_i|\mathbf{x_i})}\right)$$

Then the statistic calculated for testing the non-nested hypothesis of Model 1 vs. Model 2 is:

$$v = \frac{\sqrt{n}\left[\frac{\sum_{i=1}^{n} m_i}{n}\right]}{\sqrt{\frac{\sum_{i=1}^{n}(m_i-\overline{m})^2}{n}}}$$

Where the following results are obtained: ($v > 2 => statistically\ significant$) :

$v - statistic$:	Corn	Soybean
$v - Burr\ vs. Normal$	0.3431	2.4815
$v - Burr\ vs. LogNormal$	62.3386	67.6529

The Burr distribution is significantly better than the Normal distribution for the case of soybeans, yet it is not conclusive for the case of corn; perhaps more data may help to improve this assessment (only 48 observations are considered). However, the Burr distribution is significantly better than the Log-Normal distribution in both cases of corn and soybean prices. This confirms what had been stated before regarding the Burr distribution, having parameters than can estimate higher moments of the data (specifically third and fourth moments), is able to better capture the skewness and kurtosis (fat tails) that the crop price data have.

Crop yields have been modeled via a Beta distribution, making use of previous literature mentioned that confirms its proper goodness of fit.

Estimated parameters obtained via maximum likelihood are:

Beta Distribution:

Parameters:	Corn	Soybean
$\alpha\ (shape1) =$	11.871	19.611
$\beta(shape2) =$	5.574	9.993
$\phi\ (scale) =$	252	72

As mentioned before, two different Copula methods were used to model the correlation between the marginal distributions of our crop yields and crop prices. These are an Elliptic Copula – the Normal, and an Archimedian Copula – The Frank family.

Estimations were made by two different maximum likelihood methods as per Yan, 2007 (see Appendix 1 for details). Following results were obtained (see Appendix 2 for various parameter calculation results):

Elliptic (Normal) Copula:

	Corn	Soybean
Kendall's Tau	**-0.06745**	**-0.08893**
Log-Likelihood	**-2282.175**	**-2780.231**
Theta ($\theta - Normal\ Copula$)	**-0.077167**	**-0.130798**

Archimedean – Frank Copula:

	Corn	Soybean
Kendall's Tau	**-0.08987**	**-0.17771**
Log-Likelihood	**-2279.137**	**-2791.327**
Theta ($\theta - Frank\ Copula$)	**-0.79886**	**-1.6466**

In the case of corn, the results for the co-dependence factor - Kendall’s tau, show only a small difference between the two copula methods used (-0.09 vs. -0.07), having the frank copula a bit higher inverse relation. In addition, the maximum likelihood values obtained for each method are quite similar, having a difference of about 0.1%.

For the case of soybean, the Kendall’s tau obtained is significantly more negative in the case of the Frank copula. This difference is almost a 100%, as it goes from -0.089 in the normal copula to -0.178 in the Frank copula. Also, the maximum likelihood is a bit better for the case of the Frank copula, having a small edge of about 0.3% (-2791.3 vs. -2780.2 for a normal copula).

The co-dependence (or correlation) factor or value, such as Kendall’s tau or Spearman’s rho, calculated for each copula can vary across these, as mentioned before. In other words, for the same marginal distributions, different values of Kendall’s tau may be obtained from different copulas as in our prior case, see Nelsen (1999). In this study, a larger inverse relation was obtained with the Frank copula. For graphs denoting this inverse relation between crop prices and yields, see Appendix 3, which includes three dimensional plots and contour graphs at different level curves for the copula parameter (theta or rho).

According to Kendall’s Tau of correlation coefficient, there is an inverse relation obtained between the two marginal distributions, as anticipated. In addition, by simply comparing log-likelihoods (or by AIC criteria: both methods have same number of parameters) - the Frank archimedian copula seems to characterize slightly better the relationship between prices and yields.

Discussion

The shortcomings and difficulties pertaining to the current crop revenue insurance methods listed previously (i.e. CRC and RA) have been noted extensively in the literature (see GAO 98-111, Goodwin and Smith 1995, Goodwin 2001). The perennial excessive payouts compared to the premiums paid is a point that we address here by analyzing a different method for calculating these premiums. In first term, the crop prices have been better characterized by using a Burr distribution instead of the normal or log-normal distribution regularly used in the present CRC and RA methods, respectively. Second, the level of relation between crop prices and yields has been gauged by applying copula methods to the marginal distributions of crop prices and yields, making use of the Burr distribution for prices and the Beta distribution for yields. The aim of the study is to provide some tools for practical analysis and use in the new methods of crop revenue insurance that will be soon offered

In order to assess a potential application of the studied copula methods to the case of crop revenue insurance, a simple analysis will be made by comparing both Normal (elliptical) and Frank (archimedian) type copulas assuming a situation in which a payout may be necessary. This may be the case when the crop price has fallen below a certain level, and/or the crop yield has fallen below a certain level, such that the combination of these cases – despite their inverse relation, results in revenue which is below the minimum insured.

This expected payout function may be represented as follows, assuming U: yield and V: price (see Embrechts et. al, and Goodwin):

$$E\{f(U, u_{\min}, V, v_{min})\} = p(U \leq u_{min}; V \leq v_{min}) * [\min(u * v) - E(U * V | U \leq u_{min}; V \leq v_{min})]$$

The expected payout can introduce the use of the copula method for the probability term on the right, enabling the application of the pair wise rank correlation provided by the marginal distributions of U and V. In other words, since it is difficult to estimate with accurateness a proper joint distribution for U and V, we replace that function by its copula, calculated previously. The second term on the right is just the difference between the minimum insured revenue level and the expected revenue level obtained, given its below that minimum revenue level.

Hence the probability for a payout can be estimated through:

$$p(U \leq u_{min}; V \leq v_{min}) = H(u, v) = C(F_1(u), F_2(v))$$

Being C the copula given for any $u, v \epsilon$ [0,1]:

$C(u, v) = H\left(F_1^{-1}(u), F_2^{-1}(v)\right)$ with $F^{-1}(t) = \inf\{x \,\epsilon\, \mathbb{R} \mid F(x) \geq t\}$ for all t in [0,1].

From our previous calculation of copulas we have:

i. Normal: $C\left(F_1(u), F_2(v)\right) = \Phi_\theta[\Phi^{-1}(\mathrm{F_1(u)}), \Phi^{-1}(\mathrm{F_2(v)})]$

ii. Frank: $C(F_1(u), F_2(v)) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta F_1 u}-1)(e^{-\theta F_2 v}-1)}{e^{-\theta}-1}\right)$

With $F_1(u)$: $c.d.f.\quad for\ Beta$ distribution of crop yields.

$F_2(v)$: $c.d.f.\quad for\ Burr\ XII$ distribution of crop prices.

By setting specific minimum yields (u_{min}) and prices (v_{min}), such that their product results in a minimum insured revenue level, we are able to measure the probabilities of being below that point through the use of our copula function. This may be done through simulation providing a better portrayal of instances when payouts may necessary.

Specifically we take both copulas previously calculated for each crop, and we separately simulate 100,000 observations to obtain estimated crop yields and prices that belong to the specific copula. Once these simulated prices and yields are obtained, we take their product as the revenue, and calculate particular minimum revenues for which the insurers would receive a payment if they are below it. Minimum revenues were estimated at 70%, 75% and 80% of the expected revenue, for each copula and crop. So there are three different minimum revenue scenarios for four different copula scenarios. These scenarios were compared to the case were prices and yields are assumed to be unrelated or independent. This latter is the case for the theta or rho parameter being equal to zero for both copulas.

	Frank Copula		Normal Copula		Independent
Expected Rev:	*0.238*		*0.243*		*0.250*
Corn	Min rev.	%Δ wrt Indpdt.	Min rev.	%Δ wrt Indpdt.	Min rev.
70%	0.167	-4.80%	0.170	-2.80%	0.175
75%	0.179	-4.80%	0.182	-2.80%	0.188
80%	0.190	-4.80%	0.194	-2.80%	0.200

	Frank Copula		Normal Copula		Independent
Expected Rev:	*0.228*		*0.239*		*0.250*
Soybean	Min rev.	%Δ wrt Indpdt.	Min rev.	%Δ wrt Indpdt.	Min rev.
70%	0.160	-8.80%	0.167	-4.40%	0.175
75%	0.171	-8.80%	0.179	-4.40%	0.188
80%	0.182	-8.80%	0.191	-4.40%	0.200

See Appendix 4 for plots depicting the minimum revenue level for each copula method. From the previous table, in both cases of Frank copula, there is a larger difference of revenue payment compared to the case of the price and yield variables not being related; specifically, 4.8% and 8.8% less minimum revenue for corn and soybean, respectively. Hence, the minimum revenue for both crops is less with the application of the copula method, than in the case of assuming crop yields and prices being independent or having a positive relation. This result is anticipated, responding to the previous inverse relation between prices and yields. At the same time, this means that there are fewer instances in which a payout may be necessary, because the minimum revenue would be below the case considering there is no relation between price and yield.

