

IMPACTS OF COPULA MODELING AND PARAMETRIC VARIATION
ON REVENUE POLICY PREMIUM RATES IN
MULTIPLE PERIL CROP INSURANCE

by

Seth Neil Simonds

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Applied Economics

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2015

©COPYRIGHT

by

Seth Neil Simonds

2015

All Rights Reserved

ACKNOWLEDGMENTS

I would like to thank my committee and co-chairs. Dr. Smith allowed me to work on this topic and provided the opportunity to work with Dr. Atwood, which has been a transformative experience. His legacy will continue well into the future through the snippets of code he so generously allowed me to use and more generously spent time explaining. Dr. Bekkerman's efforts are self evident anywhere the thoughts of this thesis are communicated clearly and concisely.

Jane Boyd and Tamara Moe operate behind the scenes, taking care of the details that allowed me to focus on my work. Thank you to the economics department and Dr. Rucker for funding me and providing me the opportunity to attend this program: I have learned so much about learning, work, the intelligence of others and myself.

My father offered innumerable research topics and was always willing to talk cowboy-philosophy and basketball. My mother allowed me to hear my home, when I needed that tune the most. Frances Wright spent hours listening, thank you. Finally, I could not have accomplished this project and degree without the cohort of individuals that I entered this program with.

TABLE OF CONTENTS

1. INTRODUCTION	1
2. HISTORY OF CROP INSURANCE	5
3. RATEMAKING CONCEPTS, POLICIES AND PRACTICES	11
Concepts	11
Policies	14
Actual Production History	15
Traditional Revenue Protection	16
Upside Revenue Protection.....	17
Ratemaking Practices	18
Summary	19
4. CRITICISM OF RATEMAKING PRACTICES	21
Moral Hazard	22
Adverse Selection	23
Indirect Evidence of Adverse Selection.....	24
Direct Estimation of Adverse Selection	26
Summary	35
5. COPULAS	37
Major Copula Families and Members of Interest	38
Elliptical Copulas	39
Archimedean Copulas.....	40
Rotated Copula Representations	42
Theoretical Price-Yield Representations of Copulas	44
6. COPULA GOODNESS OF FIT STATISTICS PERFORMANCE	46
Design	46
Results	49
Summary	51
7. EMPIRICAL PREMIUM RATE DISPERSION BETWEEN COPULAS	52
Ratemaking Procedures.....	52
The Pure Empirical Method	53
General Copula Ratemaking	54
Parametric Copula Ratemaking.....	56

TABLE OF CONTENTS – CONTINUED

Empirical Copula Ratemaking	57
Revenue Product Formulas.....	57
Summary.....	58
Data for Empirical Experiments	59
Prices	60
Yields	62
Price and Yield Dependence	65
Design	66
Results	67
Copula Compared to Pure Empirical.....	67
Dispersion Between Copulas	77
Deviations Between Copulas	81
Summary	85
8. SIMULATED PREMIUM RATE DISPERSION BETWEEN COPULAS.....	88
Design and Data Generation.....	89
Results	90
Dispersion Between Copulas	90
Deviation Between Copulas	93
Summary	97
9. SIMULATED PREMIUM RATE DISPERSION WITHIN COPULAS.....	99
Design and Data Generation.....	100
Results	101
True and Sample Premium Rates	102
Parameter Variation Causing Dispersion	118
Summary	122
10. AKAIKE INFORMATION CRITERION PERFORMANCE	124
Design and Data Generation.....	125
Results	126
Summary	127
11. COPULA SELECTION STRATEGY APPLIED	129
Data.....	129
Design	130
Results	132
Empirical Dispersion Within Copulas	132

TABLE OF CONTENTS – CONTINUED

Akaike Information Criterion Selection	140
Copula Selection Premium Rates	145
Summary	155
12. CONCLUSION	157
REFERENCES CITED	161
APPENDICES	169
APPENDIX A: Tables	170
APPENDIX B: Figures	179
APPENDIX C: Taxonomy of Crop Insurance Policies	182
APPENDIX D: Current Ratemaking Practices	185
APPENDIX E: Gumbel and Clayton Copula Tail Dependence.....	193
APPENDIX F: Itô's Process for Prices	195
APPENDIX G: Yield Trends and Heteroskedasticity	200
APPENDIX H: Motivation for Deviation Measure	203
APPENDIX I: Goodness-of-Fit Studies.....	206
APPENDIX J: Corn as Special Case of Simulations	210

