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

Barry K. Goodwin and Ashley Hungerford

The federal crop insurance program has been a major fixture of U.S. agricultural policy since the 1930s, and continues to grow in size and importance. Indeed, it now represents the most prominent farm policy instrument, accounting for more government spending than any other farm commodity program. The 2014 Farm Bill further expanded the crop insurance program and introduced a number of new county-level revenue insurance plans. In 2013, over $123 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 available 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 when 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. In particular, we use 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 premium rates in crop insurance.

Key words: Copula models, crop insurance, systemic risk.

JEL codes: G22, Q14.

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 $123 billion and insures 295 million acres (Risk Management Agency 2014). 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 approximately $7.3 billion annually in subsidies to farmers and insurance companies, making it the most expensive agricultural commodity program (Risk Management Agency 2014). The 2014 Farm Bill further expanded crop insurance programs by introducing options to select "Agricultural Risk Coverage" (ARC), which functions as an aggregate county-level or whole-farm level revenue insurance program, as well as the "Supplemental Coverage Option" (SCO), which provides farmers with an optional, county-level insurance program that covers a portion of the deductible on existing crop insurance plans. Cotton has its own version of SCO in the "Stacked Income Protection Plan" (STAX). The crop insurance title also mandates development of a

Barry Goodwin is a William Neal Reynolds Distinguished Professor in the Departments of Economics and Agricultural and Resource Economics at North Carolina State University. Ashley Hungerford is a research agricultural economist at ERS-USDA. Earlier versions of this article were presented at the National Bureau of Economic Research Conference on Insurance Markets and Catastrophe Risk and in invited seminars at Georgia State University, Kansas State University, and Mississippi State University. Research support from the U.S. Forest Service and the North Carolina Agricultural Research Service is gratefully acknowledged. The views expressed here are those of the authors and should not be attributed to the ERS or the USDA. Correspondence to be sent to: barry_goodwin@ncsu.edu.

Amer. J. Agr. Econ. 97(3): 879–896; doi: 10.1093/ajae/aau086
Published online October 16, 2014

revenue-minus-cost insurance plan. All of these new insurance plans provide revenue coverage, meaning that payments may be triggered by multiple, dependent sources of risk—low prices, low yields, or a combination of both that results in a revenue shortfall.

Critics of subsidized crop insurance, such as Smith (2011), often 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 simply 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 (e.g., 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" risks. Empirical evidence has demonstrated that the spatial correlation of crop yields tends to be significantly stronger during extreme weather conditions such as droughts than is the case in a typical year (e.g., Goodwin 2001). As shown by Embrechts, McNeil, and Straumann (2002), standard models of systemic risk and insurance portfolio diversification nearly always assume that risks are linearly correlated and that this dependence is constant. In reality, the extent to which these risks may change across various states of nature has important implications for the pricing of revenue insurance and the availability of reinsurance.

This article 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. Our empirical analysis investigates two specific aspects of dependence in measuring the risks associated with crop insurance contracts and insurance portfolios. First, we consider the pricing of aggregate, county-level revenue insurance contracts. Revenue coverage currently accounts for about 70% of the total liability of the federal crop insurance program. County-level revenue insurance has existed under the Group Risk Income Protection (GRIP) program since 1999. According to unpublished Risk Management Agency (RMA) data, over 3 million acres and $3.5 billion in liability were insured under the GRIP program in 2013. In a second segment of the analysis, we consider the dependency relationships relevant to pricing yield insurance at the individual crop insurance unit level. Using unpublished RMA policy-level data on individual yield histories, we evaluate the risks associated with portfolios comprised of individual unit-level contracts. 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 non-zero tail-dependence—a critical factor when risks are state-dependent.

Empirical Framework

Though the concept dates back to work by Sklar (1959), copula models have recently realized widespread application in empirical models of joint probability distributions. Details on the construction and properties of copulas are provided by Joe (1997) and Nelsen (2006). 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 to form a joint probability distribution function. Much of the work on copulas has been motivated by their applicability to issues in risk management, insurance, and financial economics (see, among others, Rodriguez (2003); Cherubini, Luciano, and Vecchiato (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:

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

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

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

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

$$(3) \quad c(u_1, u_2, \ldots u_p) = \frac{f(F_1^{-1}(u_1), \ldots, F_p^{-1}(u_p))}{\prod_{i=1}^{p} f_i(F_i^{-1}(u_i))}.$$

There are several 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 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 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

$$(4) \quad f(x_1, x_2, \ldots, x_k) = f_k(x_k) \cdot f(x_{k-1}|x_k) \cdot f(x_{k-2}|x_{k-1}, x_k) \ldots \cdot f(x_1|x_2, \ldots, x_k).$$

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

$$(5) \quad f(x_1, x_2, \ldots, x_k) = c_{1\ldots k}(F_1(x_1), \ldots, F_k(x_k)) \cdot \prod_{i=1}^{k} f_i(x_i).$$

In the case of two random variables, this reduces to

$$(6) \quad f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) f_2(x_2).$$

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

$$(7) \quad f(x_1|x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1).$$

Following this line of reasoning, Joe (1996) demonstrated that each of the terms in equation (4) can be decomposed into the product of a pair-wise copula and a conditional marginal density:

$$(8) \quad f(x|v) = c_{x, v_k|v_{-k}}(F(x|v_{-k}), F(v_k|v_{-k})) \cdot f(x|v_{-k}).$$

Thus, as Aas et al. (2009) demonstrate, a multivariate density can be expressed as a product of pair-wise copulas. Following Joe's (1996) observations, recent research has focused on the notion of vine copulas as a means for representing high-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," also known as the "C-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 collection of "trees," where the distribution of each variable is represented by conditional distributions at a higher level on the tree. The D-vine and canonical vine 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.

It is important to note that many different factorizations and combinations of pair-wise copulas are possible. In the case of a canonical vine or D-vine copula representation of a set of k random variables, a total of $k!/2$ different specifications is possible.[1] 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. Examples of the trees used to depict canonical and D-vines—similar to the examples given in Bedford and Cooke (2002)—can be seen in figures 1 and 2, respectively. An additional inferential limitation of vine copulas has also been recently noted by Acar, Genest, and Nešlehová (2012). In particular, the factorization implicitly assumes a restricted relationship whereby the pair-wise conditional copula is invariant for all values of the conditioning variable. Although such structural restrictions and simplifying assumptions are typical in hierarchical structural models, it is important to note that the flexibility gained through vine copulas does indeed impose restrictions on the resulting multivariate distribution.