This is a very relevant finding when comparing to the present situation of CRC insurance, for example, where a payment may occur if there is no drop in yields, but there is a drop in prices. See GAO 98-111. The same occurs in the case of RA. Both these revenue insurance cases would pertain to the insured farmer facing no relation between their actual anticipated yields and a drop in price; hence obtaining an indemnity which would be higher than in the case of the revenue insurance copula method. The latter method considers the inverse relation between crop prices and yields, hence there is less chance of excessive payouts.

Conclusions

A critical issue regarding crop insurance coverage has been the excessive amounts of indemnity payouts compared to premiums charged, many times exceeding ratios of 2:1. This factor is compounded by the fact that the resources involved are in the hundreds of millions of dollars, with large amounts being subsidized by the government. In order to gauge a better alternative to the current methods available, which are in the process of being replaced, two aspects have been studied.

First a Burr distribution was used to characterize crop prices, and compared its goodness of fit versus the current normal and lognormal distributions currently being used in crop revenue insurance. Corn and Soybean future prices were used, in accordance with the practices of the current CRC and RA programs. This resulted in a better fit of the Burr distribution when compared to the log-normal, and an initial better fit when compared to the Normal, though more data may be need for this latter case, since for the corn crop there was not a significant difference in fit.

Second, two different copula families – Normal (elliptical) and Frank (archimedian) were used to measure the correlation between these crop prices and their yields. Crop yields were modeled with a Beta distribution, and the copula method made use of the price and yield distributions to provide a correlation level among them, using Kendall’s tau as means of correlation coefficient. The MLE method was used to calculate the copula method’s best fit. Results show that there is a negative correlation between the price and yield distributions, as anticipated, and they were corroborated by two different estimation MLE methods. An analysis of the implications of these results was made by calculating probabilities of indemnity payouts, and the extent of increased degree of certainty they provide in calculating the required premiums was presented.

Further Analysis

Various avenues may be pursued as topics of future research. One direct calculation could be made by using other copula distributional families, such as the t-copula, which has different dependence in the tails, and gauge their level of fitness and correlation. Another venue is to directly calculate premium rates based on the previous relations obtained through the copula method, with prices and yields. Another expansions may include calculating copula methods for other crops such as wheat, or others.

References

Agricultural Risk Protection Act of 2000 – Public Law 106-224; June 20, 2000

Embrecths, P. et. al, Modelling *Dependence with Copulas and Applications to Risk Management*. Handbook of heavy tailed distributions in Finance. Ed.by Rachev, S.T. 2003

Frees, E.W. and Valdez, E.A., "Understanding Relationships using Copulas", *North American Actuarial Journal*, Vol.2, Number 1, 1998.

GA/RCED – 98-111 Problems with New Crop Revenue Insurance Plans, 1998

Gallagher, P. "U.S. Soybean Yields; Estimation and Forecasting with Nonsymmetric Disturbances", *American Journal of Agricultural Economics*: 69, (November 1987): 796-803

Genest, C. and Rivest, L.P., "Statistical Inference Procedures for Bivariate Archimedian Copulas", *Journal of the American Statistical Association*, Vol. 88, September 1993, 1034-1043.

Genest, C. "Frank's family of bivariate distributions", *Biometrika* (1987), 74, 549-555

Genest, C. and MacKay, J. "The Joy of Copulas: Bivariate Distributions with Uniform Marginals", *The American Statistician*, November 1986, Vol. 40, 280-283

Goodwin, B.K., "Problems with Market Insurance in Agriculture"; *American Journal of Agricultural Economics: 83* (Aug. 2001): 643-649.

Goodwin, B.K, and Ker A. P., "Nonparametric Estimation of Crop Yield Distributions: Implications for rating Group-Risk Insurance Contracts", *American Journal of Agricultural Economics: 80,* (February 1998): 139 - 153

Goodwin, BK, and Ker, A.P, *Modeling Price and Yield Risk.* A comprehensive Assesment of the Role of Risk in U.S. Agriculture. Ed. R.E. Just and R.D. Pope, 2002

Goodwin, B.K., and Mahul, O., "Risk Modeling Concepts Relating to the Design and Rating of Agricultural Insurance Contracts", *World Bank Policy Research Working Paper 3392*, September 2004.

Goodwin, B.K., and Smith V.H., "The Economics of Crop Insurance and Disaster Aid.", AEI Studies in Agricultural Policy, *The AEI Press*, 1995

Goodwin, B.K., and Vados L.A., "Public Responses to Agricultural Disasters: Rethinking the Role of Government" *Canadian Journal of Agricultural Economics: 55* (2007) 399–417

Greene, W.H; Econometric Analysis, Fifth Edition, Prentice-Hall 2003.

Johnson, R., "Accelerated Life Testing and the Burr XII Distribution" PhD Thesis – University of Wales Swansea, 2003.

Joe, H., Xu, J., "The estimation method of inference functions for margins for multivariate models". Technical Report 166, Dept. of Statistics, University of British Columbia, 1996

Kotz S., and Johnson, N.L., "Encyclopedia of Statistical Sciences", Vol.1, 1982

Lindsay, S.R. et. al, "Modelling the diameter distribution of forest stands using the Burr distribution", *Journal of Applied Statistics*, Vol. 23, 1996, 609-619.

Nelsen, R.B., "An Introduction to Copulas", Lecture Notes in Statistics, Springer-Verlag vol. 139,1999

Voung, Q.H, "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses", *Econometrica,* Vol. 57, No. 2, (Mar., 1989), pp. 307-333

Watkins, A.J. "An algorithm for maximum likelihood estimation in the three parameter Burr XII Distribution", *Computational Statistics & Data Analysis* 32 (1999), 19-27

Yan, J., "Enjoy the Joy of Copulas" *Journal of Statistical Software* October 2007, vol. 21, Issue 4

Table 1.

	Prices[2] (n=47)				Yields[3] (n=81)			
	Corn		Soybean		Corn		Soybean	
	Regular	Detrend	Regular	Detrend	Regular	Detrend	Regular	Detrend
Mean	221.68	235.44	510.7	519.98	89.2	171.76	29.1	47.71
Stand. Dev	71.3	48.63	178.03	118.38	45.32	28.24	11.45	6.14
Max	376.57	353.04	826.43	883.54	185.68	229.16	54.41	64.87
Min	110.19	147.65	207.62	312.96	19.93	80.23	11.12	24.08
Predicted T	*234.08*		*512.17*		*171.29*		*47.65*	

[2] Price unit in Cents per bushel (bu). Minimum contract is for 5,000 bushels.
[3] Bushels per Planted acre.

Appendix #1: Copula estimation via two different MLE methods (Using R-code), both arrive at very similar results: see Yuan, J. 2007

One Step method:

Take n independent realizations from a multivariate distribution, $\{(X_{i1}, \ldots \ldots \ldots, X_{ip})^T : i = 1, \ldots \ldots, n\}$. Our de-trended yield and price data may still be considered 'sequentially' correlated, hence not formally independent; yet estimation through this method may be considered as a second-best approach with respect to a different method that doesn't make use of this.

Assume the multivariate distribution may be specified by p marginal cumulative distributions cdf F_i & density distributions pdf $f_i, i = 1 \ldots . p$; and a copula with density $\boldsymbol{c}$.

Consider β the vector of marginal parameters and α the vector of copula parameters. The parameter vector to be estimated is $\theta = (\beta^T, \alpha^T)^T$. The loglikelihood function is:

$$l(\theta) = \sum_{i=1}^{n} \log \boldsymbol{c}\{F_1(X_{i1}; \beta), \ldots \ldots \ldots, F_p(X_{ip}; \beta); \alpha\} + \sum_{i=1}^{n} \sum_{j=1}^{p} \log f_i (X_{ij}; \beta)$$

Being the ML estimator of θ:

$$\hat{\theta}_{ML} = \arg \max_{\theta \epsilon \Theta} l(\theta) \quad \text{where } \Theta \text{ is the parameter space.}$$

Two Step method:

Considering a substantial increase in the dimensions (p) of the Multivariate distribution, the previous method may be more difficult. Hence a two step optimization method may be more expeditious and reach similar results, proposed by Joe and Xu, 1996.