LIST OF TABLES

Table	Page
1. Summary of Copula Characteristics	42
2. Percent Rejection of Null Hypothesis with $n = 50$	50
3. Futures Price Data and Deviations from 1969 to 2013.....	61
4. Detrended County Yields (Bu./Ac.) from 1945 to 2013	64
5. Summary of County and Farm Level Corn Yields.....	65
6. Empirical County Yield and Price Dependence	66
7. Empirical Revenue Policy Premium Rates	68
8. LRM of RP Premium Rates	70
9. LRM of RPHPE Premium Rates	71
10. Within Copula True Premium Rate Dispersion.....	103
11. Within Copula Sample Premium Rate Dispersion	104
12. Sample Parameter Deviation Compared to True Parameter.....	120
13. Dispersion of Kendall's τ Compared to True Parameter.....	121
14. Percent Identification Akaike Information Criterion	127
15. AIC Copula Selection for Empirical Study	141
16. Macon, IL AIC and Normal Copula Premium Rates	146
17. McDonough, IL AIC and Normal Copula Premium Rates	147
18. Hancock, IL AIC and Normal Copula Premium Rates.....	148
19. McLean, IL AIC and Normal Copula Premium Rates	149

LIST OF TABLES – CONTINUED

Table	Page
20. Tazwell, IL AIC and Normal Copula Premium Rates	150
21. Logan, IL AIC and Normal Copula Premium Rates	151
22. Delaware, IA AIC and Normal Copula Premium Rates	152
23. Cass, IA AIC and Normal Copula Premium Rates	153
24. Lucas, IA AIC and Normal Copula Premium Rates	154
25. Abbreviations	171
26. Experimental Design for Goodness-of-Fit Studies	172
27. Percent Rejection of Null Hypothesis with $n = 75$	173
28. Percent Rejection of Null Hypothesis with $n = 150$	174
29. Copula Versus Pure Empirical RP Premium Rates	175
30. Copula Versus Pure Empirical RPHPE Premium Rates	176
31. Characteristics of Selected Counties	177
32. Number of Observations for Empirical Statistics	178

LIST OF FIGURES

Figure	Page
1. Contours of Copula Family Densities	41
2. Rotated Copula Representations.....	43
3. State Level Corn Yield Trends from 1925-2013	63
4. Empirical Revenue Premium Rate Pairs	73
5. Copula and Pure Empirical RP Premium Rates.....	75
6. Copula and Pure Empirical RPHPE Premium Rates	76
7. Empirical RP Premium Rate Dispersion	79
8. Empirical RPHPE Premium Rate Dispersion	80
9. Between Copulas Copula RP Deviations	83
10. Between Copulas Copula RPHPE Deviations	84
11. Between Copulas Simulated RP Dispersion	91
12. Between Copulas Simulated RPHPE Dispersion.....	92
13. Between Copulas Simulated RP Deviations.....	95
14. Between Copulas Simulated RPHPE Deviations	96
15. True RP Dispersion Within Copulas 50% Coverage	105
16. Sample RP Dispersion Within Copulas 50% Coverage.....	106
17. True RP Dispersion Within Copulas 65% Coverage	107
18. Sample RP Dispersion Within Copulas 65% Coverage.....	108
19. True RP Dispersion Within Copulas 85% Coverage	109

LIST OF FIGURES– CONTINUED

Figure	Page
20. Sample RP Dispersion Within Copulas 85% Coverage	110
21. True RPHPE Dispersion Within Copulas 50% Coverage	111
22. Sample RPHPE Dispersion Within Copulas 50% Coverage	112
23. True RPHPE Dispersion Within Copulas 65% Coverage	113
24. Sample RPHPE Dispersion Within Copulas 65% Coverage	114
25. True RPHPE Dispersion Within Copulas 85% Coverage	115
26. Sample RPHPE Dispersion Within Copulas 85% Coverage	116
27. Empirical RP Deviations 50% Coverage	134
28. Empirical RPHPE Deviations 50% Coverage	135
29. Empirical RP Deviations 65% Coverage Level	136
30. Empirical RPHPE Deviations 65% Coverage Level	137
31. Empirical RP Deviations 85% Coverage	138
32. Empirical RPHPE Deviations 85% Coverage	139
33. Density and Q-Q plot for Corn Futures Price Data.....	180
34. State Level Detrended Corn Yields from 1945-2013	181
35. Tail Dependence of Clayton and Gumbel Copulas	194
36. Motivation for Deviation Metric	205
37. Simulated and Empirical Corn RP Dispersion	212
38. Simulated and Empirical Corn RPHPE Dispersion	213