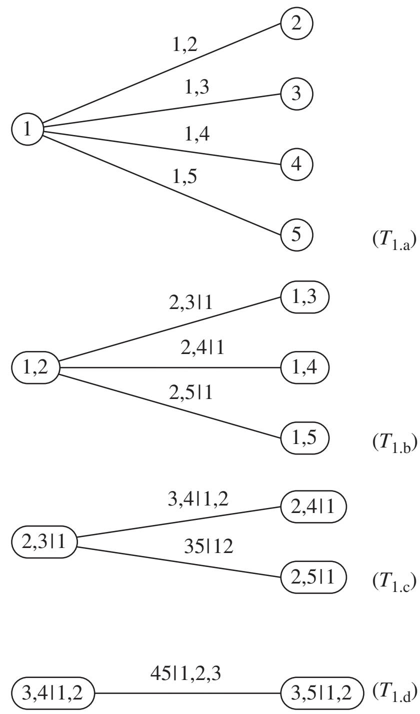

Figure 1. Trees for a canonical vine with five variables

As noted, vine copula models are determined by a particular ordering of the variables. Different heuristic data-driven specification selection mechanisms have been suggested in the literature. In the case of a canonical vine copula, Brechmann and Czado (2013) suggest adopting the ordering that maximizes the sum of pair-wise dependencies (measured by Kendall's τ) in the root node of the vine (i.e., the node with maximum column

[1] In the more general case of a regular vine, $\frac{k!}{2} \times \frac{(k-2)!}{2!}$.

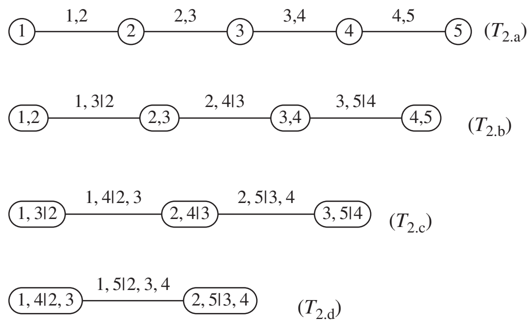

Figure 2. Trees for a D-vine with five variabless

sum in Kendall's τ matrix). For D-vines, we choose the specification that minimizes the Hamiltonian path of the nodes.[2]

Our estimation strategy involves the application of sequential 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 wide variety of copula functions (thirty-two in all) are considered for each combination. These copula functions include the Student *t*, Gaussian, Clayton, Gumbel, and Joe copulas, as well as other variations of the Archimedean copula family. 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.

Empirical Application

Our application consists of two empirical models of crop insurance contracts. The first addresses the specification and rating of county-level corn and soybean revenue insurance contracts. These contracts are fully analogous to the GRIP policies currently available in the U.S. program. Although this segment of the analysis applies to aggregate, county-level insurance contracts, its implications are entirely relevant to the rating of revenue insurance contracts at the individual farm level. In particular, indications of tail dependence in county-level revenue distributions may suggest a reevaluation of the current application of Gaussian copulas in farm-level insurance contracts. The second segment of our analysis utilizes unit-level data drawn from individual crop insurance policies to consider dependencies among yields among insurance units in a single county. 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.

County-Level Revenue Insurance

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 United States. We focus on corn and soybeans—the two

---

[2] The Hamiltonian or traceable path visits each node only once. Brechmann (2010) describes a solution to the "traveling salesman problem" that selects the ordering of variables in a D–vine. We adopt her suggestion in the analysis that follows.

[3] Estimation and inferences were accomplished using the "COPULA" procedure of SAS and the "copula," "VineCopula," and "CDVine" packages of the *R* language. Details are available in Chvosta, Erdman, and Little (2011), Schepsmeier and Brechmann (2012), Schepsmeier et al. (2013), and Yan and Kojadinovic (2012). Excellent overviews of the *R* packages and implementation issues are presented by Yan (2007) and Czado (2011).

most prominent crops grown in the United States. In 2013, these two crops made up 68% of the total liability of $123 billion in the U.S. federal crop insurance program (Risk Management Agency 2014). We further focus on four specific counties in Illinois that are among the largest producers of corn and soybeans in the U.S. 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 53-year period spanning 1960–2012.

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 regression (loess) to represent trends for each county-crop combination in a nonparametric fashion.[5] We then recenter yields on 2012 by adding the deviations to the predicted 2012 yield. Specifically, we estimate the nonparametric trend equation

$$y_t = g(t) + \epsilon_t \tag{9}$$

and generate a sample of detrended yields as

$$\hat{y}_t = \hat{y}_{2012} + \epsilon_t. \tag{10}$$

The nonparametric loess estimates are illustrated below in figure 3.

We use rank-based empirical distributions to represent the marginals. An alternative approach is to estimate parametric marginal distributions and use the cumulative distribution function (CDF) values of the estimates at each data point to estimate the copula. Using the nonparametric, empirical marginals is preferred in that the asymptotic distributions of the copula estimates are not affected by the first-stage estimation of the marginals, as has been shown by Chen and Fan (2006). Further, Charpentier, Fermanian, and Scaillet (2007) have noted that copula estimates based upon the empirical CDFs may be preferred because this approach can lead to smaller estimation variations compared to those based on the true marginals, even if known.[6] It should be noted that the estimates are made using detrended data and thus are conditional on the models used to detrend the yield data. This approach is very common when working with yield data collected over time. This approach is also followed in practice in the rating of county-level yield and revenue insurance contracts by the RMA. There is also an element of uncertainty in the estimation of the shape and location parameters of the parametric marginals used to simulate rates. Again, this first-stage estimation error has implications for the variability of the rate estimates. The fact that the data are detrended in a first stage necessarily results in an "Inference Functions for Margins" (IFM) estimator, even though the copulas are estimated using conventional maximum likelihood techniques. Joe (2005) provides additional details on the distinction and efficiency differences between full maximum likelihood and the IFM estimator.

In order to actually implement the estimated copula models in simulating random yields and prices to derive rate estimates, we need some representation of the marginal distributions. One could use either nonparametric or parametric marginals, though a parametric specification provides an explicit functional representation of the marginals. Thus, we also independently estimate parametric marginals for each of the yield and price distributions. These estimates are not used in estimating the copula models but are applied in the simulation of the random distributions of yields and prices. 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.[7] 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

[4] In the discussion of results that follows below, we denote these four counties as 1, 2, 3, and 4. This particular ordering reflected prominence in terms of planted corn acreage in 2012.

[5] The loess is estimated in SAS using PROC GAM. The supplementary online appendix provides a detailed explanation on loess and its estimation. For our analysis, we utilize a quadratic function in the loess regression.

[6] We are grateful to an anonymous referee for pointing out these advantages to the nonparametric approach.