This method, called inference functions for margins (IFM) estimates the marginal parameters β,

in a first step:

$$\hat{\beta}_{IFM} = \arg \max_{\beta} \sum_{i=1}^{n} \sum_{j=1}^{p} \log f_i (X_{ij};\ \beta)$$

And then estimates the parameters of association α given by $\hat{\beta}_{IFM}$, by:

$$\hat{\alpha}_{IFM} = \arg \max_{\alpha} \sum_{i=1}^{n} \log \boldsymbol{c} \left(F_1(X_{i1};\ \hat{\beta}_{IFM}), \ldots \ldots \ldots \ldots F_p(X_{ip};\ \hat{\beta}_{IFM});\ \alpha\right)$$

When each marginal distribution F_i has its own parameter set β_i , such that $\beta = (\beta_1^T, \ldots \ldots \ldots \beta_p^T)^T$, *then* the first step involves a MLE for each margin $j = 1, \ldots \ldots \ldots, p$:

$$\hat{\beta}_{jIFM} = \arg \max_{\beta j} \sum_{i=1}^{n} log f\, (X_{ij};\ \beta_j)$$

One Step Results:

Normal Elliptical

Both ML estimations are based on 500 observations taken from the copula with known marginals Beta and Burr XII – according to their respective parameter results.

Corn Crop: | Soybean Crop:

Margin 1 : Beta – Crop yields

Corn Crop	*Estimate*	*Std. Error*	Soybean Crop	*Estimate*	*Std. Error*
m1.shape1	**9.588089**	**3.6169760**	**m1.shape1**	**16.832515**	**8.712091**
m1.shape2	**5.543283**	**0.8138692**	**m1.shape2**	**10.841455**	**2.212099**
m1.ncp	**3.823379**	**10.7764578**	**m1.ncp**	**7.127001**	**25.965249**

Margin 2 : Burr XII – Crop Prices

Corn Crop	*Estimate*	*Std. Error*	Soybean Crop	*Estimate*	*Std. Error*
m2.shape1	**6.993879**	**0.4441981**	**m2.shape1**	**24.84527114**	**3.64096331**
m2.shape2	**1.750277**	**0.3779837**	**m2.shape2**	**0.08770315**	**0.01458844**
m2.scale	**261.663379**	**12.0750284**	**m2.scale**	**329.72494610**	**3.91907059**

Copula: | ***Copula***:

Corn Crop	*Estimate*	*Std. Error*	Soybean Crop	*Estimate*	*Std. Error*
rho.1	**-0.07716675**	**0.04450274**	**rho.1**	**-0.1307984**	**0.04439401**

***The maximized loglikelihood is* -2282.175** | ***The maximized loglikelihood is* -2780.231**

The convergence code is 0

***a.00 - Kendall's Correlation Coefficient*:** | ***a.00 - Kendall's Correlation Coefficient*:**

-0.0674469 | ***-0.08892986***

***One Step Results*:**

Frank Archimedian

Both ML estimations are based on 500 observations taken from the copula with known marginals Beta and Burr XII – according to their respective parameter results.

Corn Crop:

Margin 1 : Beta – Crop yields

	Estimate	*Std. Error*
m1.shape1	**10.329336**	**3.6197203**
m1.shape2	**5.129206**	**0.7478332**
m1.ncp	**1.979075**	**10.7298615**

Margin 2 : Burr XII – Crop Prices

	Estimate	*Std. Error*
m2.shape1	**6.954176**	**0.4161142**
m2.shape2	**1.700168**	**0.3212506**
m2.scale	**257.574265**	**10.4289894**

***Copula*:**

	Estimate	*Std. Error*
param	**-0.7988675**	**0.2768432**

The maximized loglikelihood is -2279.137

The convergence code is 0

***a.00 - Kendall's Correlation Coefficient*:**

-0.08987575

Soybean Crop:

	Estimate	*Std. Error*
m1.shape1	**17.552994**	**7.159249**
m1.shape2	**9.408260**	**1.795189**
m1.ncp	**2.152311**	**21.601034**

	Estimate	*Std. Error*
m2.shape1	**19.6338025**	**2.54723236**
m2.shape2	**0.1122604**	**0.01706057**
m2.scale	**328.1138322**	**4.47971444**

***Copula*:**

	Estimate	*Std. Error*
param	**-1.646662**	**0.2770623**

The maximized loglikelihood is -2791.327

***a.00 - Kendall's Correlation Coefficient*:**

-0.1777154

Appendix 2.-

Results Burr Beta Corn - Frank						**Normal**					
Rho	Rho-Hat	Std.Error	Max likelhd	K Tau		Rho	Rho-Hat	Std.Error	Max likelhd	K Tau	
-5	-4.89158	0.34074	**-2137.356**	**-0.44943**		-0.3	-0.3064	0.04043	**-2229.142**	**-0.2041**	
	-5.87449	0.36782	**-2093.084**	**-0.51024**			-0.2867	0.04118	**-2178.069**	**-0.1908**	
	-4.48190	0.32708	**-2160.083**	**-0.42964**			-0.2696	0.04161	**-2242.094**	**-0.1569**	*
	-4.96769	0.33983	**-2124.339**	**-0.45925**			-0.3160	0.04055	**-2187.173**	**-0.1950**	
-3	-2.84543	0.29645	**-2151.196**	**-0.29345**		-0.2	-0.2357	0.04211	**-2249.808**	**-0.1509**	*
	-3.32521	0.30373	**-2171.87**	**-0.33408**			-0.2528	0.04207	**-2216.702**	**-0.1630**	
	-2.69791	0.29641	**-2227.717**	**-0.27743**			-0.2329	0.04235	**-2237.304**	**-0.1378**	
	-3.19362	0.31074	**-2201.612**	**-0.31421**			-0.2304	0.04216	**-2239.865**	**-0.1609**	*
-2	-2.16985	0.28224	**-2270.434**	**-0.23137**	*	-0.15	-0.1108	0.04417	**-2210.227**	**-0.0767**	
	-2.19223	0.28364	**-2193.242**	**-0.23322**			-0.1162	0.04418	**-2216.727**	**-0.0755**	*
	-2.21766	0.28713	**-2242.464**	**-0.23163**	*		-0.0840	0.04471	**-2252.544**	**-0.0495**	*
	-2.03552	0.28661	**-2201.187**	**-0.21605**			-0.1440	0.04378	**-2215.346**	**-0.1038**	
-1	-0.98479	0.27073	**-2246.901**	**-0.10825**		-0.1	-0.1135	0.04428	**-2257.053**	**-0.0705**	
	-1.03786	0.26653	**-2231.647**	**-0.11671**			-0.1557	0.04354	**-2257.641**	**-0.0809**	
	-1.27044	0.27703	**-2264.5**	**-0.13587**	*		-0.0986	0.04440	**-2257.597**	**-0.0669**	*
	-0.91790	0.26955	**-2262.579**	**-0.09945**	*		-0.0663	0.04505	**-2276.029**	**-0.0434**	*
-0.75	**-0.79887**	0.27684	**-2279.137**	**-0.08988**	*	-0.05	-0.0040	0.04470	**-2229.841**	**0.0001**	
	-0.74716	0.27798	**-2259.575**	**-0.07779**			-0.1484	0.04406	**-2248.057**	**-0.0994**	
	-1.25368	0.28027	**-2258.553**	**-0.13303**			-0.1003	0.04456	**-2236.901**	**-0.0817**	
	-1.09866	0.27058	**-2270.498**	**-0.12184**	*		**-0.0772**	0.04450	**-2282.175**	**-0.0674**	*
-0.5	-0.46999	0.269527	**-2271.459**	**-0.05419**	*						
	-0.66675	0.266219	**-2276.008**	**-0.07546**	*						
	-0.54270	0.271863	**-2248.582**	**-0.05784**							
	-0.19054	0.264969	**-2251.489**	**-0.01889**							

