Copula-Based Models of Systemic Risk in U.S. Agriculture: Implications for Crop Insurance and Reinsurance Contracts *

Barry K. Goodwin

DRAFT

May 10, 2012

Abstract

The federal crop insurance program has been a major fixture of US agricultural policy since the 1930s. The program continues to grow in size and prominence and now represents the most prominent farm policy instrument, accounting for more government spending than any other farm commodity program. In 2011, over $114 billion in crop value was insured under the program. Crop revenue insurance, first introduced in the 1990s, now accounts for nearly 70% of the total liability in the program. The plans cover losses that result from a revenue shortfall that can be triggered by multiple, dependent sources of risk—either low prices, low yields, or a combination of both. The actuarial practices currently applied in rating these plans essentially involve the application of a Gaussian copula model to the pricing of dependent risks. We evaluate the suitability of this assumption by considering a number of alternative copula models, including a relatively new innovation in copula modeling—the vine copula. This approach uses combinations of pair–wise copulas of conditional distributions to model multiple sources of risk. We find that this approach is generally preferred by model fitting criteria in the applications considered here. We demonstrate that alternative approaches to modeling dependencies in a portfolio of risks may have significant implications for the pricing of such risks.

*William Neal Reynolds Distinguished Professor in the Departments of Economics and Agricultural and Resource Economics at North Carolina State University. Paper prepared for the NBER Conference on Insurance Markets and Catastrophe Risk. Research support from the US Forest Service and the North Carolina Agricultural Research Service is gratefully acknowledged. Direct correspondence to Box 8109, North Carolina State University, Raleigh, NC, 27695, E-mail: barry_goodwin@ncsu.edu.

Copula-Based Models of Systemic Risk in U.S. Agriculture: Implications for Crop Insurance and Reinsurance Contracts

1 Introduction

Agriculture is subject to a wide variety of risks, including many hazards arising from widespread natural disasters. The U.S. federal crop insurance program, which was initially introduced on a small scale in 1938, now carries a total liability in excess of $114 billion and insures 262 million acres. The premiums paid by farmers in this program are highly subsidized (in excess of 60% of the total premium) and private insurance companies also receive significant taxpayer subsidies to operate and administer the program. Private insurance companies are also provided with an advantageous taxpayer–supported reinsurance agreement. In recent years, the program has accounted for nearly $10 billion annually in subsidies to farmers and insurance companies, making it the most expensive agricultural commodity program. The program is currently being debated in Congress as the new 2012 Farm Bill is considered. If anything, indications are that the next farm bill may expand federal crop insurance programs by introducing a "shallow loss" program intended to offer higher coverage levels. Whether such a program is implemented through the federal crop insurance program or as a component of other farm commodity programs remains to be seen. However, all Congressional observers agree that crop insurance will continue to play a key role in US farm policy.

A central question underlying the mammoth federal crop insurance program involves the need and rationale for such immense taxpayer support. Observers question whether the government is providing welfare–enhancing support because the market has failed to provide it or, conversely, whether private insurance markets have been crowded–out by such huge government subsidies. Advocates for government intervention frequently point to the substantial systemic risk that characterizes agriculture. In particular, the argument maintains that the $114 billion in liability is too large for private insurers and reinsurers to adequately cover due to the potential systemic risk associated with natural disasters such as drought and floods (see, for example, Miranda and Glauber (1997)).

Quantifying the degree of systemic risk is central to addressing public policy issues involving the necessity of large subsidies for agricultural insurance. Of particular concern to the debate is the role of "state–dependent" risk. Empirical evidence has demonstrated that the spatial correlation of crop yields tends to be significantly stronger during extreme events such as droughts than is true in a typical year (see, for example, Goodwin (2001)). Standard models of systemic risk and insurance portfolio diversification for crop insurance nearly always assume that risks are linearly correlated and constant. In reality, the extent to which these risks may change across various states of nature has important implications for the pricing and availability of reinsurance.

This paper applies a variety of copula models to evaluate the extent to which weather and natural disaster risks in agriculture tend to be systemic and state–dependent. Copula models, though a relatively recent analytical innovation, have realized widespread use in evaluating multivariate sources of risk in the design and pricing of insurance contracts. We apply these models to two important data sources that measure agricultural risks—the USDA's extensive history on county–level crop yields and futures market assessments of expected prices. We consider the pricing of revenue insurance contracts. Revenue coverage currently accounts for about 70% of the total liability of the federal crop insurance program. Our results demonstrate that the approach adopted for measuring multiple, correlated sources of risk may have very substantial implications for the accurate measurement of portfolio risks. The standard assumption—a Gaussian copula model—is shown to significantly underprice risk. This may reflect the fact that this model does not allow for tail–dependence—a critical factor when risks are state–dependent.

The next section of the paper provides a brief overview of the US federal crop insurance program. We then discuss analytical models of risks for portfolios of insurance contracts comprised of multiple sources of risk. In particular, we consider high–ordered copula models of the joint distributions of crop yield and price risks. We present an application of these models to the case of corn and soybean revenue insurance for four prominent Illinois counties. The paper concludes with a discussion of the implications of the models for systemic risk in the U.S. crop insurance program. We quantify this risk and provide implications for the federal reinsurance pool as well as private reinsurance, which also plays an important role in the U.S. crop insurance program. Finally, we discuss the implications of our results for the viability of private crop insurance contracts and concomitant arguments for government support through subsidized premiums and reinsurance.

2 The US Federal Crop Insurance Program

The US federal crop insurance program has been in existence since 1938. However, it has taken on an increasingly prominent role in US agricultural policy in recent years. The current program exceeded $114 billion in liability in 2011 and insured 265 million acres.[1] The program involves significant subsidies, both to the farmers who "purchase" the insurance and the private crop insurance companies that administer and operate the programs. Smith (2011) notes that the private insurance companies that operate the program receive about $1.44 for every dollar of subsidy provided to farmers. Further, an evaluation of the program since 1981 reveals that the typical farmer receives about $1.88 in indemnity payments for every dollar of premium that they pay. In addition to such direct subsidies, the federal government operates a "Standard Reinsurance Agreement" that allows private insurance companies to retain low risks and to cede high risks to the taxpayer. The net result of such subsidies and the advantageous treatment of the private companies that administer the program is a very large and expensive federal program. Further, the program is currently one of the few that ratchets up the level of support as market conditions improve. Subsidies are largely based on total premium, such that higher commodity prices lead to higher premium levels and therefore to larger subsidies.

[1] Crop insurance statistics are taken from the online "Summary of Business" database of the Risk Management Agency (RMA-USDA). These statistics can be accessed at http://www.rma.usda.gov/data/sob.html.

The current program consists of a large number of different insurance instruments. Most prominent of these are the individual "yield" and "revenue" protection plans, that offer farmers federally–subsidized coverage of their individual farm's yields or crop revenues. In addition, a number of group "index" plans exist that base coverage on some index or aggregate indicator of crop yields or revenues. This includes the group risk program (GRP) and the group risk income protection (GRIP) program. In both cases, coverage and indemnities are based on county–level crop yields (or revenues). Smith and Goodwin (2010) discuss various operational aspects of the programs and provide an overview of their history and actuarial performance. In addition to crop insurance programs, livestock insurance coverage and a variety of other insurance mechanisms exist. For example, plans that base coverage on federal tax returns and satellite imagery of the "greenness" of the ground are currently available.

Since the last Farm Bill in 2007, the federal crop insurance program has cost taxpayers about $6 billion per year. This makes the crop insurance program the most costly farm program (excluding nutritional assistance programs). As commodity prices have increased, the costs associate with the federal crop insurance program have risen while other price support programs, which have fixed target and support prices, have largely been irrelevant. All indications arising from the current debate over the 2012 Farm Bill are that the programs are likely to further expand, with many calling for coverage of the "shallow losses" that comprise the deductible associated with the current programs (typically as low as 15% of liability).

3 Econometric Framework

Copula models have recently realized widespread application in empirical models of joint probability distributions.[2] The models essentially use a "copula" function to tie together two or more marginal probability functions that may (or may not) be related to one another.

[2]For details on construction and properties of copulas, see among others, Joe (1997) and Nelsen (2006).

Much of the work on copulas has been motivated by their applicability to the issues in risk management, insurance and financial economics (see, among others, Rodriguez (2003), Cherubini et al. (2004), Hu (2006), Patton (2006), and Jondeau and Rockinger (2006)). In the empirical literature, copula models have been used extensively in the design and rating of crop revenue insurance contracts, where the inverse correlation of prices and yields plays an important role in pricing revenue risk.