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

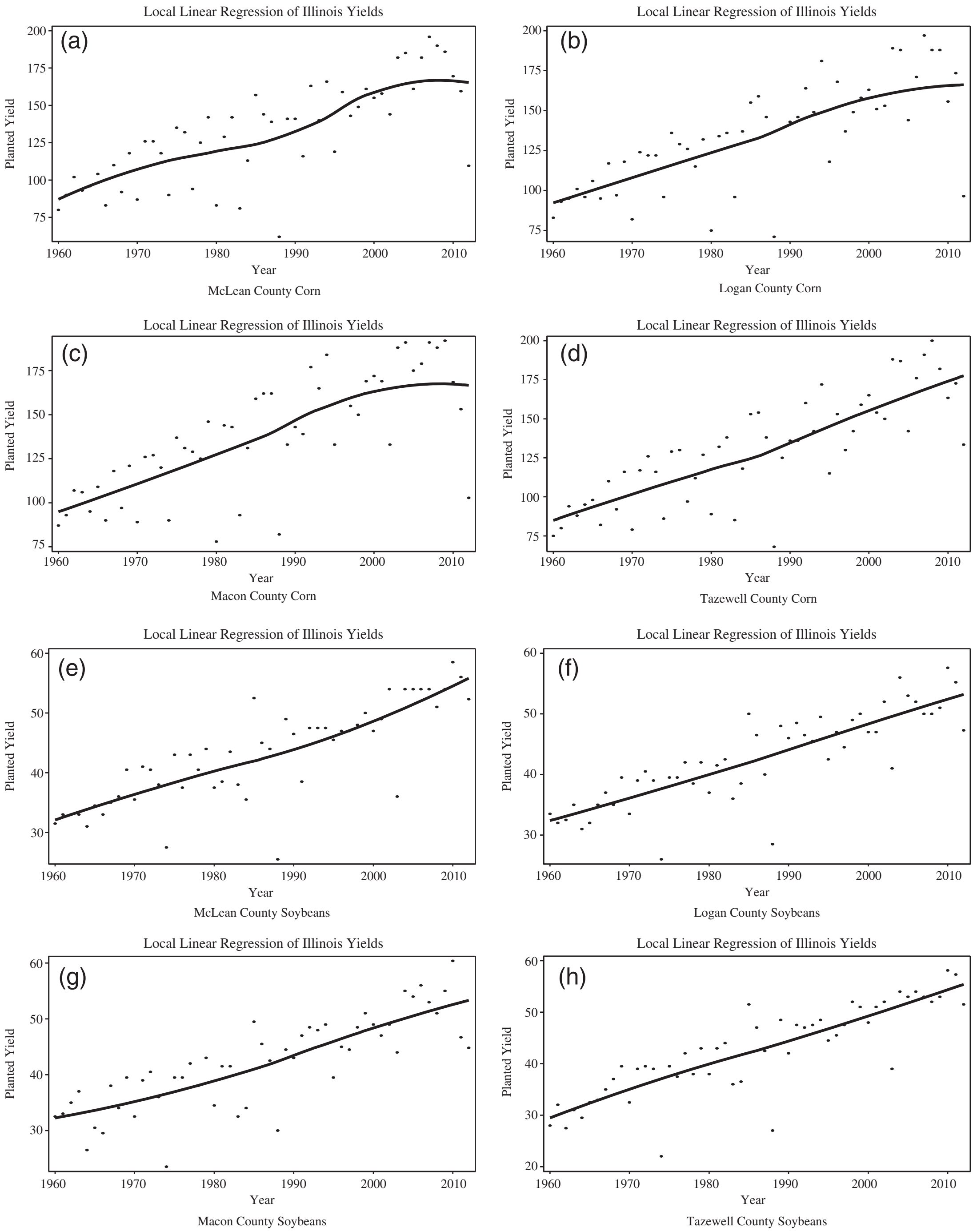

Figure 3. Local quadratic regression of Illinois yield trends

distribution. Plots of the detrended yield data and prices are presented in figures 4 and 5. 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.

The set of 10 random variables associated with corn and soybean yields and prices in the four counties results in 45 correlation