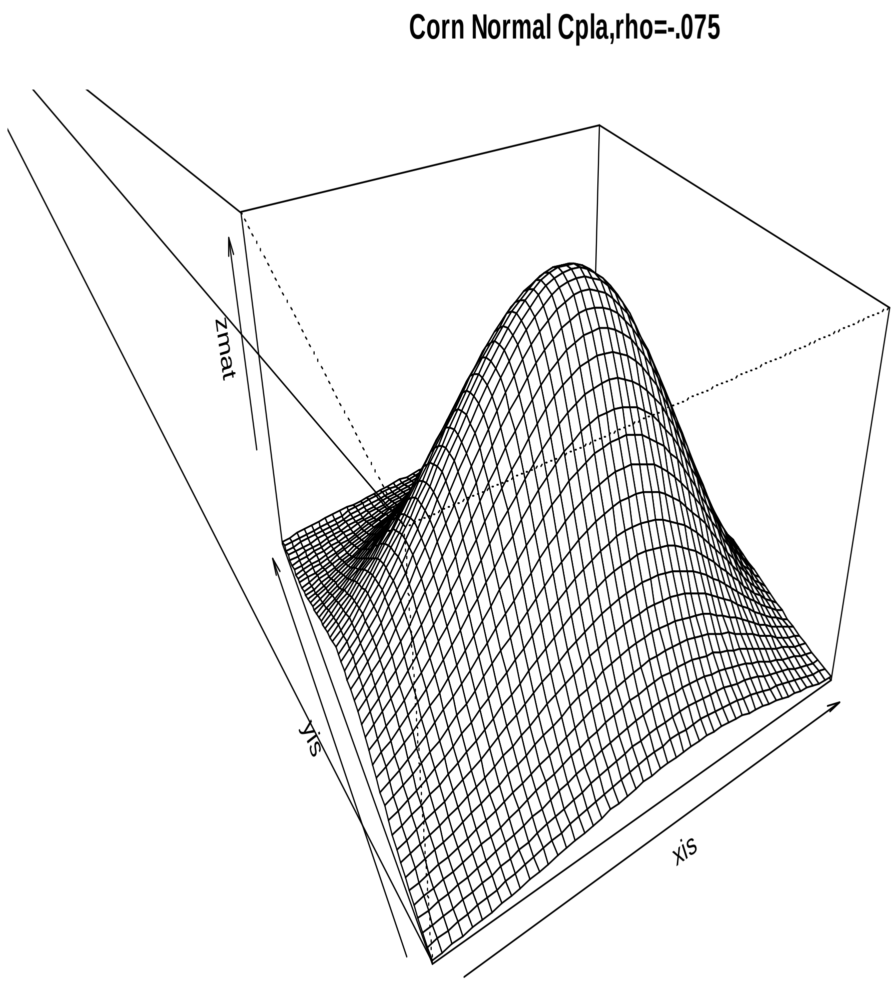
Results Burr Beta Soybean - Frank

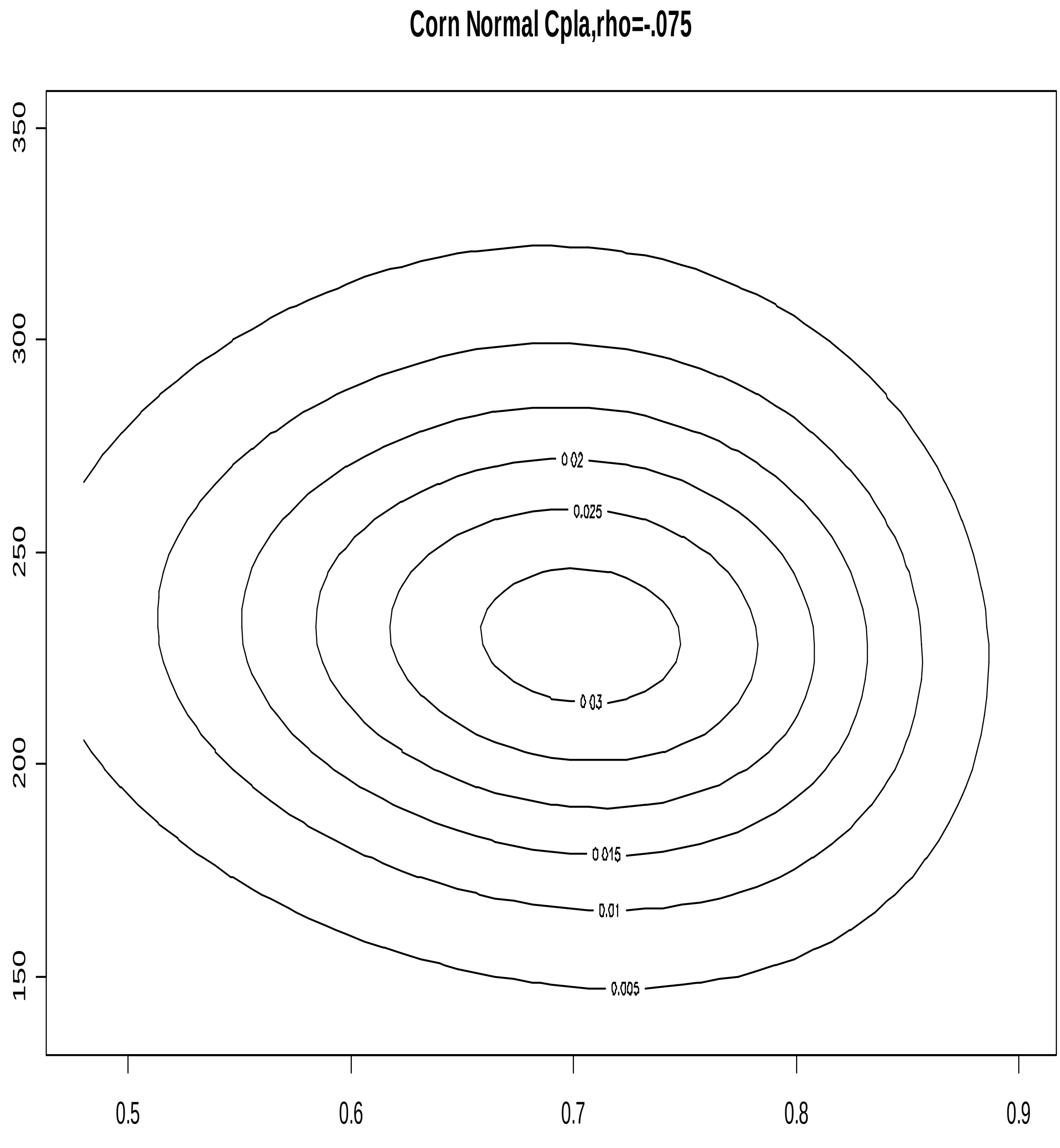
Rho	Rho-Hat	Std.Error	Max likelhd	K Tau	
-8	-8.43286	0.44770	**-2479.117**	**-0.62018**	
	-7.81429	0.42624	**-2482.854**	**-0.59356**	
	-7.71796	0.42230	**-2543.013**	**-0.59660**	
	-7.81605	0.43352	**-2438.745**	**-0.58871**	
-5	-5.11138	0.34856	**-2615.552**	**-0.46034**	
	-5.11138	0.34856	**-2615.552**	**-0.39527**	
	-4.07331	0.32226	**-2641.919**	**-0.39296**	
	-4.76770	0.34487	**-2573.277**	**-0.42703**	
-2	-1.53019	0.27593	**-2711.205**	**-0.16752**	
	-2.15697	0.28922	**-2757.247**	**-0.22455**	
	-2.20348	0.28575	**-2712.739**	**-0.23117**	
	-1.90631	0.27903	**-2726.567**	**-0.20452**	
-1.5	-1.15187	0.26585	**-2778.078**	**-0.13063**	
	-1.46727	0.27813	**-2761.454**	**-0.15700**	*
	-1.66505	0.27921	**-2753.13**	**-0.17941**	
	-1.64666	0.27706	**-2791.327**	**-0.17772**	**
-1	-1.10853	0.26680	**-2704.732**	**-0.12473**	
	-0.85754	0.26818	**-2762.636**	**-0.09544**	*
	-1.13774	0.27620	**-2766.48**	**-0.12273**	*
	-0.88314	0.26885	**-2724.841**	**-0.10020**	