A p–dimensional copula, $C(u_1, u_2, \ldots, u_p)$, is a multivariate distribution function in the unit hypercube $[0, 1]^p$ with uniform $U(0, 1)$ marginal distributions. As long as the marginal distributions are continuous, a unique copula is associated with the joint distribution, F, that can be obtained as:

$$C(u_1, u_2, \ldots u_p) = F(F_1^{-1}(u_1), \ldots, F_p^{-1}(u_p)). \tag{1}$$

In a similar fashion, given a p–dimensional copula, $C(u_1, \ldots u_p)$, and p univariate distributions, $F_1(x_1), \ldots, F_p(x_p)$, the equation 1 is a p–variate distribution function with marginals $F_1, \ldots F_p$ whose corresponding density function can be written as:

$$f(x_1, x_2, \ldots x_p) = c(F_1(x_1), \ldots, F_p(x_p)) \prod_{i=1}^{p} f_i(x_i) \tag{2}$$

Provided that it exists, the density function of the copula, c, can be derived using equation 1 and marginal density functions, f_i:

$$c(u_1, u_2, \ldots u_p) = \frac{f(F_1^{-1}(u_1), \ldots, F_p^{-1}(u_p))}{\prod_{i=1}^{p} f_i(F_i^{-1}(u_i))}$$

There is a large number of parametric families of copulas applied in the literature. Two of the most commonly used copula families are elliptical copulas and Archimedean copulas. Gaussian and t–copulas are examples of elliptical copulas while the Clayton and Gumbel are among Archimedean copulas.

A multivariate density essentially conveys information about the distribution of individual random variables (through the marginals) and the interrelationships among individual variables. A number of different conceptual metrics are commonly used to measure and communicate these interrelationships—Pearson's linear correlation, Spearman rank correlation, and Kendall's τ measure of rank correlation. Copula models differ in terms of how these interrelationships are represented. For example, a Gaussian copula assumes linear correlation and imposes zero dependence in the tails of the distributions. A t copula allows for non-zero tail dependence (which increases as the degrees of freedom parameter falls) but imposes symmetry in the dependence relationships in alternative tails of the distributions. Archimedean copulas typically allow for dependence in only one tail and often represent the dependence relationship by using a single parameter, even when the copula includes multiple random variables. Thus, the choice of a copula function determines the nature of the relationships among dependent random variables. For example, while an Archimedean copula may be used to represent a multivariate distribution, it imposes a very strong set of restrictions on the dependency relationships among the variables. Our goal in this analysis is to achieve as much flexibility as is possible in representing the joint distribution of a set of dependent random variables (prices and crop yields) while, at the same time, maintaining a tractable approach to estimation and inference in light of the significant "curse of dimensionality" that such a multivariate problem presents. To this end, we consider multivariate versions of common elliptical and Archimedean copulas as well as a relatively new innovation in the representation of multivariate distributions—vine copulas.

Following Aas, et al. (2009), a joint, multivariate density function for a set of k random variables can be written in factored form as

$$f(x_1, x_2, \ldots, x_k) = f_k(x_k) \cdot f(x_{k-1}|x_k) \cdot f(x_{k-2}|x_{k-1}, x_k) \ldots \cdot f(x_1|x_2, \ldots, x_k). \tag{3}$$

This density is unique for a given ordering of variables. The joint density can also be expressed in terms of a copula function, as noted above, as

$$f(x_1, x_2, \ldots, x_k) = c_{1\ldots k}(F_1(x_1), \ldots, F_k(x_k)) \cdot \prod_{i=1}^{k} f_i(x_i). \tag{4}$$

In the case of two random variables, this reduces to

$$f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) f_2(x_2). \tag{5}$$

Thus, with rearranging, a bivariate conditional density can be written as

$$f(x_1|x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1). \tag{6}$$

Following this line of reasoning, Joe (1996) demonstrated that each of the terms in equation 3 can be decomposed into the product of a pair-wise copula and a conditional marginal density:

$$f(x|v) = c_{x,v_k|v_{-k}}(F(x|v_{-k}), F(v_k|v_{-k})) \cdot f(x|v_{-k}). \tag{7}$$

Thus, as Aas et al. (2009) demonstrate, a multivariate density can be expressed as a product of pair-copulas. It is important to note that many different factorizations and combinations of pair-wise copulas are possible. In the case of a canonical vine copula representation of a set of k random variables, a total of $k!/2$ different specifications is possible. A vine representation of a multivariate distribution is therefore dependent upon the specific decomposition into pair-wise conditional copulas, which in implementation will be reflected in the ordering of variables.

Following Joe's (1996) observations, recent research has focused on the notion of vine copulas as a means for representing higher-ordered distributions in terms of a combination of individual pair-wise copula functions. Bedford and Cooke (2002) introduced a "regular vine" representation that allows considerable flexibility in representing multivariate densities in terms of combinations of pair-wise copulas. Kurowicka and Cooke (2006) derived two special cases of vine copulas—the "canonical vine" and the "D-vine." In both cases, a general multivariate density is represented in terms of combinations of pair-wise copula functions. Both cases afford a degree of flexibility and generality not typically available in the application of conventional copula functions to higher-ordered problems. That said, it is important to note that any such representation is unique only with regard to a particular ordering of variables. Vine copulas are best represented in terms of a "tree," where the distribution of each variable is represented by conditional distributions at a higher level on the tree. The D– and canonical copulas differ in terms of the decomposition used to represent a multivariate density as combinations of pair-wise copula functions. As Aas et al. (2009) note, a D–vine has pair-wise combinations of variables in the initial level of the tree while the canonical-vine relates a single variable to all others in the initial level of the tree. Aas et al. (2009) note that a D–vine is most appropriate when a particular ordering of variables is suggested (such as in a time series context) while a canonical–vine is suggested when variables can be ordered according to their influence on other variables. We adopt a canonical–vine in this analysis and use the heuristic data-driven specification selection mechanism suggested by Brechmann and Czado (2011) in choosing the optimal ordering of data to define the vine. In particular, we follow Aas et al. (2009), Brechmann et al. (2010) and Dißmann et al. (2011) and choose the ordering that maximizes the dependencies in the first level of the tree. To this end, we choose the specification that maximizes the sum of the absolute values of Kendall's τ statistics in the first level of the tree.

Our estimation strategy involves the application of sequential maximum likelihood estimation of the pair–wise canonical copula model. The optimal copula functions for each conditional pair are chosen (again, heuristically) using the minimized value of the Akaike information criterion (AIC). A large variety of copula functions (forty in all) are considered for each combination. Likewise, we adopt standard maximum likelihood estimation techniques to estimate the joint densities associated with higher-ordered, multivariate elliptical and Archimedean copula models.[3]

The benchmark for our applied comparisons is the Gaussian copula model, which realizes significant prominence in pricing crop revenue risk in the current federal crop insurance program. In particular, current rating methods use the Iman and Conover (1982) method with normal score functions to represent the correlation associated with prices and yields in setting rates for revenue coverage. Mildenhall (2006) demonstrates that the Iman and Conover resorting procedure, when based upon normal scores, is essentially equivalent to the use of a Gaussian copula.

[3] Estimation and inferences were accomplished using the "COPULA" procedure of SAS and the "copula" and "CDVine" packages of the *R* language. Details are available in Chvosta, Erdman, and Little (2011), Schepsmeier and Brechman (2012), and Yan and Kojadinovic (2012). Excellent overviews of the *R* packages and implementation issues are presented by Yan (2007) and Czado (2011).

4 Empirical Application

The empirical application is intended to demonstrate the relevance of the techniques to the pricing of crop insurance and reinsurance contracts and to highlight the potential consequences associated with the choice of a specific representation of the joint distribution. As we have noted above, liability in the current federal crop insurance program for most major crops is overwhelmingly skewed toward crop revenue coverage, which involves the joint distribution of crop yields and prices. We utilize county–level crop yield data taken from the USDA's National Agricultural Statistics Service (NASS) databases. Relevant crop prices are taken as the average of February closing quotes on the Chicago Board of Trade for futures contracts that expire at harvest–time (November and December). These price quotes represent a market–based assessment made at the time of planting of the expected price after harvest. Such price quotes are used in pricing crop revenue insurance in the US. We focus on corn and soybeans—the two most prominent crops in the US. In 2011, these two crops made up 67.5% of the total liability of $114 billion in the US federal crop insurance program. We further focus on four specific counties in Illinois that are among the largest producers of corn and soybeans in the US Corn Belt. These four counties are in a common crop reporting district and thus are in close proximity to one another. The specific counties are McClean, Logan, Macon, and Tazewell.[4] Our data cover the 52–year period spanning 1960–2011.