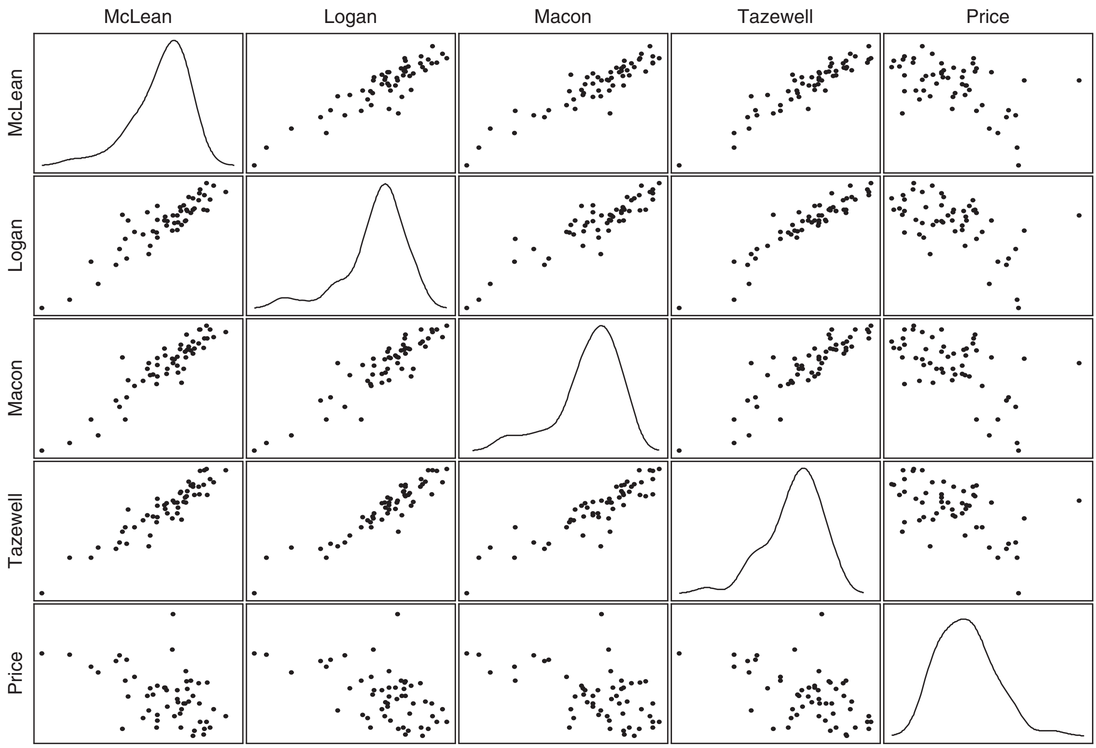

Figure 4. Empirical distributions for detrended corn yields and price

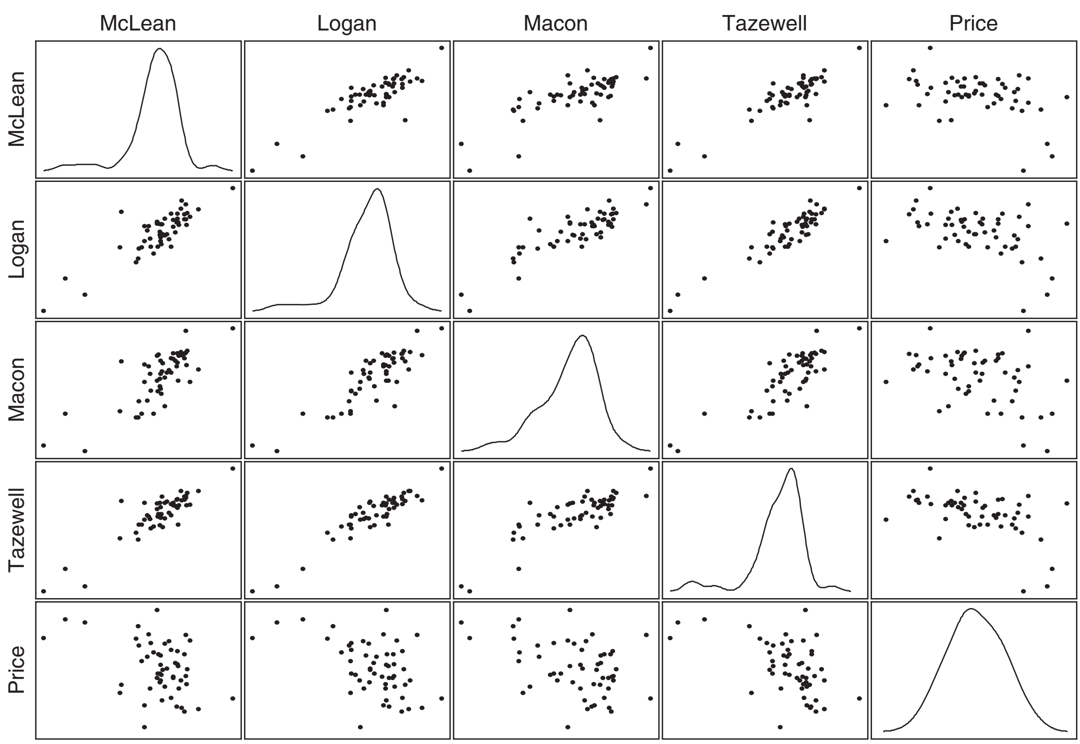

Figure 5. Empirical distributions for detrended soybeans yields and price

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. 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 5.84. The parameter estimates reflect the correlation patterns illustrated in figures 4 and 5.

We used sequential maximum likelihood procedures within the context of the IFM estimator to estimate the vine copulas; we considered both canonical and D-vine copula models. Upon estimation of both specifications, we conducted Vuong's (1989) non-nested specification test (Dißmann et al. 2013). The test statistic and other goodness-of-fit statistics (AIC and Schwarz Bayesian criteria [SBC]) are presented in table 2. The test statistic significantly favors the D-vine specification. Likewise, the D-vine model has a higher log-likelihood function value and smaller AIC and SBC values, all of which favors the D-vine over the C-vine specification. The heuristic approaches to determining the ordering of data (which was applied to both the C-vine and D-vine models) resulted in the following ordering of variables for the D-vine model: (C_2, C_4, C_1, C_3, S_3, S_2, S_4, S_1, S_P, C_P), where C and S indicate whether the crop is corn or soybeans, respectively.[8] The resulting vine copula parameter estimates along with the pair-wise copulas chosen for each node are presented in table 2. The vine copula model has a larger likelihood function value and smaller values of the AIC and Bayesian information criteria (BIC) than the Gaussian and *t* copulas, suggesting a superior fit over the more restrictive versions considered above.

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 U.S. federal crop insurance program. Expansions to the program brought about by the 2014 Farm Bill make the issues even more pressing. Yields and prices are, of course, inversely correlated and such dependence must be represented in pricing a revenue insurance contract in order to derive an actuarially-fair rate. The U.S. 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. Likewise, more complicated gross margin insurance plans that consider up to twenty-four correlated sources of risk exist for livestock products.[9] 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 U.S. 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 \$4.57 per bushel of corn and \$11.15 per bushel of soybeans (reflecting the market prices at the time of the writing of this article). 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 and corresponding premium rates

[8] Herein lies one apparent weakness in the vine copula representation. As noted, 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 Oh and 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.

[9] For example, the livestock gross margin (LGM) dairy insurance plan is based upon the combination of 24 futures contracts—12 for milk, 5 for corn, and 7 for soybean meal.

Table 1. Elliptical (Gaussian and *t*) Copula Estimates

	Gaussian Copula		*t* Copula	
Parameter	Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
$\rho_{C_1C_2}$	0.8868	0.0217*	0.8831	0.0252*
$\rho_{C_1C_3}$	0.9068	0.0180*	0.9013	0.0221*
$\rho_{C_1C_4}$	0.9286	0.0136*	0.9231	0.0170*
$\rho_{C_1S_1}$	0.5681	0.0777*	0.6466	0.0708*
$\rho_{C_1S_2}$	0.5368	0.0836*	0.6270	0.0757*
$\rho_{C_1S_3}$	0.6669	0.0636*	0.6995	0.0613*
$\rho_{C_1S_4}$	0.5281	0.0861*	0.5969	0.0798*
$\rho_{C_1C_p}$	−0.5351	0.0881*	−0.5401	0.0922*
$\rho_{C_1S_p}$	−0.3873	0.1082*	−0.4072	0.1161*
$\rho_{C_2C_3}$	0.9017	0.0190*	0.9048	0.0215*
$\rho_{C_2C_4}$	0.9378	0.0121*	0.9400	0.0134*
$\rho_{C_2S_1}$	0.4067	0.0964*	0.4639	0.0926*
$\rho_{C_2S_2}$	0.5092	0.0917*	0.5924	0.0828*
$\rho_{C_2S_3}$	0.5719	0.0769*	0.6054	0.0734*
$\rho_{C_2S_4}$	0.4552	0.0972*	0.5089	0.0919*
$\rho_{C_2C_P}$	−0.5734	0.0827*	−0.5518	0.0895*
$\rho_{C_2S_p}$	−0.4259	0.1045*	−0.4134	0.1145*
$\rho_{C_3C_4}$	0.9126	0.0169*	0.9148	0.0189*
$\rho_{C_3S_1}$	0.4875	0.0869*	0.5391	0.0869*
$\rho_{C_3S_2}$	0.4968	0.0858*	0.5580	0.0840*
$\rho_{C_3S_3}$	0.6977	0.0583*	0.6980	0.0610*
$\rho_{C_3S_4}$	0.4808	0.0895*	0.5214	0.0902*
$\rho_{C_3C_p}$	−0.5650	0.0839*	−0.5439	0.0914*
$\rho_{C_3S_p}$	−0.4078	0.1058*	−0.4097	0.1141*
$\rho_{C_4S_1}$	0.4020	0.0951*	0.4785	0.0913*
$\rho_{C_4S_2}$	0.4996	0.0917*	0.5955	0.0827*
$\rho_{C_4S_3}$	0.5942	0.0738*	0.6393	0.0705*
$\rho_{C_4S_4}$	0.4515	0.0967*	0.5252	0.0908*
$\rho_{C_4C_p}$	−0.5724	0.0829*	−0.5829	0.0861*
$\rho_{C_4S_p}$	−0.4119	0.1059*	−0.4312	0.1150*
$\rho_{S_1S_2}$	0.8011	0.0385*	0.7935	0.0471*
$\rho_{S_1S_3}$	0.7606	0.0455*	0.7566	0.0527*
$\rho_{S_1S_4}$	0.8111	0.0368*	0.8083	0.0438*
$\rho_{S_1C_p}$	−0.3520	0.1132	−0.3671	0.1156*
$\rho_{S_1S_p}$	−0.4026	0.1091	−0.4210	0.1214*
$\rho_{S_2S_3}$	0.8453	0.0293	0.8590	0.0307*
$\rho_{S_2S_4}$	0.8903	0.0213*	0.8692	0.0291*
$\rho_{S_2C_p}$	−0.4420	0.1029*	−0.4735	0.1034*
$\rho_{S_2S_p}$	−0.4964	0.0959*	−0.4924	0.1056*
$\rho_{S_3S_4}$	0.8161	0.0352*	0.8015	0.0428*
$\rho_{S_3C_p}$	−0.4225	0.1035*	−0.4088	0.1071*
$\rho_{S_3S_p}$	−0.4282	0.1043*	−0.4120	0.1121*
$\rho_{S_4C_p}$	−0.4524	0.1010*	−0.4948	0.0982*
$\rho_{S_4S_p}$	0.5315	0.0910*	0.5509	0.0969*
$\rho_{C_pC_p}$	0.7694	0.0453*	0.7717	0.0516*
ν			5.8463	1.7839*
LLF	333.7		343.5	
AIC	−577.4		−595.0	
BIC	−488.7		−504.4	