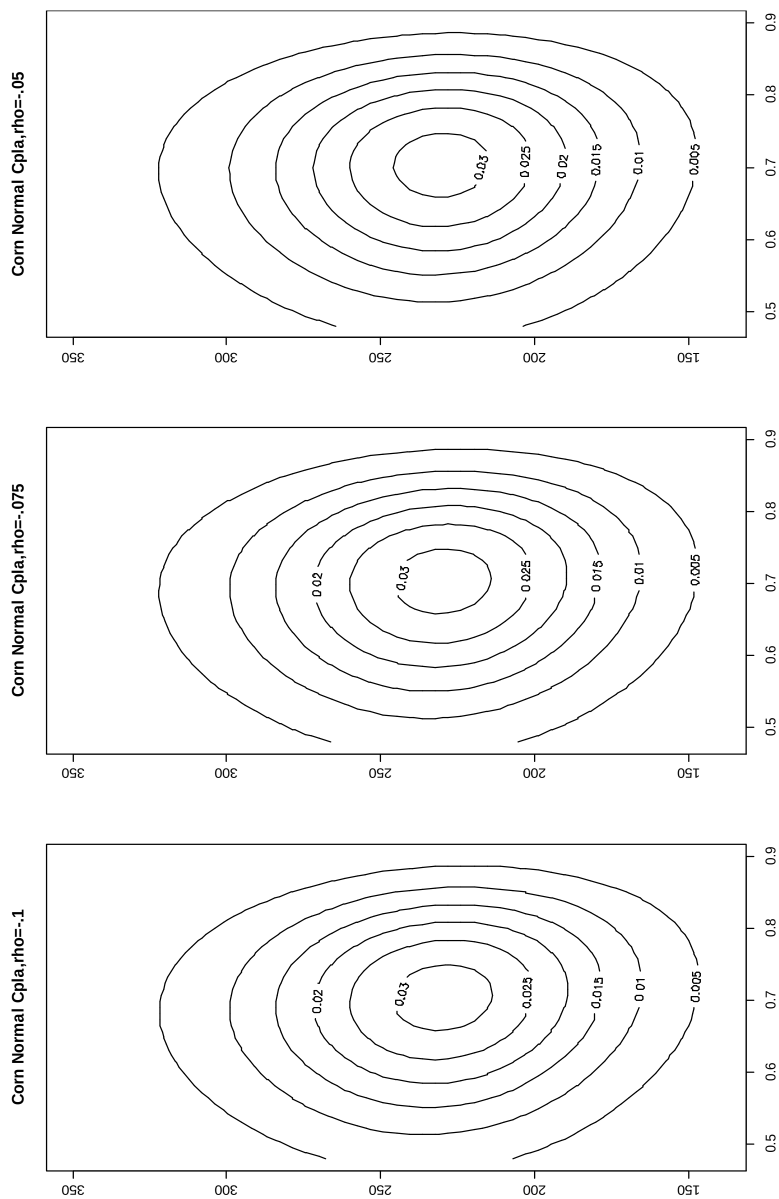
Normal

Rho	Rho-Hat	Std.Error	Max likelhd	K Tau	
-0.5	-0.4507	0.03576	**-2619.598**	**-0.2832**	
	-0.4456	0.03609	**-2713.94**	**-0.2851**	
	-0.5222	0.03249	**-2726.879**	**-0.3507**	
	-0.5595	0.03031	**-2652.884**	**-0.3781**	
-0.3	-0.3395	0.03957	**-2724.852**	**-0.2167**	
	-0.3204	0.04028	**-2692.837**	**-0.1951**	
	-0.3089	0.04058	**-2761.229**	**-0.1868**	*
	-0.3347	0.03950	**-2753.209**	**-0.2030**	*
-0.25	-0.2227	0.04243	**-2741.291**	**-0.1381**	*
	-0.2553	0.04159	**-2710.894**	**-0.1680**	
	-0.2685	0.04177	**-2719.647**	**-0.1718**	
	-0.1962	0.04271	**-2772.676**	**-0.1253**	*
-0.2	-0.2424	0.04198	**-2713.825**	**-0.1475**	
	-0.2971	0.04101	**-2713.679**	**-0.1930**	*
	-0.2116	0.04322	**-2750.007**	**-0.1368**	*
	-0.2778	0.04158	**-2700.298**	**-0.1652**	
-0.15	-0.1931	0.04359	**-2759.248**	**-0.1225**	*
	-0.1031	0.04459	**-2741.615**	**-0.0621**	*
	-0.1341	0.04417	**-2759.977**	**-0.0861**	*
	-0.1720	0.04317	**-2715.891**	**-0.1226**	
-0.1	-0.06167	0.043938	**-2754.97**	-0.0234	*
	-0.1115	0.043238	**-2754.517**	-0.0753	*
	-0.16992	0.043779	**-2717.68**	-0.0994	
	-0.11969	0.044255	**-2710.982**	-0.0657	
-0.05	-0.14546	0.043579	**-2775.621**	**-0.0833**	*
	-0.18739	0.043671	**-2731.419**	**-0.1237**	
	-0.1308	0.044394	**-2780.231**	**-0.0889**	*
	-0.18285	0.042991	**-2717.087**	**-0.1093**	

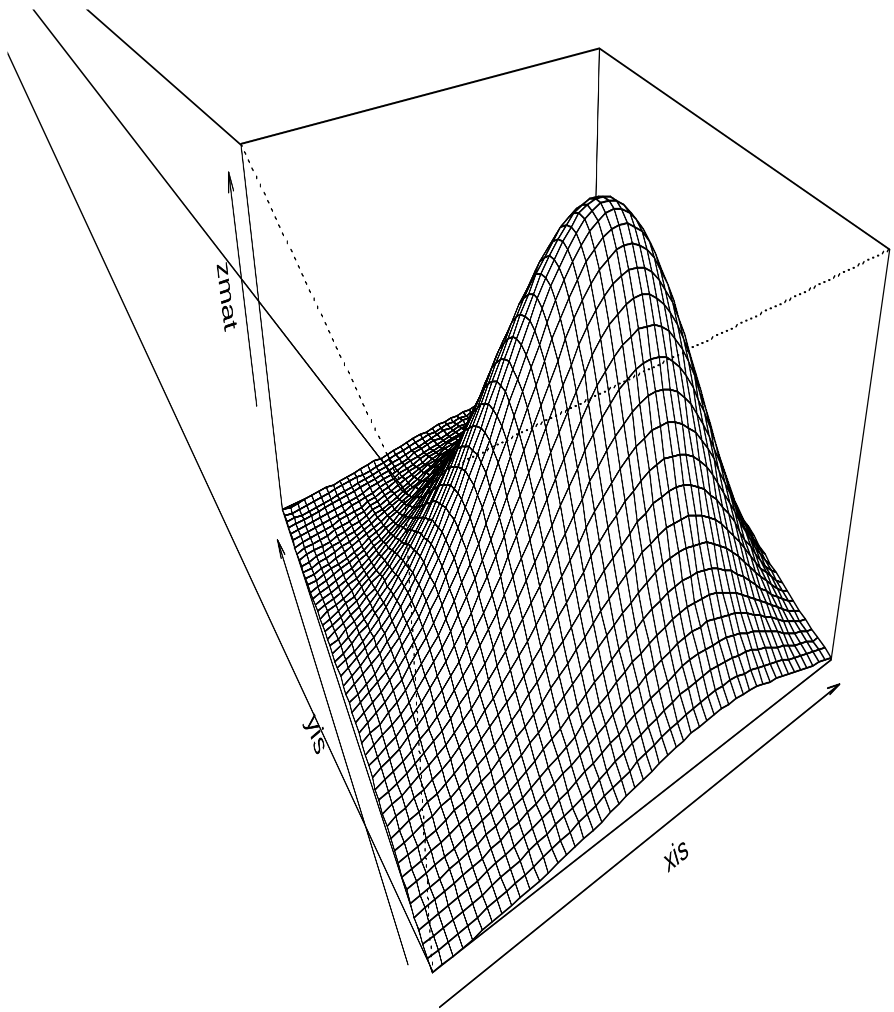
Appendix 3.