An initial complication pertinent to any modeling of crop yields observed over time involves an adjustment for the significant upward trend that has characterized crop yields. This is commonly handled by applying a detrending process, with deviations from trend being "recentered" or "recalibrated" to a common time period. In our case, we utilize local quadratic regression (LOESS) to represent trends for each county–crop combination in a nonparametric fashion. We then recenter yields on 2011 by adding the deviations to the predicted 2011 yield. Specifically, we estimate the nonparametric trend equation

$$y_t = g(t) + \epsilon_t \tag{8}$$

[4] In the discussion of results that follows below, we denote these four counties as 1, 2, 3, and 4. This particular ordering reflected the prominence in terms of planted corn acreage in 2011.

and generate a sample of detrended yields as

$$\hat{y}_t = \hat{y}_{2011} + \epsilon_t. \tag{9}$$

The nonparametric LOESS estimates are illustrated below in figure 1.

We also independently estimate parametric marginals for each of the yield and price distributions. A number of different parametric specifications have been used to represent crop yield distributions. Common choices include the beta and Weibull distributions, both of which can accommodate the negative skewness commonly observed for crop yield distributions. We use the Weibull here in light of its simplicity.[5] In the case of prices, we adopt the common assumption of log–normality and model the log of the ratio of planting–time and harvest–time prices using a normal distribution. Plots of the detrended yield data and prices are presented in figures 2 and 3. As expected, a high degree of positive dependence among yields is apparent while negative dependence between yields and prices is also confirmed. This is consistent with the high degree of systemic risk that is reflected in the impact on yields of common weather conditions. Maximum likelihood estimates of the price and yield densities are presented in figures 4 and 5, respectively. The figures serve to validate the parametric choices for representing the marginals.

The set of 10 random variables associated with corn and soybean yields and prices in the four counties results in 45 correlation coefficients to be estimated in the Gaussian and t copula models. The t copula also requires estimation of an additional degrees of freedom parameter. As the value of this parameter rises, the t copula converges to the Gaussian (they are essentially equivalent for values above 30). In addition to estimating the Gaussian and t copulas, we also estimate single–parameter Archimedean copulas—the Clayton and Gumbel copulas. As noted above, these copulas provide a much more restrictive representation of the dependency structure.

[5] The Weibull distribution is represented by two parameters whereas the beta requires estimation of three or four parameters. Direct estimation of the minimum and maximum possible values for a beta can be challenging and we therefore opt in favor of the Weibull.

Maximum likelihood estimates and summary statistics of the Gaussian and t copulas are presented in table 1. Parameter estimates are very similar and the degrees of freedom parameter estimate is 20.69, reflecting the similarity between the two distributions. Maximum likelihood estimates of the two Archimedean copulas are presented in table 2. Log–likelihood function values and values of the AIC/SBC criteria strongly favor the more richly parameterized elliptical copula models. A consequence of the inflexibility associated with single parameter copula models is revealed in a consideration of the resulting dependency structure. Table 3 presents Kendall's τ for the actual sample and for the estimated Clayton copula. The limited flexibility associated with the single–parameter Archimedean copulas is apparent.

We used sequential maximum likelihood procedures to estimate the canonical vine copula. The ordering was based upon maximizing rank correlation in the first level of the vine, as is discussed above. This criterion yielded the following ordering of variables—(C_1, C_4, S_4, S_2, C_2, C_P, C_3, S_P, S_1, S_3).[6] The resulting vine copula parameter estimates along with the pair-wise copulas chosen for each node are presented in table 4. The vine copula model has a larger likelihood function value and smaller values of the AIC criterion, suggesting a superior fit over the more restrictive versions considered above. However, the SBC criterion, which applies a stronger penalty for additional parameters, narrowly favors the more parsimonious models. The canonical tree copula model structure implied by the estimates is presented in figure 6.

The critical question to be addressed in this research involves the extent to which pricing of insurance contracts based upon multiple sources of risk may be affected by the approach used to measure and represent dependence. This question is relevant on several levels. First, revenue insurance contracts, which consider two sources of risk—yield and price—are very common in the US federal crop insurance program. Yields and prices are, of course, inversely correlated and such dependence must be represented in pricing a revenue insurance contract

[6]Herein lies one apparent weakness in the vine copula representation. The resulting estimates are not invariant with respect to ordering. Given the set of ten random variables, 1,814,400 possible orderings exist. Other approaches to copula models of high–ordered multivariate distributions, such as Patton's (2012) factor model approach, do not suffer from this shortcoming but do involve other specification issues, such as defining the factors. Sensitivity of the estimates to such specification issues remains an important area of research.

in order to derive an actuarially–fair rate. The US federal program also offers "whole farm" revenue coverage for farmers growing both corn and soybeans. In this case, the total revenue from both crops provides the basis for coverage. Rates for such coverage are lower by virtue of the imperfect correlation of losses across crops. Finally, from a reinsurer's perspective, the pricing of a portfolio of risks is a critical factor in determining the terms of reinsurance treaties and contracts for coverage. In spite of the significant federal involvement in the US crop insurance program, reinsurance plays a very significant role in the industry. We consider the pricing of synthetic contracts that cover all revenue risks for a single crop across the four counties as well as pooled coverage across both crops in all counties (e.g., total revenue). We consider two levels of coverage—75% and 95% of expected revenue. Although the rates and loss probabilities are transparent to the commodity price for individual crop revenues, we use prices of $5.37 per bushel of corn and $13.62 per bushel of soybeans (reflecting the market prices at the time of the writing of this paper). Of course, the pooling of revenue across crops is impacted by the relative prices. We do not adjust the portfolio for differences in exposure (i.e., different levels of acreage) across the counties and therefore assume an identical level of acreage for all counties and both crops.

Using simulated, correlated uniform variates from each respective copula model and the estimated marginal distributions, we estimated loss probabilities and actuarially–fair premium rates for each contract. Loss probabilities are presented in table 5 and rates are presented in table 6. As expected, the probabilities and rates reflect the lower risks associated with pooling across various risks that are not perfectly correlated. The rates generally fall when the contract includes coverage across multiple counties and crops. The loss probabilities indicate that the probabilities of a payable claim also fall as the risks are further aggregated.

Most striking is the fact that the premium rates differ very significantly across the alternative copula models. For example, the rate for covering 95% of expected revenue for corn in county 1 is 4.66% according to the vine model while a Gaussian copula implies a rate of only 3.65%. In the case of a 95% coverage contract for total revenue, the vine copula model suggests an actuarially–fair premium rate of 3.75% while the Gaussian and t copula models imply rates of about 2.9%. To put this in perspective, total crop insurance liability for these crops in these four counties in 2010 was $650,308,546. Thus, the rate differences, if applied to the 2010 total crop insurance book, suggests a potential difference of over $5.5 million for these four counties alone. Thus, assumptions underlying the representation of dependencies among multiple sources of risk definitely have important impacts on the pricing, viability, and profitability of crop insurance contracts. In light of the huge magnitude of the federal program ($114 billion in liability in 2011), such seemingly small differences may translate into very significant implications for private insurers and taxpayers.

One interesting result is that the pricing that results from the canonical vine copula model are often closer to those implied by the very restrictive Archimedean copulas than the more flexible elliptical copulas. The Gaussian and t copula models tend to suggest less risk than either the vine or Archimedean models. This may reflect the "state–dependent" nature of agricultural yield and price risk, which is not captured in the Gaussian and t copula estimates. In particular, one expects that the imposition of zero tail dependence, as is the case for the Gaussian and the t copula estimates (which are very close to the Gaussian) may result in significantly underpricing portfolio risk. One expects that periods of significant yield shortfalls, such as in a drought, may experience a higher degree of correlation among yields in individual areas and therefore yields and prices on an aggregate level. Again, this reflects the systemic nature of weather and the fact that weather extremes may tend to impact a larger geographic area. Such conditions were observed in the 1988 drought and the 1993 Midwest floods, which caused widespread crop losses.

5 Concluding Remarks

The federal crop insurance program has been a major fixture of US agricultural policy since the 1930s. The program continues to grow in size and significance and now represents the most prominent farm policy instrument, accounting for more government spending than any other farm commodity program. Rising prices have expanded the scale of the program and the introduction of revenue insurance plans and increasing government premium subsidies have led to a program with a very high level of farmer participation.