Note: An asterisk indicates statistical significance at the $\alpha = .10$ or smaller level.

Table 2. D-vine Copula Model Estimates

Factorization	Copula Family	Parameter 1 Estimate	Standard Error	Parameter 2 Estimate	Standard Error
C_2, C_4	R_{180} Gumbel	4.8626	0.5685*		
C_4, C_1	R_{180} Gumbel	4.3485	0.4956*		
C_1, C_3	R_{180} Gumbel	3.7854	0.4337*		
C_3, S_3	Clayton	1.6533	0.3401*		0.6848*
S_3, S_2	BB7	2.1508	0.4449*	3.2092	
S_2, S_4	BB7	2.6232	0.5094*	3.7795	0.8183*
S_4, S_1	BB7	2.5324	0.4776*	2.4664	0.6429*
S_1, S_P	R_{90} Joe	−1.6194	0.2390*		
S_P, C_P	Student t	0.7737	0.0632*	2.9862	1.9575
$C_2, C_1 \| C_4$	Frank	0.6322	0.8196		
$C_4, C_3 \| C_1$	R_{180} Gumbel	1.3238	0.1411*		
$C_1, S_3 \| C_3$	Joe	1.2067	0.1815*		
$C_3, S_2 \| S_3$	Student t	−0.1619	0.1620		
$S_3, S_4 \| S_2$	Frank	1.3875	0.7733*		
$S_2, S_1 \| S_4$	R_{180} Clayton	0.3700	0.1850*		
$S_4, S_P \| S_1$	Frank	−2.5358	0.8926*		
$S_1, C_P \| S_P$	R_{270} Clayton	−0.2250	0.1921		1.2410*
$C_2, C_3 \| C_4, C_1$	Student t	0.3141	0.1448*	2.6997	
$C_4, S_3 \| C_1, C_3$	R_{90} Clayton	−0.2671	0.1804		
$C_1, S_2 \| C_3, S_3$	Gaussian	0.2389	0.1206*		
$C_3, S_4 \| S_3, S_2$	R_{90} Joe	−1.2063	0.1569*		
$S_3, S_1 \| S_2, S_4$	Frank	1.0841	0.8475		
$S_2, S_P \| S_4, S_1$	R_{90} Clayton	−0.1956	0.1712		
$S_4, C_P \| S_1, S_P$	Gaussian	−0.0816	0.1310		
$C_2, S_3 \| C_4, C_1, C_3$	R_{90} Clayton	−0.1926	0.1538		
$C_4, S_2 \| C_1, C_3, S_3$	Frank	2.2234	0.8666*		
$C_1, S_4 \| C_3, S_3, S_2$	Joe	1.2615	0.1787		
$C_3, S_1 \| S_3, S_2, S_4$	Clayton	0.1042	0.1360		
$S_3, S_P \| S_2, S_4, S_1$	Frank	0.6354	0.8823		
$S_2, C_P \| S_4, S_1, S_P$	R_{270} Joe	−1.1270	0.1346*		
$C_2, S_2 \| C_4, C_1, C_3, S_3$	Gaussian	0.3806	0.1070*		
$C_4, S_4 \| C_1, C_3, S_3, S_2$	Frank	−0.8525	0.7817		
$C_1, S_1 \| C_3, S_3, S_2, S_4$	Gaussian	0.2421	0.1201*		
$C_3, S_P \| S_3, S_2, S_4, S_1$	R_{90} Clayton	−0.3240	0.2045		
$S_3, C_P \| S_2, S_4, S_1, S_P$	R_{90} Joe	−1.0953	0.1083*		
$C_2, S_4 \| C_4, C_1, C_3, S_3, S_2$	R_{90} Joe	−1.1636	0.1631*		
$C_4, S_1 \| C_1, C_3, S_3, S_2, S_4$	Frank	−4.3714	1.0200*		
$C_1, S_P \| C_3, S_3, S_2, S_4, S_1$	Gaussian	0.0885	0.1352		
$C_3, C_P \| S_3, S_2, S_4, S_1, S_P$	Gaussian	−0.3706	0.1142*		
$C_2, S_1 \| C_4, C_1, C_3, S_3, S_2, S_4$	Frank	−1.1901	0.8107		
$C_4, S_P \| C_1, C_3, S_3, S_2, S_4, S_1$	R_{90} Gumbel	−1.0859	0.0872*		
$C_1, C_P \| C_3, S_3, S_2, S_4, S_1, S_P$	R_{90} Clayton	−0.2097	0.1868		
$C_2, S_P \| C_4, C_1, C_3, S_3, S_2, S_4, S_1$	Gaussian	−0.0620	0.1388		
$C_4, C_P \| C_1, C_3, S_3, S_2, S_4, S_1, S_P$	Clayton	0.0831	0.1324		
$C_2, C_P \| C_4, C_1, C_3, S_3, S_2, S_4, S_1, S_P$	R_{180} Joe	1.0956	0.1865*		
C-Vine Log-Likelihood	358.7699				
D-Vine Log-Likelihood	380.6401				
C-Vine AIC	−623.5399				
D-Vine AIC	−659.2802				
C-Vine SBC	−530.9361				
D-Vine SBC	−558.7953				
Vuong Test (C-Vine vs. D-Vine)	−2.4193				
Vuong Test p-Value	0.0156				

Note: An asterisk indicates statistical significance at the $\alpha = .10$ or smaller level.