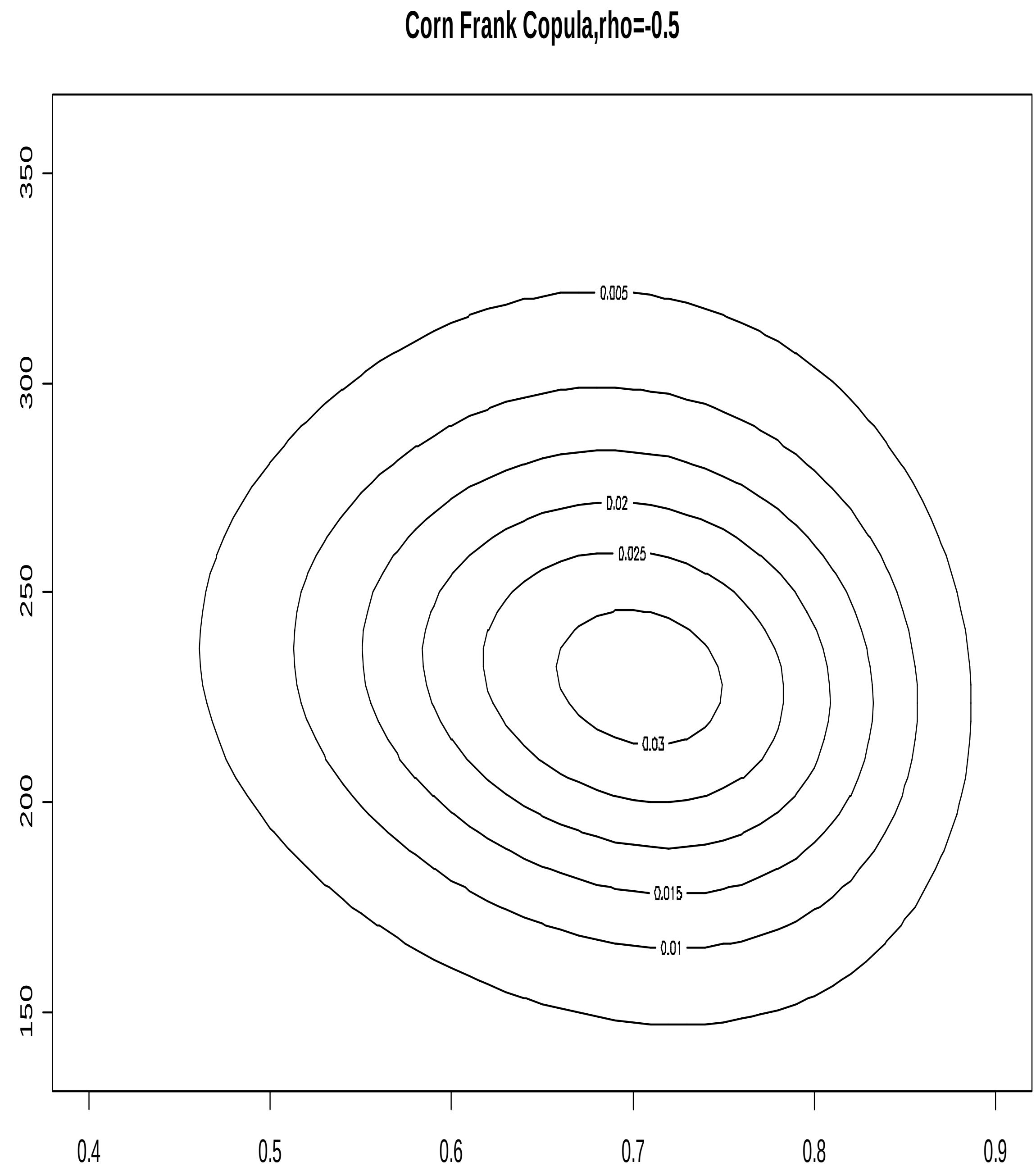



Corn Normal Cpla,rho=-.075
350
300
250
200
150
0.5
0.6
0.7
0.8
0.9
0.02
0.025
0.03
0.015
0.01
0.005