Revenue insurance, which was introduced in the mid–1990s, involves multiple sources of dependent risk (i.e., prices and yields). Revenue coverage accounts for nearly 70% of the total liability in the program. The plans cover losses that result from a revenue shortfall that can be triggered by either low prices, low yields, or a combination of both. The actuarial practices currently applied in rating these plans essentially involve the application of a Gaussian copula model to the pricing of dependent risks. We evaluate the suitability of this assumption by considering a number of alternative copula models. We apply a relatively new innovation in copula modeling—the vine copula. This approach uses combinations of pair–wise copulas of conditional distributions to model multiple sources of risk. We find that this approach is generally preferred by model fitting criteria in the applications considered here. We also demonstrate that alternative approaches to modeling dependencies in a portfolio of risks may have significant implications to the pricing of such risks. Although this point is obvious to any observer of contemporary financial conditions, the implications for pricing crop revenue insurance have yet to be explored. Our paper is a first step in such an exploration.

The multivariate vine copulas presented here are not without their own limitations. In particular, the estimates are not invariant with respect to the factoring of the multivariate density, which is reflected in the ordering of individual variables in the model. In light of the substantial number of possible specifications that could be used to characterize dependency relationships, vine copulas have an inherent curse of dimensionality problem. Future research should explore more formal approaches to determining the most appropriate specification. Likewise, other approaches to higher–ordered copula models merit consideration and comparison to the estimates presented here.

References

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). "Pair-copula constructions of multiple dependence," Insurance: Mathematics and Economics 44 (2), 182-198.

Bedford, T. and R. M. Cooke (2002). "Vines—A New Graphical Model For Dependent Random Variables," *Annals of Statistics* 30, 1031-1068.

Brechmann, E. C. and C. Czado (2011). "Risk Management with High-Dimensional Vine Copulas: An Analysis of the Euro Stoxx 50," unpublished manuscript.

Brechmann, E. C. (2010). "Truncated and Simplified Regular Vines and Their Applications," Diploma Thesis, Faculty of Mathematics, Technische Universität München, Germany.

Breymann, W., Dias, A. and Embrechts, P. (2003), "Dependence structures for multivariate highfrequency data in finance". *Quantitative Finance,* 3, 1–14.

Cherubini, U., Luciano, E. and Vecchiato, W. (2004), *Copula Methods in Finance.* John Wiley and Sons, Chichester.

Chvosta, J. D. J. Erdman, and Mark Little, (2011), "Modeling Financial Risk Factor Correlation with the COPULA Procedure," SAS Global Forum, Paper 340-2011, SAS Institute Inc., Cary, NC.

Czado, C. (2011) "The World of Vines," Technische Universität München.

Deheuvels, P. (1979), "La Fonction de Dépendance Empirique et ses Propriétés: Un Test Non–Paramétrique d Indépendance." Académie Royale de Belgique. *Bulletin de la Classe des Sciences*, 5e Série 65, 274-292.

Demarta, S. and McNeil, A.J. (2005), "The t Copula and Related Copulas." *International Statistical Review*, 73, 111–129.

Dißmann, J. F., E. C. Brechmann, C. Czado, and D. Kurowicka (2011). "Selecting And Estimating Regular Vine Copulae And Application To Financial Returns," Unpublished Manuscript.

Embrechts, P., McNeil, A. and Straumann, D. (2002), Correlation and dependence in risk management: Properties and pitfalls. In: Risk Management: Value at Risk and Beyond, M.A.H. Dempster (ed.), pp. 176–223. Cambridge University Press, Cambridge.

Fang, H.B., Fang K.T. and Kotz, S. (2002), "The meta-elliptical distributions with given marginals." *Journal of Multivariate Analysis* 82, 1–16.

Glasserman, P. (2004), *Monte–Carlo Methods in Financial Engineering.* Springer–Verlag, New York.

Goodwin, B. K. (2001) "Problems with market insurance in agriculture," *American Journal of Agricultural Economics*, 83, 643-49.

Hu, L. (2006), "Dependence patterns across financial markets: A mixed copula approach." *Applied Financial Economics*, 10, 717–729.

Iman, R. and W. Conover (1982). "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables," *Communications in Statistics and Simulation Computing*, 11(3):311-334.

Joe, H. (1996). Families Of M-Variate Distributions With Given Margins And m/(m-1)=2 Bivariate Dependence Parameters. In L. Rüschendorf, B. Schweizer, and M. D. Taylor (Eds.), *Distributions with Fixed Marginals and Related Topics,* pp. 120-141. Hayward: Institute of Mathematical Statistics.

Joe, H. (1997), *Multivariate Models and Dependence Concepts.* Chapman and Hall, London.

Jondeau, E. and Rockinger, M. (2006), "The Copula-GARCH model of conditional dependencies: An international stock-market application." *Journal of International Money and Finance*, 25, 827–853.

Kurowicka, D. and R. M. Cooke (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling.* Chichester: John Wiley.

Mildenhall, S. J. (2005) "The Report of the Research Working Party on Correlations and Dependencies Among All Risk Sources; Part I: Correlation and Aggregate Loss Distributions With An Emphasis On The Iman-Conover Method," Report of the CAS Research Working Party on Correlations and Dependencies Among All Risk Sources.

Miranda, M. and Glauber, J. (1997) "Systemic risks, reinsurance, and the failure of crop insurance markets," *American Journal of Agricultural Economics*, 79, 206215.

Nelsen, R.B. (2006), *An Introduction to Copulas.* Springer-Verlag, New York .

Oh, D. H. and A. J. Patton (2011) "Modelling Dependence in High Dimensions with Factor Copulas," Unpublished Working Paper, Duke University October 2011.

Patton, A.J. (2006), "Modelling asymmetric exchange rate dependence." *International Economic Review,* 47, 527–556.

Reboredo, J.C. (2011), "How do crude oil prices co-move?: A copula approach." *Energy Economics*, 33, 948–955.

Rodriguez, J. C. (2003), "Measuring financial contagion: A copula approach." *Journal of Empirical Finance,* 14, 401–423.

Schweizer, B. and Sklar, A. (1983), *Probabilistic Metric Spaces.* Elsevier Science, New York.

Schepsmeier, U. and E. C. Brechmann. (2012) "Package CDVine: Statistical inference of C- and D-vine Copulas," *R*-Project CRAN Repository, February, 2012.

Sklar, A. (1959), "Fonctions de rèpartition àn dimensions et leurs marges." Publ. Inst. Statist. Univ. Paris 8: 229–231

Smith, V. (2011) "Premium Payments: Why Crop Insurance Costs Too Much," American Enterprise Institute, Washington, DC.

Smith, Vincent H., and Barry K. Goodwin (2010). Private and Public Roles in Providing Agricultural Insurance in the United States, in Jeffrey Brown (editor), *Private and Public Roles in Insurance,* AEI Press, Washington D.C.

Smith, M. S., Gan, Q. and Kohn, R.J. (2011). "Modelling Dependence Using Skew t Copulas: Bayesian Inference and Applications." *Journal of Applied Econometrics*, (in press).

Yan, J. (2007) "Enjoy the Joy of Copulas: With a Package copula," *Journal of Statistical Software,* Vol. 21, Issue 4.

Yan, J. and I. Kojadinovic (2012). "Package 'copula,': Multivariate Dependence With Copulas," *R*-Project CRAN Repository, February, 2012.

Table 1: Elliptical (Gaussian and t) Copula Estimates (Data Ordered as C1, C4, S4, S2, C2, CP, C3, SP, S1, S3)

	Gaussian Copula		t Copula	
Parameter	Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
ρ_{12}	0.9112	0.0170	0.9110	0.0182
ρ_{13}	0.5253	0.0835	0.5544	0.0828
ρ_{14}	0.5246	0.0833	0.5622	0.0829
ρ_{15}	0.8608	0.0262	0.8631	0.0270
ρ_{16}	-0.4709	0.0950	-0.4945	0.0977
ρ_{17}	0.8896	0.0214	0.8934	0.0219
ρ_{18}	-0.3281	0.1090	-0.3569	0.1139
ρ_{19}	0.5562	0.0777	0.5827	0.0791
ρ_{110}	0.6276	0.0684	0.6538	0.0676
ρ_{23}	0.4406	0.0953	0.4761	0.0947
ρ_{24}	0.4749	0.0933	0.5215	0.0929
ρ_{25}	0.9328	0.0132	0.9342	0.0138
ρ_{26}	-0.5025	0.0925	-0.5317	0.0945
ρ_{27}	0.8833	0.0227	0.8898	0.0228
ρ_{28}	-0.3574	0.1088	-0.3904	0.1137
ρ_{29}	0.3717	0.0957	0.4054	0.0982
ρ_{210}	0.5427	0.0804	0.5774	0.0795
ρ_{34}	0.9082	0.0179	0.9020	0.0205
ρ_{35}	0.4262	0.0957	0.4471	0.0953
ρ_{36}	-0.3737	0.1079	-0.3957	0.1091
ρ_{37}	0.4633	0.0875	0.4780	0.0888
ρ_{38}	-0.4918	0.0941	-0.4992	0.0977
ρ_{39}	0.8772	0.0241	0.8788	0.0259
ρ_{310}	0.8218	0.0337	0.8091	0.0382
ρ_{45}	0.4768	0.0926	0.5085	0.0922
ρ_{46}	-0.3637	0.1107	-0.3775	0.1152
ρ_{47}	0.4607	0.0862	0.4814	0.0882
ρ_{48}	-0.4531	0.0997	-0.4524	0.1070