Table 3. Revenue Insurance Probabilities of Claims and Rates

Insurance	Gaussian		Student *t*		D-vine	
Instrument	Prob(Loss)	Rate	Prob(Loss)	Rate	Prob(Loss)	Rate
75% Coverage Level						
Corn Revenue County 1	0.0458	0.0029	0.0452	0.0035	0.0483	0.0034
Corn Revenue County 2	0.0415	0.0026	0.0446	0.0035	0.0490	0.0034
Corn Revenue County 3	0.0422	0.0026	0.0454	0.0036	0.0523	0.0038
Corn Revenue County 4	0.0391	0.0023	0.0393	0.0029	0.0423	0.0027
Soybean Revenue County 1	0.0201	0.0010	0.0209	0.0014	0.0243	0.0014
Soybean Revenue County 2	0.0131	0.0006	0.0157	0.0009	0.0161	0.0008
Soybean Revenue County 3	0.0203	0.0010	0.0233	0.0017	0.0229	0.0013
Soybean Revenue County 4	0.0115	0.0005	0.0129	0.0007	0.0173	0.0009
Corn Revenue Total	0.0382	0.0022	0.0406	0.0030	0.0452	0.0030
Soybean Revenue Total	0.0129	0.0005	0.0151	0.0009	0.0183	0.0010
Total Revenue	0.0149	0.0006	0.0194	0.0012	0.0211	0.0012
95% Coverage Level						
Corn Revenue County 1	0.4259	0.0449	0.4203	0.0437	0.4168	0.0446
Corn Revenue County 2	0.4253	0.0436	0.4213	0.0436	0.4217	0.0452
Corn Revenue County 3	0.4259	0.0439	0.4207	0.0438	0.4183	0.0458
Corn Revenue County 4	0.4224	0.0428	0.4166	0.0416	0.4181	0.0431
Soybean Revenue County 1	0.3844	0.0334	0.3759	0.0321	0.3583	0.0322
Soybean Revenue County 2	0.3764	0.0301	0.3696	0.0297	0.3538	0.0289
Soybean Revenue County 3	0.3846	0.0334	0.3786	0.0329	0.3680	0.0325
Soybean Revenue County 4	0.3735	0.0292	0.3646	0.0281	0.3467	0.0285
Corn Revenue Total	0.4223	0.0425	0.4170	0.0420	0.4166	0.0438
Soybean Revenue Total	0.3760	0.0300	0.3678	0.0294	0.3518	0.0295
Total Revenue	0.3902	0.0321	0.3819	0.0318	0.3718	0.0319

are presented in table 3. 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.

The premium rates differ across the alternative copula models, and the differences are substantial in some cases. For example, the rate for covering 75% of expected revenue for corn in county 3 is 0.38% according to the vine copula model, while a Gaussian copula implies a rate of only 0.26%. In the case of a 75% coverage contract for total revenue, the vine copula model suggests an actuarially-fair premium rate of 0.12%, while the Gaussian copula implies a rate of 0.06%. To put this into perspective, in 2013 total crop insurance liability for these crops in these four counties was $988,647,318. Thus, the rate differences, if applied to the 2013 total crop insurance book, suggests a potential difference of over $593,000 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 ($123 billion in liability in 2013), such seemingly small differences may translate into very significant implications for private insurers and taxpayers.

Perhaps most important is the finding that the rate differences are much greater because one considers coverage of lower-probability events, which correspond to deep losses occurring in the tails of the distributions. The average percentage rate differences between the D-vine copula and the Gaussian copula across all contracts is −0.6% at the 95% coverage level, and 50% at the 75% coverage level. The rate differences between the Gaussian copula and the *t* copula are similar to the rate differences between the Gaussian copula and D-vine copula. Thus, the differences in rates and implications of applying a copula that is less supported by the data are exaggerated in the tails of the revenue distribution, which reflects significant differences in tail dependence and the risks of deep losses across the alternative copula models.

One interesting result is that the pricing that results from the D-vine copula model are closer to those implied by the *t* copula than is the case for the Gaussian copula that is currently used to rate revenue contracts. Thus, the imposition of zero tail dependence in the Gaussian copula model appears to result in greater pricing deviations compared to the results of the *t* copula. This is not unexpected since the *t* copula allows for non-zero tail dependence, though such dependence is restricted to be symmetric in alternate tails. The estimated degrees of freedom parameter in the *t* copula was 5.85 (table 1), which indicates a substantial degree of tail dependence compared to the Gaussian copula. Put differently, the *t* copula, which converges to a Gaussian copula as the degrees of freedom parameter increases, offers additional flexibility that is supported by the data and that results in substantial differences in estimates of loss-probabilities and premium rates. This may imply that improvements in the accuracy of revenue insurance premium rates could be possible by considering a *t* copula as an alternative to the Gaussian copula currently in use.

The Gaussian copula model tends to suggest less tail risk and lower rates than is the case with the D-vine copula. This may reflect the "state-dependent" nature of agricultural yield and price risk, which is not captured in the Gaussian copula estimates. In particular, one expects that the imposition of zero tail dependence, as is the case for 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 and 2012 droughts and the 1993 Midwest floods, which caused widespread crop losses.

Hierarchical Unit-level Copulas

The preceding analysis considered D-vine copula estimates and calculated rates for county-level revenue insurance contracts. Such county-level revenue insurance programs have played an important role in the U.S. crop insurance program over the last fifteen years. As noted above, recent changes to U.S. farm policy that include the Agricultural Risk Coverage (ARC) and Supplementary Coverage Option (SCO) will offer a significantly expanded set of county-level revenue insurance options, making accurate pricing of these revenue risks even more pressing. However, the majority of liability in the U.S. crop insurance program exists in the form of individual (farm unit-level) insurance. Farmers currently have the option to insure farm units (generally defined as all land within a section/township/range) individually or as an aggregated group. Aggregation lowers risk, and premium rates are thus adjusted downward to account for this lower risk. The extent to which correlation may exist among yields on individual crop insurance units is an important consideration in the operation of the program and in the reinsurance decisions made by insurers. Until recently, differences in unit structure were priced using assumed, fixed discount rates. For example, units aggregated to form a "basic unit" received a 10% premium rate discount. The RMA recently adopted variable unit discounts; the conceptual and empirical methods used to determine these unit discounts are discussed by Knight et al. (2010).

In addition to high-ordered vine copulas, we also consider a hierarchical copula model that evaluates dependencies and rating at the individual unit level. To this end, we utilize unpublished actual production history (APH) records for individual corn crop insurance units in each of the Illinois counties evaluated above. These data, which are typically proprietary, were made publicly available by the RMA in 1998.[10] The unit-level data include the production histories for each unit for up to 10 years. We utilize unit-level records containing actual yields for units in the counties that had a complete yield history (i.e., 10 years of actual yields). For this analysis, we assume that the unit size remains constant over time.[11]

The model utilizing hierarchical copulas estimates the marginal distributions of

[10] Unfortunately, this is the only publicly available source of unit-level or farm-level data. The histories covered by these yield records largely apply to conventional yield insurance since revenue coverage did not exist over most of the history covered by the yield records. To the extent that more current data on revenue coverage can be obtained from the RMA, this line of research may benefit from an updated consideration of the dependence relationships that are considered here.