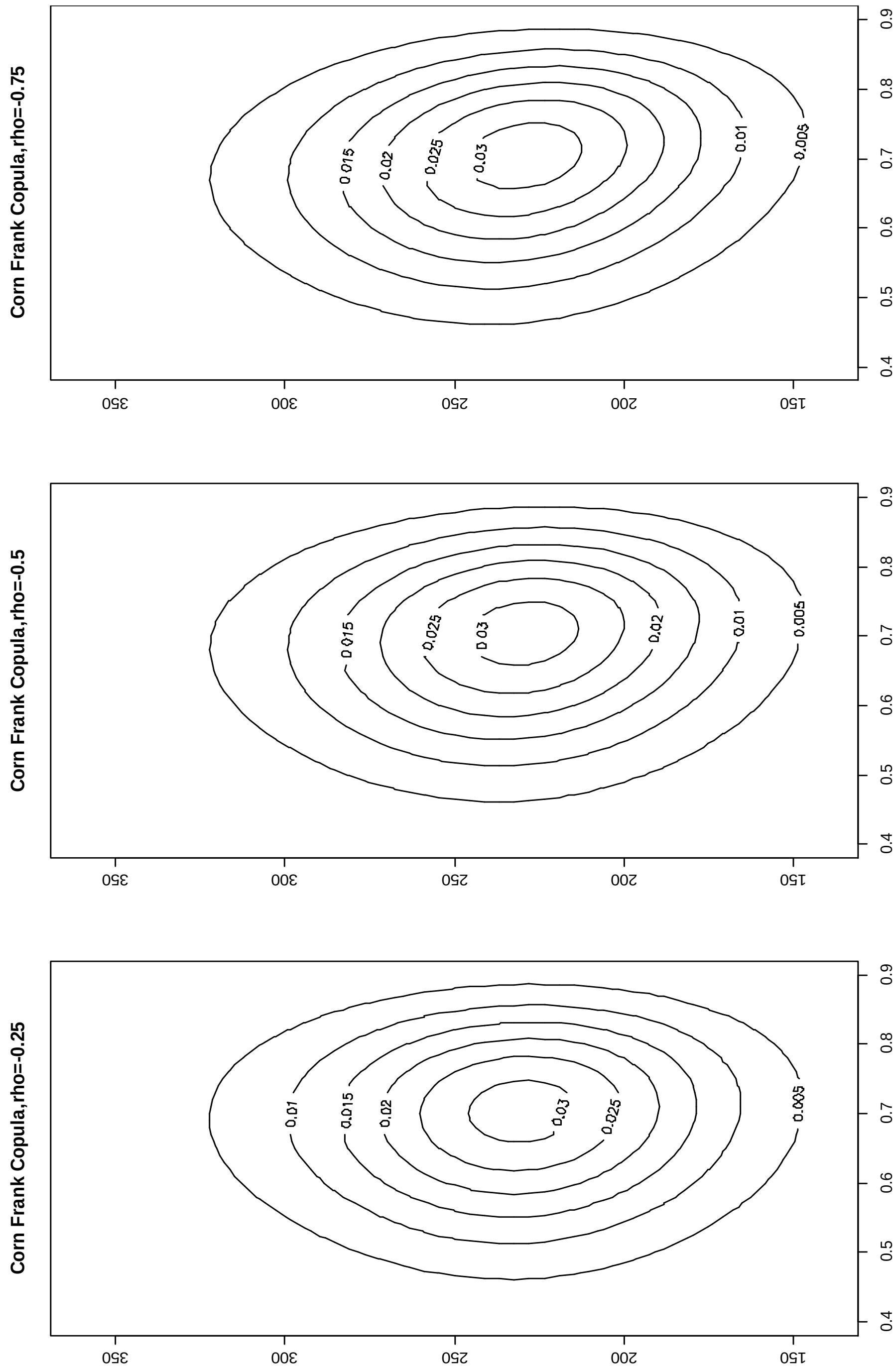

Corn Normal Cpla,rho=-.05
Corn Normal Cpla,rho=-.075
Corn Normal Cpla,rho=-.1

Corn Frank Copula,rho=-0.5

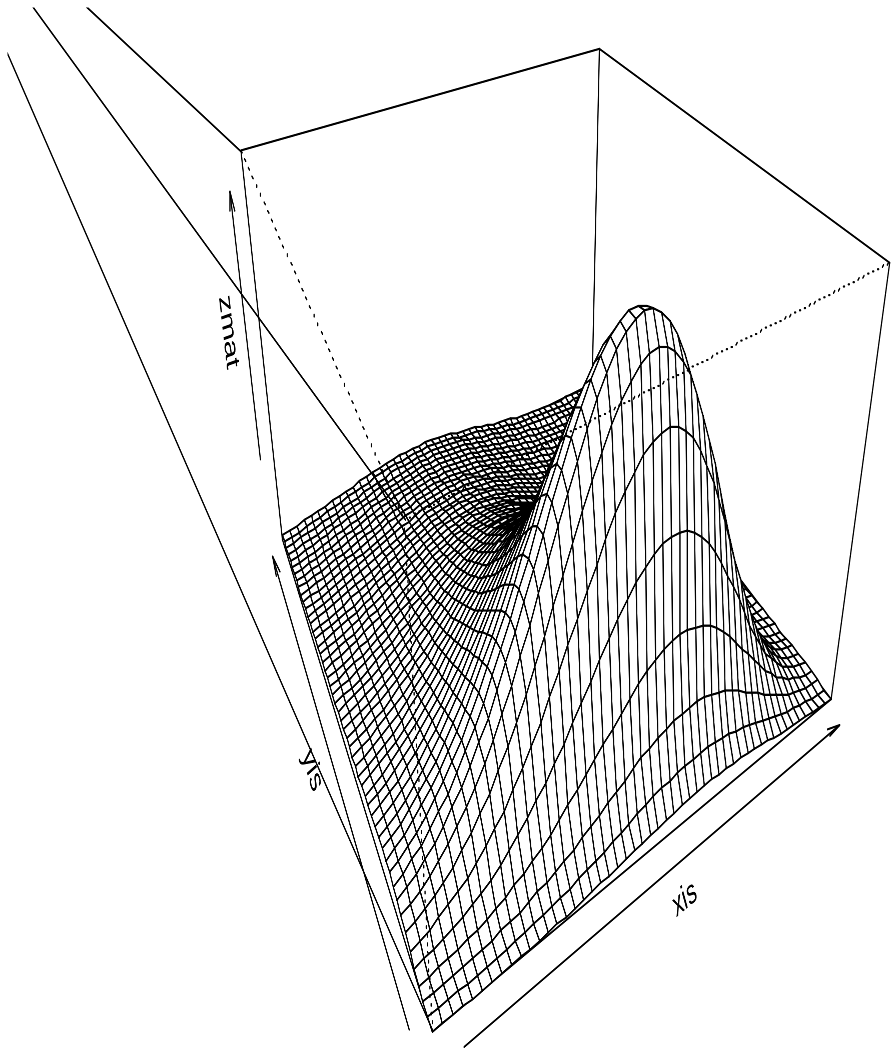


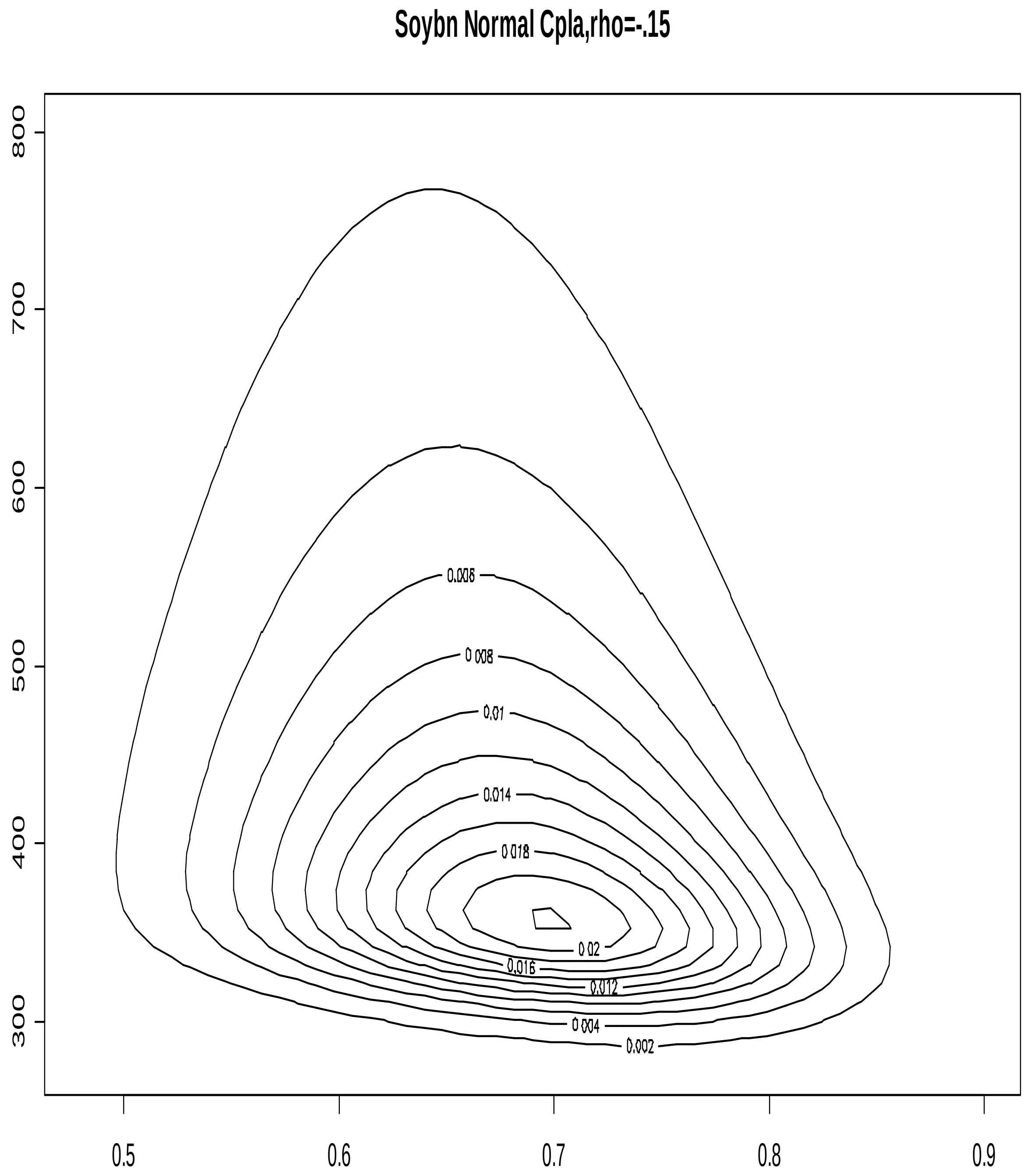


Corn Frank Copula,rho=-0.5
0.005
0.01
0.015
0.02
0.025
0.03
350
300
250
200
150
0.4
0.5
0.6
0.7
0.8
0.9

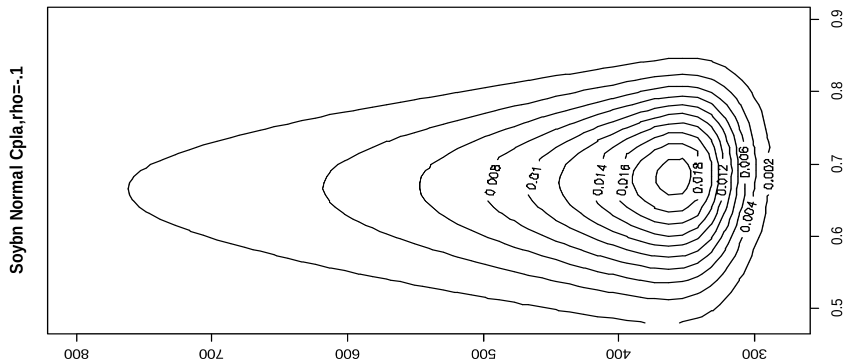

Corn Frank Copula,rho=-0.75
Corn Frank Copula,rho=-0.5
Corn Frank Copula,rho=-0.25

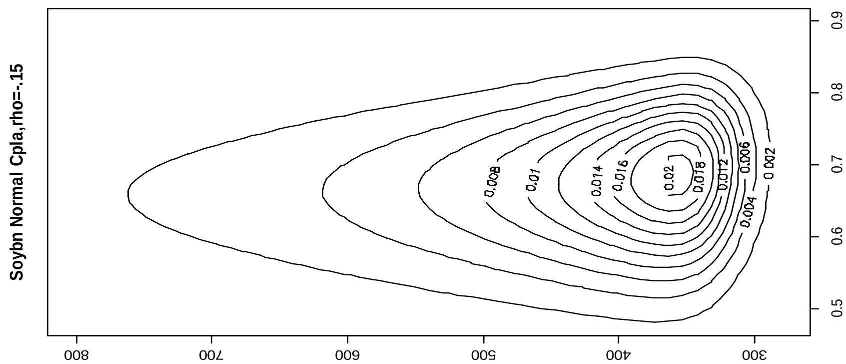
Soybn Normal Cpla,rho=-.15

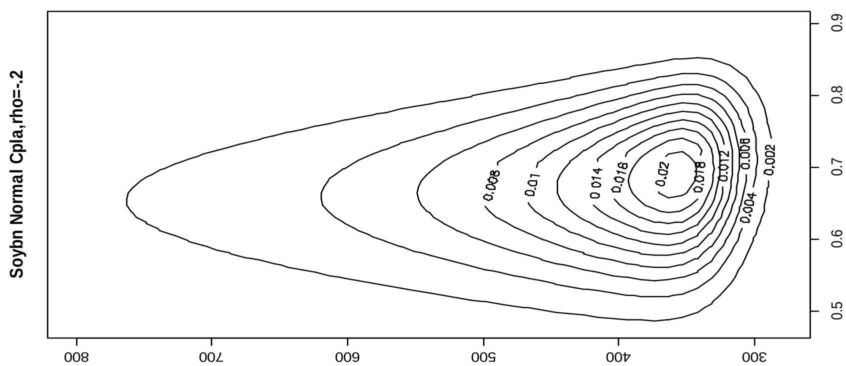




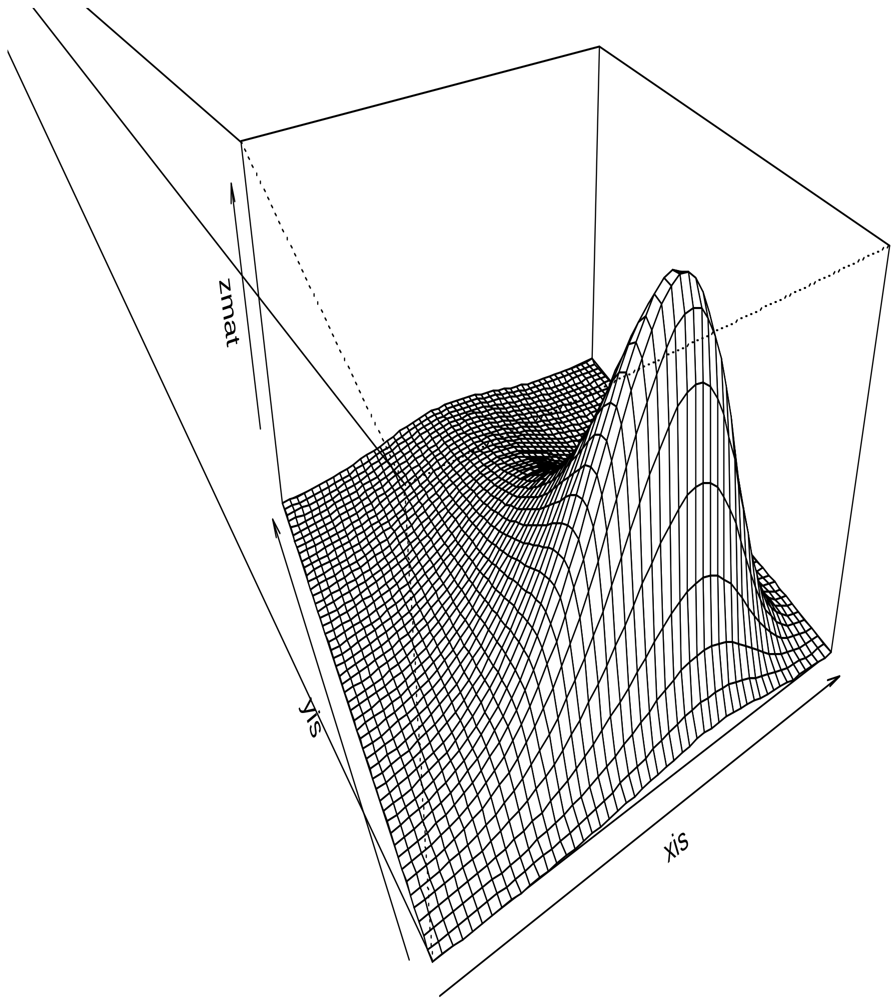
Soybn Normal Cpla,rho=-.15
800
700
600
500
400
300
0.5
0.6
0.7
0.8
0.9
0.006
0.008
0.01
0.014
0.018
0.02
0.016
0.012
0.004
0.002