Table 2: (Continued)

	Gaussian Copula		t Copula	
Parameter	Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
ρ_{49}	0.8373	0.0319	0.8368	0.0352
ρ_{410}	0.8534	0.0277	0.8483	0.0307
ρ_{56}	-0.5301	0.0873	-0.5454	0.0894
ρ_{57}	0.8854	0.0219	0.8878	0.0225
ρ_{58}	-0.3773	0.1046	-0.3992	0.1078
ρ_{59}	0.3689	0.0969	0.3890	0.0988
ρ_{510}	0.5156	0.0818	0.5352	0.0816
ρ_{67}	-0.5161	0.0886	-0.5247	0.0911
ρ_{68}	0.7430	0.0500	0.7427	0.0537
ρ_{69}	-0.2850	0.1198	-0.2768	0.1244
ρ_{610}	-0.3692	0.1089	-0.3681	0.1132
ρ_{78}	-0.3667	0.1050	-0.3830	0.1089
ρ_{79}	0.4347	0.0884	0.4450	0.0920
ρ_{710}	0.6527	0.0641	0.6636	0.0648
ρ_{89}	-0.3659	0.1113	-0.3551	0.1204
ρ_{810}	-0.3569	0.1069	-0.3399	0.1140
ρ_{910}	0.7679	0.0439	0.7576	0.0485
ν			20.6906	10.7926
LLF	352.5038		354.9279	
AIC	-615.0076		-617.8558	
BIC	-527.2016		-528.0986	

Table 2: Archimedean Copula Estimates (Clayton and Gumbel)

Clayton Copula		Gumbel Copula	
Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
0.3903	0.0579	1.2474	0.0398
LLF	56.4828		33.9548
AIC	-110.9655		-65.9096
BIC	-109.0143		-63.9584

Table 3: Kendall's τ Values from Actual Marginals (Upper Triangle) and Clayton Copula (Lower Triangle)

	C1	C2	C3	C4	S1	S2	S3	S4	CP	SP
C1		0.6878	0.7089	0.7602	0.3590	0.2941	0.4133	0.2640	-0.3273	-0.1991
C2	0.1720		0.7195	0.7707	0.2308	0.2624	0.3695	0.2262	-0.3710	-0.2549
C3	0.1585	0.1665		0.7164	0.2881	0.2534	0.4570	0.2564	-0.3439	-0.2157
C4	0.1574	0.1656	0.1546		0.2278	0.2624	0.3816	0.2383	-0.3469	-0.2308
S1	0.1544	0.1573	0.1516	0.1607		0.5701	0.5294	0.5762	-0.2021	-0.2006
S2	0.1587	0.1559	0.1595	0.1532	0.1625		0.6365	0.6652	-0.2609	-0.2926
S3	0.1659	0.1625	0.1635	0.1671	0.1666	0.1631		0.5641	-0.2685	-0.2278
S4	0.1603	0.1557	0.1586	0.1648	0.1482	0.1633	0.1713		-0.2730	-0.3469
CP	0.1634	0.1700	0.1579	0.1652	0.1586	0.1618	0.1683	0.1673		0.5520
SP	0.1653	0.1669	0.1637	0.1585	0.1587	0.1544	0.1603	0.1654	0.1575	

Table 4: Canonical Vine Copula Model Estimates

Copula Family	Parameter 1 Estimate	Standard Error	Parameter 2 Estimate	Standard Error
Rotated Gumbel 180	3.5987	0.4138		
Rotated BB7 180	1.3056	0.1918	0.5354	0.2592
Rotated BB7 180	1.4271	0.2252	0.5317	0.2593
Gaussian	0.8821	0.0270		
Frank	-3.4179	0.9165		
Rotated Gumbel 180	3.2173	0.3741		
Gaussian	-0.3077	0.1164		
Rotated BB1 180	0.3151	0.2299	1.3606	0.1889
BB7	1.6674	0.3239	0.7501	0.2800
Rotated Clayton 180	0.0556	0.1756		
Joe	1.1892	0.1635		
Rotated Gumbel 180	1.8636	0.2151		
Rotated Clayton 90	-0.3101	0.2136		
t	0.3899	0.1399	3.3926	1.7618
Rotated Clayton 90	-0.4049	0.2091		
Frank	-2.4208	0.9897		
Joe	1.0364	0.1230		
Gaussian	0.8668	0.0262		
Rotated Clayton 180	0.0431	0.1451		
Rotated Clayton 90	-0.4349	0.2296		
Rotated Clayton 90	-0.0630	0.1350		
Rotated Gumbel 90	-1.3123	0.1444		
Rotated Gumbel 180	2.7090	0.3157		
Rotated Gumbel 180	2.0319	0.2368		
Gaussian	0.0882	0.1338		
Rotated Joe 180	1.2554	0.1907		
Rotated Gumbel 90	-1.0909	0.1047		
Frank	0.6035	0.9004		
Gaussian	0.3767	0.1098		
Frank	3.1177	0.9071		
Rotated Clayton 90	-0.2784	0.1897		
Rotated BB7 180	1.1134	0.1344	0.4504	0.2098
Rotated Joe 270	-1.2088	0.1637		

Table 4: (Continued)

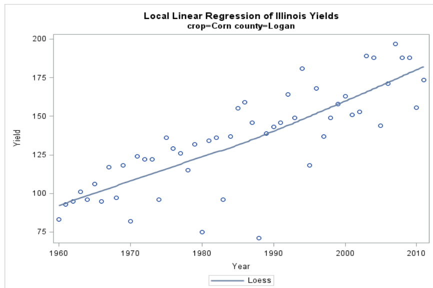
Copula Family	Parameter 1 Estimate	Standard Error	Parameter 2 Estimate	Standard Error
Rotated Clayton 270	-0.1519	0.1656		
Rotated Clayton 270	-0.1342	0.1819		
Rotated Joe 90	-1.0965	0.1685		
Rotated BB1 180	0.3605	0.2559	1.5048	0.2282
Rotated Joe 180	1.2361	0.1681		
Joe	1.1241	0.1579		
Clayton	0.0979	0.1540		
t	-0.0114	0.1625	2.8130	1.3422
Gaussian	0.6083	0.0798		
Rotated Clayton 270	-0.1589	0.1821		
Frank	1.8188	0.8460		
Rotated Joe 90	-1.2483	0.1696		
LLF	365.1950			
AIC	-624.3899			
BIC	-520.9740			

Table 5: Simulated Probabilities of Loss Claim

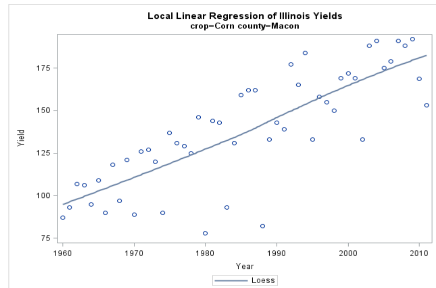
Insurance Instrument	Clayton	Gumbel	Gaussian	t	Canonical Vine
......75% Revenue Guarantee......					
Corn Revenue County 1	0.1117	0.1054	0.0257	0.0273	0.0553
Corn Revenue County 2	0.1078	0.0987	0.0195	0.0205	0.0467
Corn Revenue County 3	0.1101	0.1029	0.0217	0.0250	0.0470
Corn Revenue County 4	0.1153	0.1026	0.0198	0.0205	0.0503
Soybean Revenue County 1	0.0876	0.0849	0.0401	0.0452	0.0554
Soybean Revenue County 2	0.0831	0.0792	0.0297	0.0318	0.0408
Soybean Revenue County 3	0.0966	0.0927	0.0414	0.0436	0.0538
Soybean Revenue County 4	0.0869	0.0831	0.0306	0.0302	0.0380
Corn Revenue Total	0.1012	0.0756	0.0160	0.0185	0.0484
Soybean Revenue Total	0.0809	0.0688	0.0287	0.0319	0.0455
Total Revenue	0.0704	0.0337	0.0071	0.0128	0.0299
......95% Revenue Guarantee......					
Corn Revenue County 1	0.3942	0.4305	0.3916	0.3947	0.4234
Corn Revenue County 2	0.3942	0.4212	0.3941	0.3954	0.4291
Corn Revenue County 3	0.3914	0.4236	0.3991	0.4040	0.4320
Corn Revenue County 4	0.4011	0.4261	0.3949	0.3921	0.4311
Soybean Revenue County 1	0.3590	0.3765	0.3519	0.3584	0.3953
Soybean Revenue County 2	0.3629	0.3803	0.3489	0.3527	0.4039
Soybean Revenue County 3	0.3657	0.3865	0.3612	0.3627	0.4135
Soybean Revenue County 4	0.3590	0.3801	0.3459	0.3519	0.3936
Corn Revenue Total	0.3960	0.4299	0.3945	0.3945	0.4292
Soybean Revenue Total	0.3612	0.3893	0.3529	0.3543	0.3986
Total Revenue	0.3392	0.3876	0.3551	0.3563	0.4038