[11] We had an overall sample of, respectively, 498, 1,658, 656, and 165 units in Logan, Macon, McLean, and Tazewell counties.

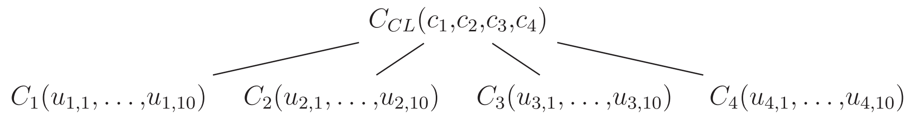

Figure 6. Hierarchical copula model with one county-level copula and four unit-level copulas

yields for policy units, the copulas for joining together policy units in a given county (the unit-level copula), and the copula joining together the unit-level copulas (the county-level copula). Figure 6 illustrates this hierarchical copula, with $C_1, \ldots, C_4$ representing the unit-level copulas for the four counties and C_{CL} representing the county-level copula. For the marginal distributions of the unit-level copulas, we again use rank-based, empirical distributions. The probabilities acquired from the unit-level copulas are used as the marginal distributions in the county-level copula that connects the unit-level copulas of the four counties. We utilize four types of copulas that have previously been discussed: Clayton, Gumbel, Gaussian, and *t*. The Gaussian and *t* copulas have exchangeable correlation matrices; in other words, the correlation is assumed to be the same among all pairs of units for a given county.[12] This assumption is reasonable in that we expect the units to be exposed to relatively homogeneous production conditions within a particular county.

We randomly select ten insurance units from each county. With the ten units, we estimate each of the four types of copula models under consideration here: Clayton, Gumbel, Gaussian, and *t*. This random selection is repeated for 5,000 iterations to estimate the sampling distribution for the parameter of each copula model for each county. These estimated copulas are then used in the county-level copulas of the corresponding copula. Percentile coverage intervals for the replicated parameter estimates of the unit-level and county-level copulas are then constructed, and actuarially fair premium rates are estimated based on the unit-level copulas.

Table 4 summarizes the distributional aspects of the replicated copula estimates. We also present summaries of the fit (AIC and BIC values) and statistical significance of the replicated parameter estimates. For each of the 5,000 iterations, we observe which of the four copulas (Clayton, Gumbel, Gaussian, and *t*) has the lowest AIC and BIC. The frequency in which a certain copula model has the lowest AIC or BIC is recorded in the "AIC" and "BIC" columns of table 4. In all of the counties, the coverage intervals are relatively wide, though the proportions of the 5,000 replicates that have statistically significant parameter estimates is high in nearly every case.[13] For all four counties, the unit-level copulas are best estimated by the *t* copula according to the AIC and BIC values. We also see that the median for the degrees of freedom for the unit-level *t* copulas is low, indicating a substantial degree of platykurtosis. The more restrictive Archimedean copula models receive much less support, particularly in the case of the Gumbel copula. For the county-level copulas, the parameter estimates are higher than the parameter estimates for the unit-level copulas. Therefore, at the county-level, the rank-based correlation is higher than at the unit-level. This is likely due to more noise at the unit-level compared to the county-level. Also, for the county-level *t* copulas, the median estimate for the degrees of freedom is much higher at 13.93, which indicates that the tails of the county-level *t* copulas are slimmer than the tails of the unit-level *t* copulas. Table 4 shows the Clayton copula to be the preferred copula at the county-level according to AIC and BIC. The Clayton copula is an intuitive choice at the county-level because widespread droughts or other natural disasters causing lower yields would lead to lower tail dependence. However, it is important to note that none of the coefficients are significant for the Clayton copula.

We evaluate the premium rates implied by the replicated estimates. Table 5 presents the premium rates for individual units. There is very little variation among the rates derived

[12] Note that the limited span of our data—ten years—necessarily limits the scope of our empirical analysis. Our exchangeability assumption is made necessary by the data limitations.

[13] The null hypotheses for the Clayton, Gaussian, and *t* copulas are parameters equal to zero. However, the null hypothesis for the Gumbel copula is a parameter equal to one. These null hypotheses correspond with a correlation equal to zero.

Table 4. Unit and County–level Copula Estimates Based upon Unit–level Data

Copula	Lower 2.5%	Median	Upper 97.5%	Median df for t	AIC Count	BIC Count	No. Significant
Unit-level Copula Estimates							
Logan							
Clayton	0	1.991	3.04		189	198	4,573
Gumbel	1	1.99	2.57		4	4	4,327
Gaussian	0.68	0.80	0.88		143	165	5,000
t	0.45	0.73	0.89	1.15	4,664	4,633	4,940
Macon							
Clayton	0.66	1.68	2.67		234	249	4,857
Gumbel	1.364	1.85	2.38		9	10	4,575
Gaussian	0.61	0.76	0.85		169	197	5,000
t	0.35	0.68	0.84	1.21	4,588	4,544	4,974
McLean							
Clayton	0	2.11	3.28		466	485	4,743
Gumbel	1	1.98	2.53		0	0	4,547
Gaussian	0.72	0.81	0.88		183	201	5,000
t	0.58	0.77	0.87	1.26	4,351	4,314	4,999
Tazewell							
Clayton	0	1.49	2.81		165	170	4,142
Gumbel	1	1.80	2.52		12	12	3,767
Gaussian	0.65	0.78	0.87		163	178	5,000
t	0.39	0.66	0.85	1.09	4,660	4,640	4,977
County-level Copula Estimates							
Clayton	0	2.18	4.33		2,441	2,454	0
Gumbel	1	3.12	4.35		284	286	299
Gaussian	0.81	0.90	0.95		1,458	1,493	4,996
t	0.65	0.87	0.94	13.93	817	767	4,271

Note: The first four columns present the distribution (at percentiles of 2.5%, 50%, and 97.5%) for the copula parameter estimates obtained from the 5,000 replicated sets of estimates. To conserve space, we only present the median for the *t* copula degrees of freedom. The AIC Count and BIC Count represent the number of occurrences out of 5,000 samples where a given copula (Clayton, Gumbel, Gaussian, or t) has the lowest value for the respective model selection criteria. Likewise, "No. Significant" represents the number of occurrences where the parameter estimates of the copula is statistically significant at the $\alpha = .05$ level.