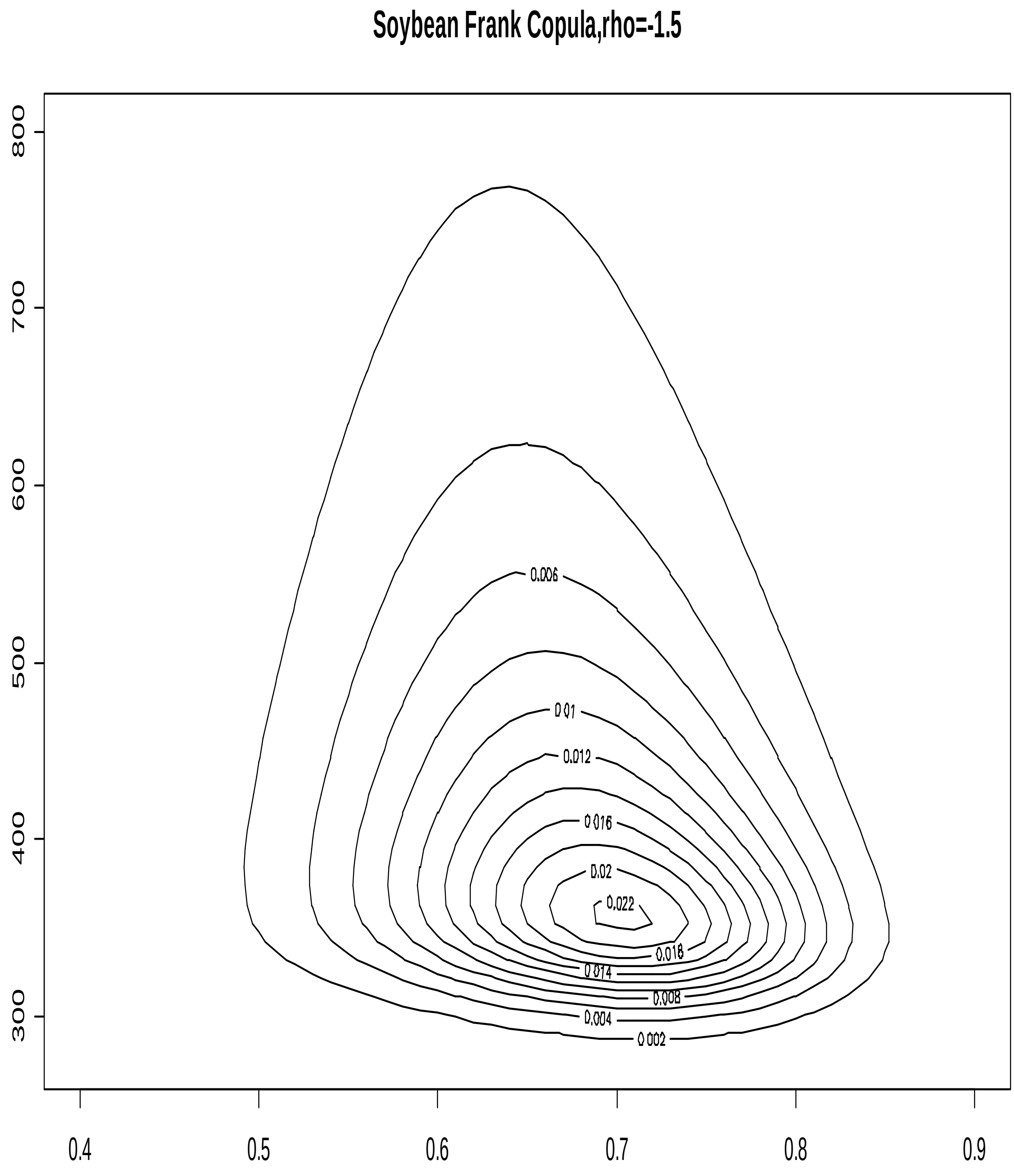

Soybn Normal Cpla,rho=-.1

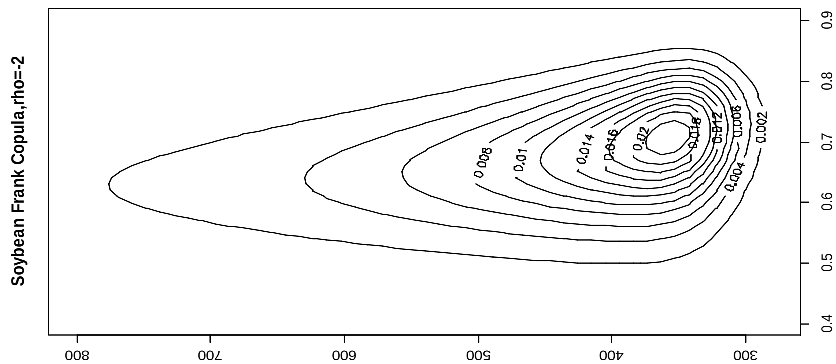

Soybn Normal Cpla,rho=-.15

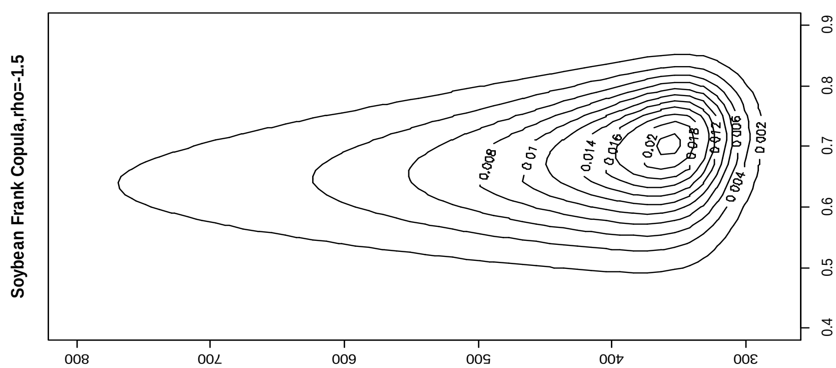

Soybn Normal Cpla,rho=-.2

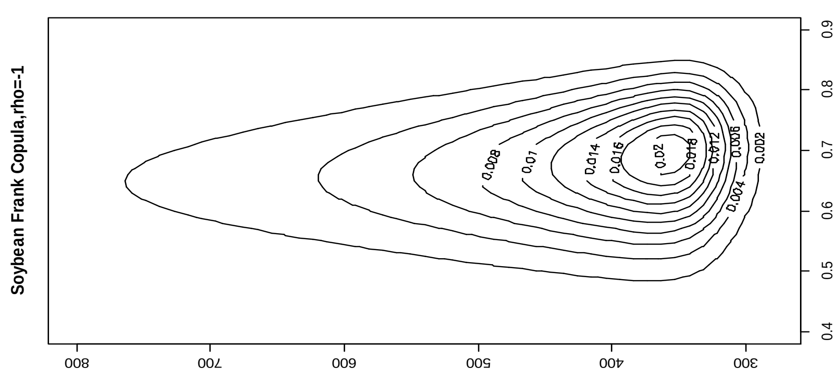
Soybean Frank Copula,rho=-1.5




Soybean Frank Copula,rho=-1.5
800
700
600
500
400
300
0.4
0.5
0.6
0.7
0.8
0.9
0.006
0.01
0.012
0.016
0.02
0.022
0.018
0.014
0.008
0.004
0.002


Soybean Frank Copula,rho=-2


Soybean Frank Copula,rho=-1.5


Soybean Frank Copula,rho=-1

Appendix 4.