Table 6: Simulated Revenue Insurance Premium Rates

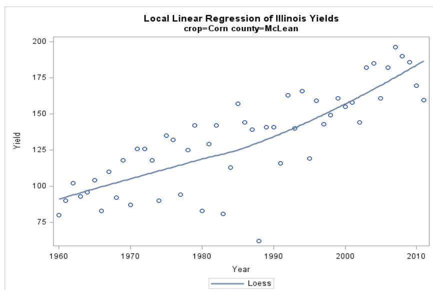
Insurance Instrument	Clayton	Gumbel	Gaussian	t	Canonical Vine
75% Revenue Guarantee					
Corn Revenue County 1	0.0142	0.0113	0.0017	0.0020	0.0042
Corn Revenue County 2	0.0151	0.0111	0.0014	0.0017	0.0030
Corn Revenue County 3	0.0153	0.0118	0.0014	0.0020	0.0035
Corn Revenue County 4	0.0134	0.0099	0.0011	0.0013	0.0035
Soybean Revenue County 1	0.0125	0.0094	0.0025	0.0032	0.0041
Soybean Revenue County 2	0.0100	0.0072	0.0013	0.0016	0.0024
Soybean Revenue County 3	0.0113	0.0087	0.0022	0.0027	0.0037
Soybean Revenue County 4	0.0124	0.0084	0.0015	0.0015	0.0024
Corn Revenue Total	0.0102	0.0043	0.0009	0.0013	0.0032
Soybean Revenue Total	0.0088	0.0049	0.0012	0.0017	0.0028
Total Revenue	0.0070	0.0015	0.0003	0.0006	0.0017
95% Revenue Guarantee					
Corn Revenue County 1	0.0628	0.0634	0.0365	0.0365	0.0466
Corn Revenue County 2	0.0622	0.0619	0.0346	0.0347	0.0446
Corn Revenue County 3	0.0625	0.0626	0.0361	0.0367	0.0458
Corn Revenue County 4	0.0619	0.0610	0.0345	0.0344	0.0457
Soybean Revenue County 1	0.0530	0.0522	0.0380	0.0401	0.0452
Soybean Revenue County 2	0.0509	0.0506	0.0353	0.0354	0.0418
Soybean Revenue County 3	0.0544	0.0542	0.0390	0.0399	0.0461
Soybean Revenue County 4	0.0529	0.0513	0.0351	0.0346	0.0405
Corn Revenue Total	0.0579	0.0559	0.0340	0.0343	0.0447
Soybean Revenue Total	0.0500	0.0478	0.0351	0.0358	0.0424
Total Revenue	0.0437	0.0392	0.0288	0.0297	0.0375



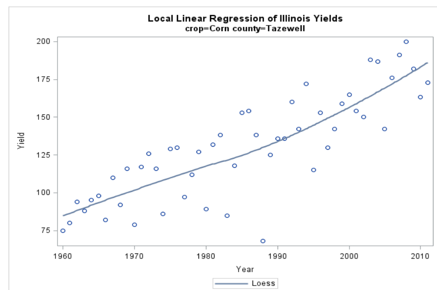
(a)



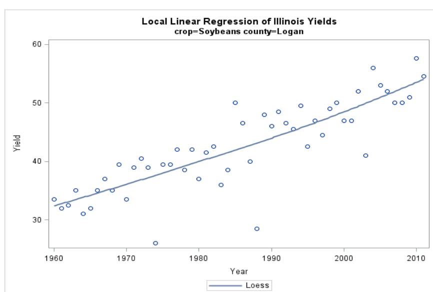
(b)



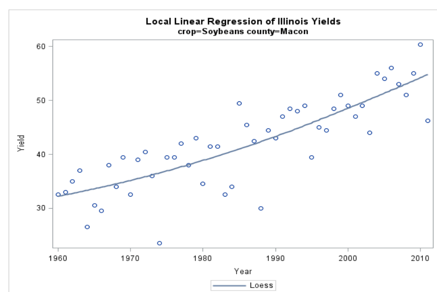
(c)



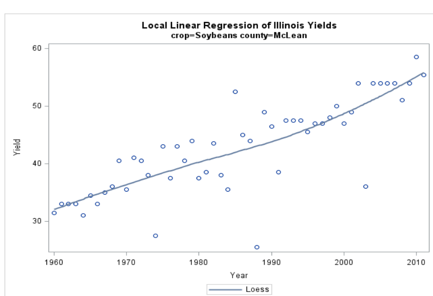
(d)



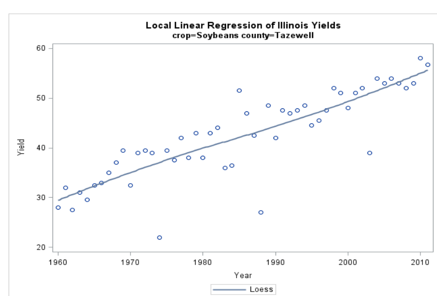
(e)



(f)



(g)



(h)