Table 5. Yield Insurance Premium Rates Based upon Unit–level Copulas

Copula	75% Coverage			95% Coverage		
	Lower 2.5%	Median	Upper 97.5%	Lower 2.5%	Median	Upper 97.5%
Logan						
Clayton	0.0220	0.1733	0.5569	0.0207	0.1593	0.5251
Gumbel	0.0219	0.1731	0.5549	0.0205	0.1594	0.5246
Gaussian	0.0225	0.1724	0.5565	0.0210	0.1591	0.5241
t	0.0225	0.1731	0.5513	0.0207	0.1588	0.5236
Macon						
Clayton	0.0201	0.1751	0.5614	0.0186	0.1612	0.5335
Gumbel	0.0221	0.1748	0.5642	0.0196	0.1608	0.5336
Gaussian	0.0208	0.1750	0.5645	0.0193	0.1604	0.5371
t	0.0204	0.1745	0.5627	0.0191	0.1602	0.5308
McLean						
Clayton	0.0214	0.1738	0.5563	0.0200	0.1599	0.5282
Gumbel	0.0221	0.1748	0.5642	0.0207	0.1611	0.5270
Gaussian	0.0217	0.1745	0.5592	0.0202	0.1600	0.5275
t	0.0216	0.1735	0.5609	0.0205	0.1603	0.5250
Tazewell						
Clayton	0.0229	0.1733	0.5306	0.0211	0.1600	0.5006
Gumbel	0.0228	0.1729	0.5341	0.0212	0.1584	0.5043
Gaussian	0.0227	0.1728	0.5345	0.0217	0.1592	0.5090
t	0.0227	0.1733	0.5352	0.0211	0.1591	0.5076

from the four copula functions for each county at the unit-level. This result is not surprising considering that the correlation among units is relatively weak compared to the correlation at the county-level. In the premium rates determined at the county-level found in table 3, we see much greater variation depending on the model selected. At the county-level, the presence of systemic risk is more pronounced and the observations are less noisy than at the unit-level. Therefore, the presence of tail-dependence can more readily be observed through the premium rates at the county-level.

Conclusions

The federal crop insurance program has been a major fixture of U.S. agricultural policy since the 1930s. The scale of the program continues to grow and a number of new revenue and margin insurance instruments will be introduced as a result of the 2014 Farm Bill. 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 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. The margin insurance plans that currently exist for many livestock products involve combinations of many correlated instruments in deriving the terms of coverage and rates. Dependencies among these individual instruments is an even greater concern here.

We evaluate the suitability of these assumptions by considering a number of alternative copula models. We utilize combinations of pair-wise copulas of conditional distributions to model multiple sources of risk within the framework of a vine copula model. 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 article 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, as well as comparison to the estimates presented here.

Supplementary Material

Supplementary online appendix is available at http://oxfordjournals.org/our_journals/ajae/ online.

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.

Acar, E.F., C. Genest, and J. Nešlehová. 2012. Beyond Simplified Pair-copula Constructions. *Journal of Multivariate Analysis* 110: 74–90.

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. 2010. Truncated and Simplified Regular Vines and their Applications. MS thesis, Technical University Munich.

Brechmann, E.C., and C. Czado. 2013. Risk Management with High-dimensional Vine Copulas: An Analysis of the Euro Stoxx 50. *Statistics & Risk Modeling* 30 (4): 307–342.

Charpentier, A., J. Fermanian, and O. Scaillet. 2007. The Estimation of Copulas: Theory and Practice. In *Copulas: From Theory to Application in Finance*, ed. J. Rank, 35–60. London: Risk Books.

Chen, X., and Y. Fan. 2006. Estimation and Model Selection of Semiparametric Copula-based Multivariate Dynamic Models under Copula Misspecification. *Journal of Econometrics* 135 (1): 125–154.

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

Chvosta, J., D.J. Erdman, and M. Little. 2011. Modeling Financial Risk Factor Correlation with the COPULA Procedure. SAS Global Forum. Paper 340-2011, SAS Institute, Inc.

Czado, C. 2011. The World of Vines. Technical University Munich, Germany.

Dißmann, J.F., E.C. Brechmann, C. Czado, and D. Kurowicka. 2013. Selecting and Estimating Regular Vine Copulae and Application to Financial Returns. *Statistics and Data Analysis* 59: 52–69.

Embrechts, P., A. McNeil, and D. Straumann. 2002. Correlation and Dependence in Risk Management: Properties and Pitfalls. In *Risk Management: Value at Risk and Beyond,* ed. M.A.H. Dempster, 176–223. Cambridge: University Press.

Goodwin, B.K. 2001. Problems with Market Insurance in Agriculture. *American Journal of Agricultural Economics* 83: 643–649.

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 *Distributions with Fixed Marginals and Related Topics*, ed. L. Rüschendorf, B. Schweizer, and M.D. Taylor, 120–141. Hayward: Institute of Mathematical Statistics.

———. 1997. *Multivariate Models and Dependence Concepts.* London: Chapman and Hall.

———. 2005. Asymptotic Efficiency of the Two-stage Estimation Method for Copula-based Models. *Journal of Multivariate Analysis* 94 (2): 401–419.

Jondeau, E., and M. Rockinger. 2006. The Copula-GARCH Model of Conditional Dependencies: An International Stock-market Application. *Journal of International Money and Finance* 25: 827–853.

Knight, T.O., K.H. Coble, B.K. Goodwin, R.M. Rejesus, and S. Seo. 2010. Developing Variable Unit-structure Premium Rate Differentials in Crop Insurance. *American Journal of Agricultural Economics* 92 (1): 141–151.

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

Mildenhall, S.J. 2006. Correlation and Aggregate Loss Distributions with an Emphasis on the Iman-Conover Method. Casualty Actuarial Society Forum.

Miranda, M., and J. Glauber. 1997. Systemic Risks, Reinsurance, and the Failure of Crop Insurance Markets. *American Journal of Agricultural Economics* 79: 206–215.

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

Oh, D.H., and A.J. Patton. 2012. Modelling Dependence in High Dimensions with Factor Copulas. Working paper, Department of Statistics, Duke University.

Patton, A.J. 2006. Modelling Asymmetric Exchange Rate Dependence. *International Economic Review* 47: 527–556.

Risk Management Agency. 2014. Federal Crop Insurance Corporation: Nationwide Summary – by Crop. Washington, DC.

Rodriguez, J.C. 2003. Measuring Financial Contagion: A Copula Approach. *Journal of Empirical Finance* 14: 401–423.

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

Schepsmeier, U., J. Stoeber, E.C. Brechmann, and B. Graeler. 2013. Package VineCopula: Statistical Inference of Vine Copulas. *R*-Project CRAN Repository.

Sklar, A. 1959. Distribution Functions in n Dimensions and Their Margins. *Statistics Publications, University of Paris* 8: 229–231.

Smith, V. 2011. Premium Payments: Why Crop Insurance Costs Too Much. In

American Boondoggle — Fixing the 2012 Farm Bill. American Enterprise Institute, Washington, DC.

Vuong, Q.H. 1989. Likelihood Ratio Tests for Model Selection and Non-nested Hypothesis. *Econometrica* 57 (2): 307–333.

Yan, J. 2007. Enjoy the Joy of Copulas: With a Package "copula". *Journal of Statistical Software* 21 (4): 1–21.

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