Figure 1: Local Linear Regression of Illinois Yield Trends

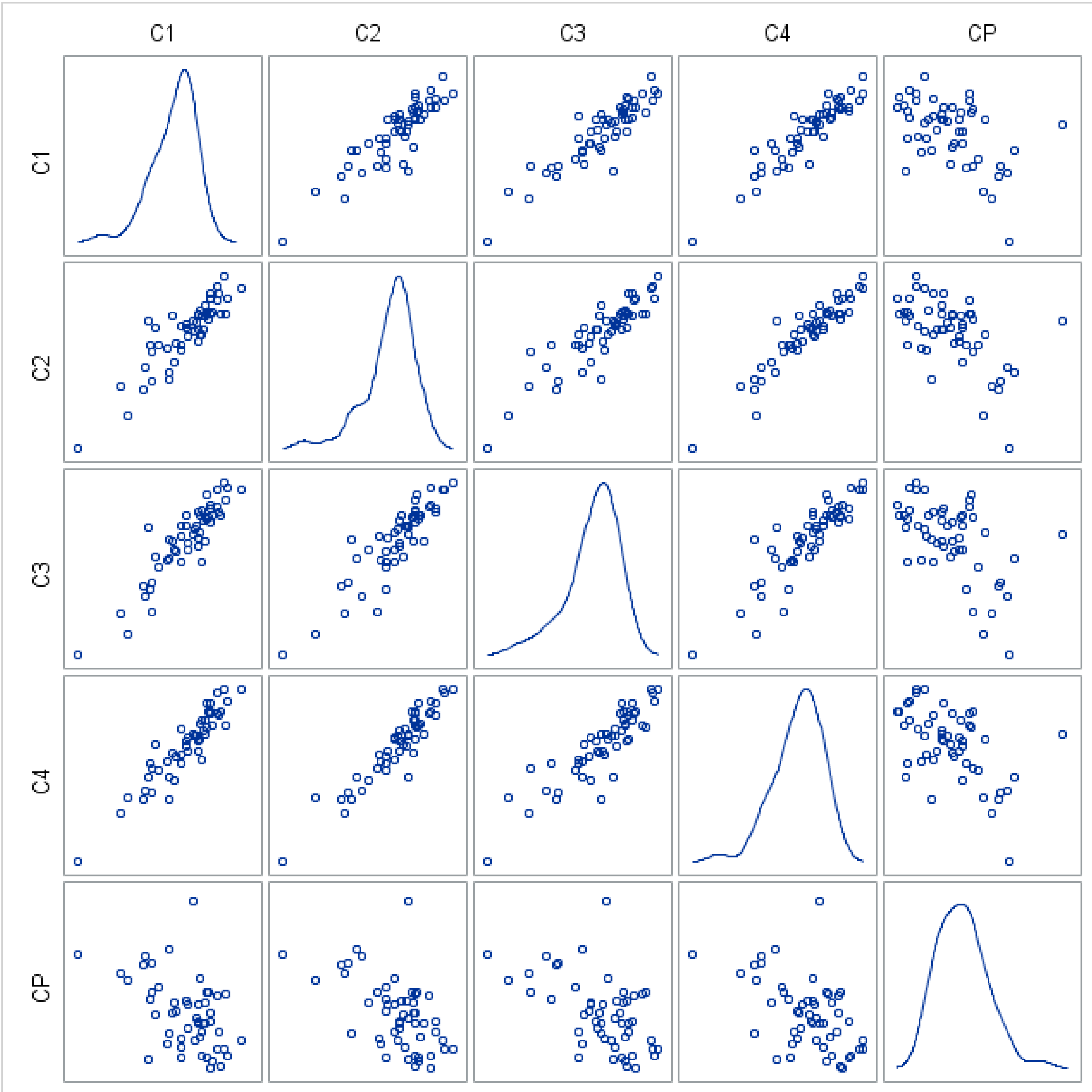


Figure 2: Empirical Distributions for Detrended Corn Yields and Price

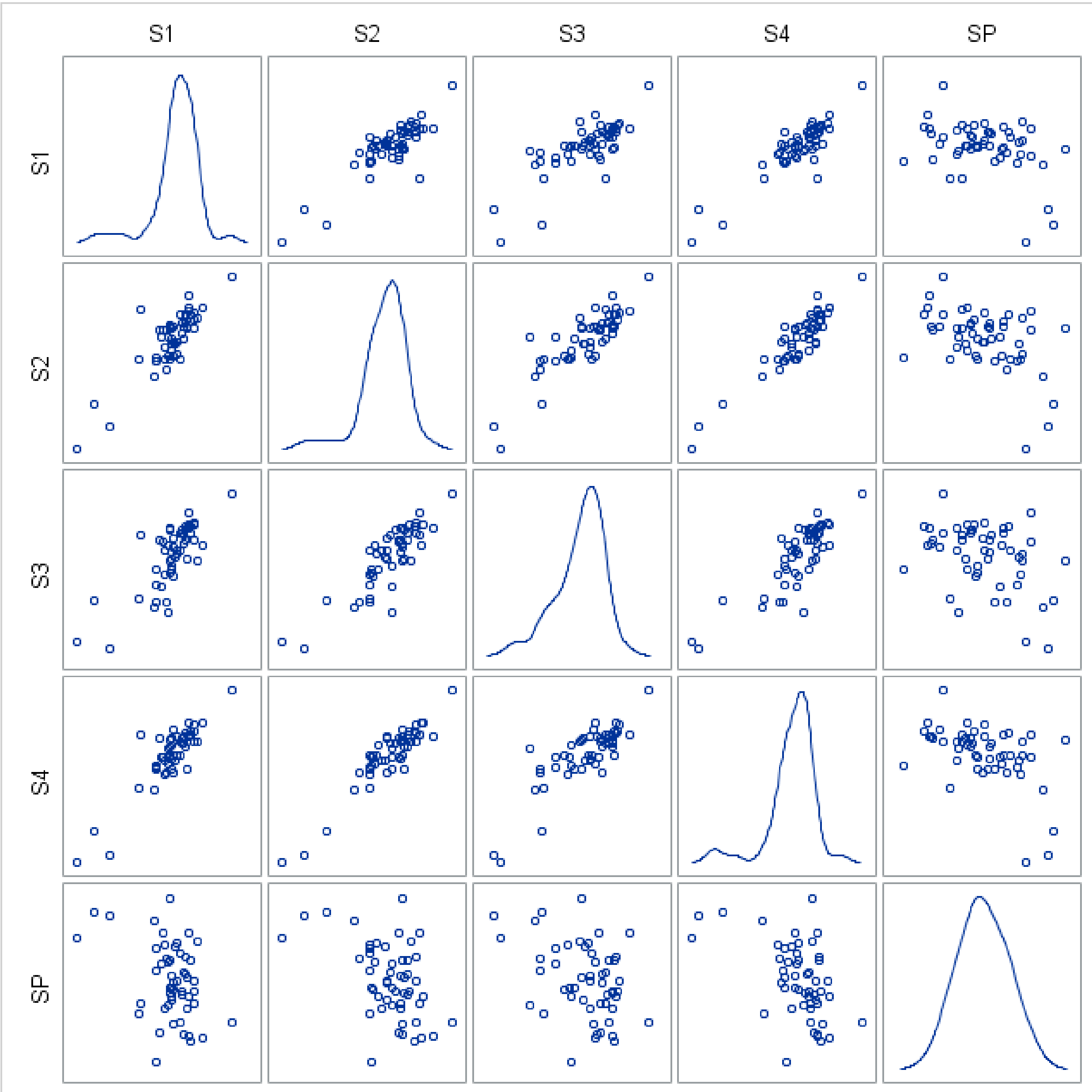
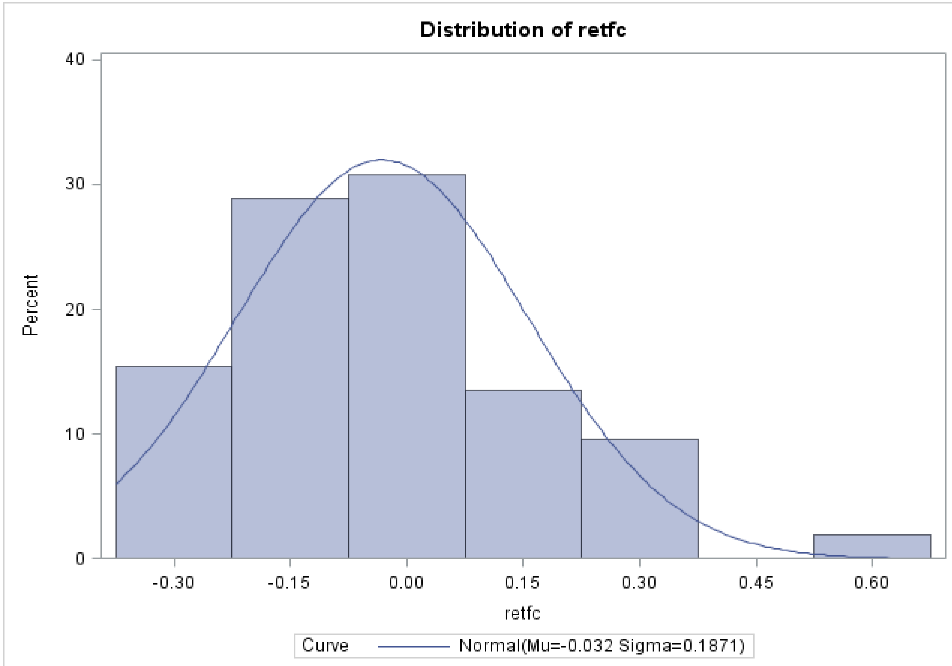
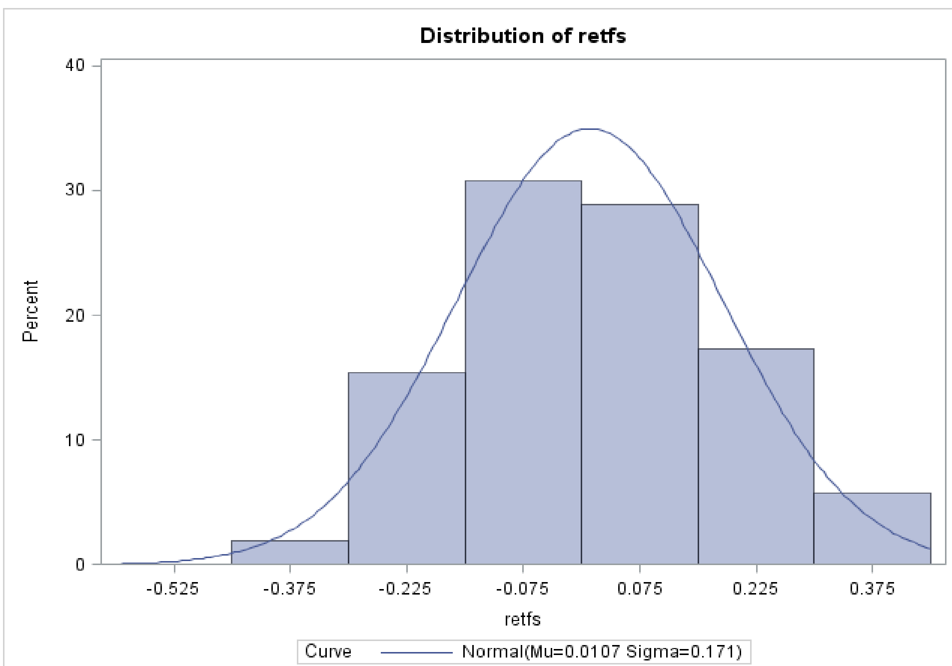


Figure 3: Empirical Distributions for Detrended Soybeans Yields and Price

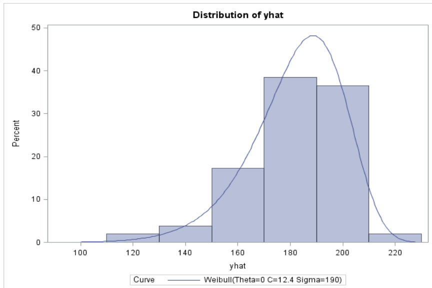


(a) Corn Price

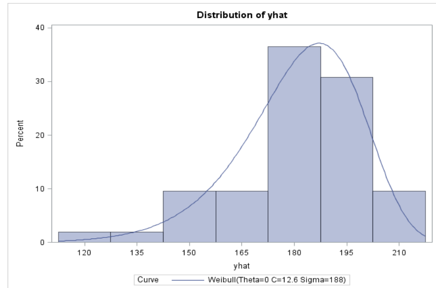


(b) Soybean Price

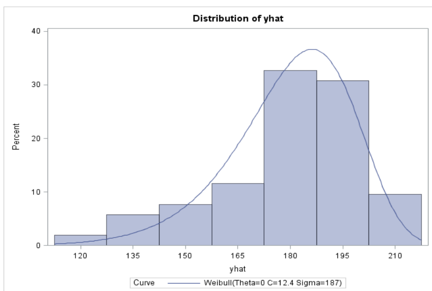
Figure 4: Fitted Normal Distributions for Logarithmic Price Changes



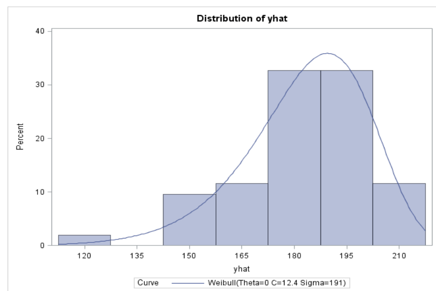
(a) Corn 1



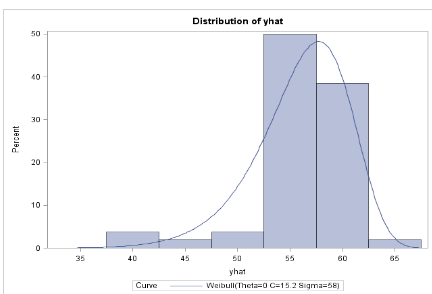
(b) Corn 2



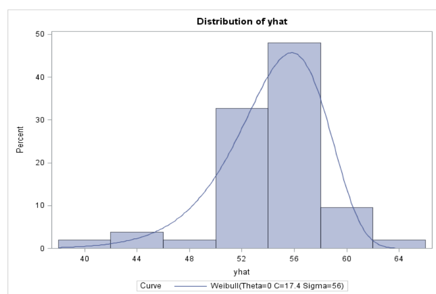
(c) Corn 3



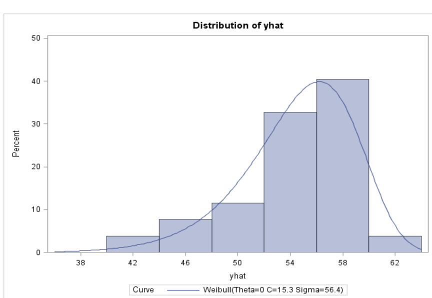
(d) Corn 4



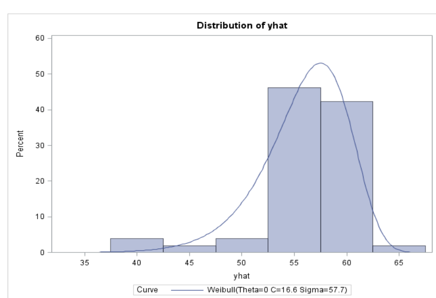
(e) Soy 1



(f) Soy 2

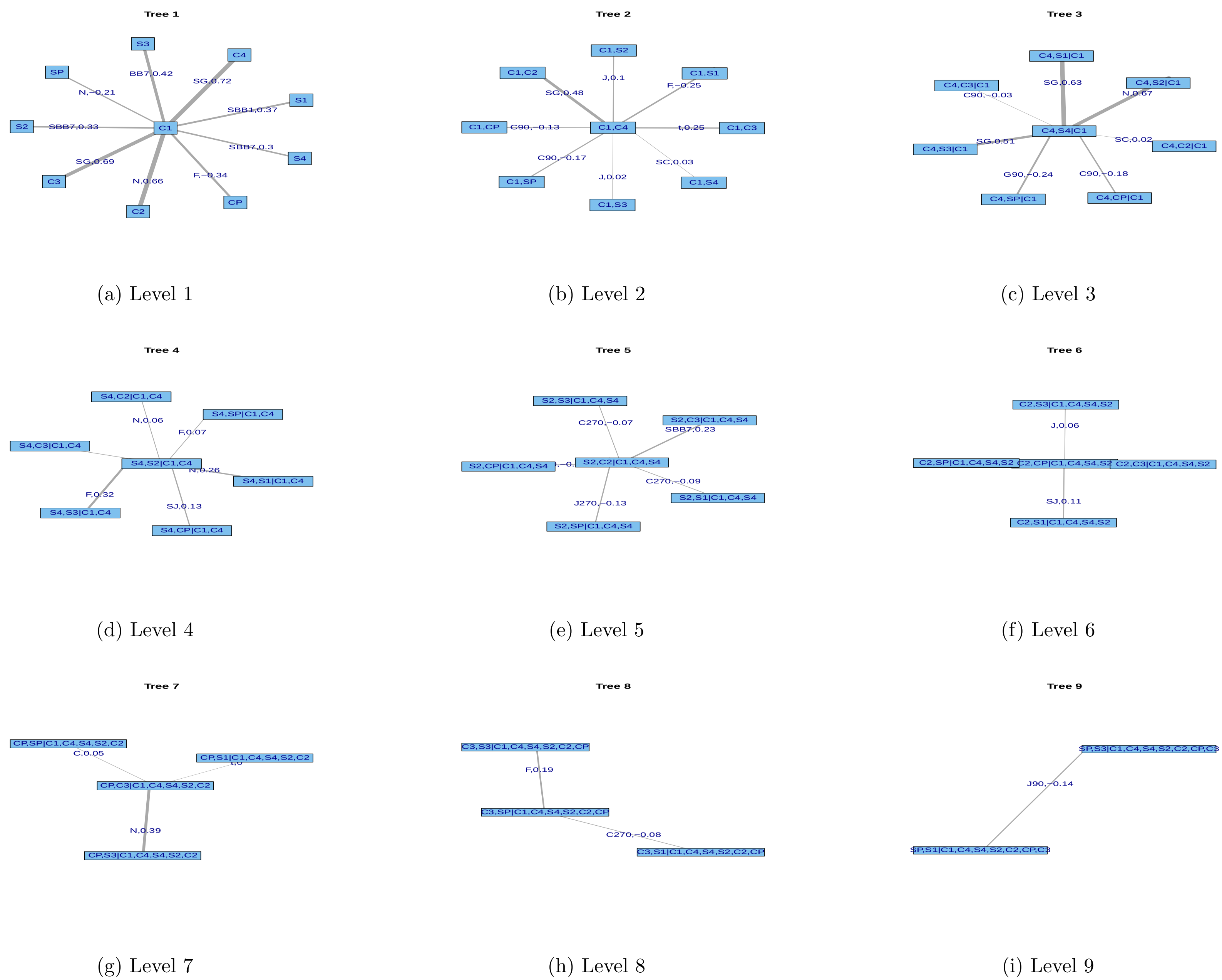


(g) Soy 3



(h) Soy 4

Figure 5: Fitted Weibull Distributions for Yields



(a) Level 1 (b) Level 2 (c) Level 3

(d) Level 4 (e) Level 5 (f) Level 6

(g) Level 7 (h) Level 8 (i) Level 9

Figure 6: Fitted Canonical Vine Copula Tree