ABSTRACT

Federally subsidized multiple peril crop insurance is the primary mechanism by which U.S. farmers receive public income. This study investigates the role of copula modeling in developing revenue product premium rates for multiple peril crop insurance. Simulation and empirical experiments are used to examine the viability of a ratemaking practice that relies on an assumed Normal copula. This study shows that the assumption of a copula cannot be statistically justified and that premium rates generated within copulas and between alternative copulas can diverge as a function of the marginal price and yield distributions, their relationship and the level of protection a producer elects. The current ratemaking practice does not account for the imprecision of premium rates implicit to a copula based approach. A copula selection method is proposed and examined in order to reduce premium rate imprecision resulting from copula misspecification. A non-copula based ratemaking method may better meet the overt policy objectives of multiple peril crop insurance.

CHAPTER ONE

INTRODUCTION

Subsidized multiple peril crop insurance (MPCI) is a major risk management tool for U.S. agriculture producers.¹ Over time, subsidies have increased to encourage participation and reduce adverse selection, which occurs if premium rate inaccuracies allow higher risk participants to acquire coverage at lower than actuarially fair prices. Since 1983, the program has expanded the range of counties and crops protected, as well as creating revenue product that protects farmers against shortfalls in prices and yields. Revenue products are the most widely purchased MPCI policies and are therefore the primary mechanism by which public wealth is transferred to individual farmers.

The Risk Management Agency (RMA) oversees the development and administration of MPCI products and must develop insurance policies and associated premium rates that accurately represent producers' probability of loss in order to prevent the potential for adverse behavioral responses endemic to insurance markets. An obstacle to this process is the limited number of historical price and yield observations upon which probabilities of loss and therefore premium rates can be estimated. To overcome this constraint, the RMA uses a simulation process by which price and yield realizations are generated from assumed marginal distributions. Revenue product premium rates are generated from these joint

¹Commonly used abbreviations are listed in Table 25 in Appendix A.

outcomes and attempt to reflect the county and crop specific inverse relationship between prices and yields.

Much of the criticism of the MPCCI ratemaking procedures contends that the technical innovations meant to overcome limited data do not. The current ratemaking procedure for revenue policies uses the equivalent of a Normal copula to bind price and yield realizations together in a manner that reflects the specific price and yield relationship of a county and crop. Copulas are mathematical structures that allow marginal distributions to be joined into a multivariate distribution. Copulas impose their multivariate shapes and the county-crop price-yield dependence, while retaining the unique properties of the marginal distributions. Because there are an infinite number of copulas and available techniques cannot discriminate between them at small sample size and when dependence is weak, the selection of any particular copula to use in the MPCCI ratemaking process is subject to criticism. Consequently, this raises the question of whether the RMA's current Normal copula based ratemaking process, appropriately represents the loss probabilities or whether alternative modeling strategies can estimate more accurate premium rates.

This study investigates the viability and traces the implications of using alternative copula functions to generate premium rates for multiple peril crop insurance (MPCCI) revenue policies and investigates a possible copula selection method through six simulation and empirical experiments. The first experiment is a large scale Monte-Carlo simulation that explores the finite sample performance of

available copula goodness-of-fit tests, which test the hypothesis that a given set of data come from a hypothesized copula at the small sample sizes and weak price-yield dependence, which characterize MPCCI data. Next, we use an empirical experiment to investigate whether premium rates vary for the two major revenue policies Revenue Protection (RP), which offers upside price protection and Revenue Protection with Harvest Price Exclusion (RPHPE) for corn in Illinois, Indiana and Iowa, the three largest corn producing states. Premium rates for both policies are generated by alternative parametric copulas, as well as an empirical copula and a non-copula based pure empirical rating methods. We then use a simulation experiment to investigate the potential for dispersion between copula generated premium rates when prices and yields exhibit greater variability than in the available empirical data. Using simulation allows the variation in revenues to be precisely identified and the results extended beyond corn.

The first set of experiments investigate the potential for dispersion between premium rates generated by alternative copulas, but do not offer insights about the potential for premium rate dispersion within copulas. To examine this issue, we first use a simulation experiment and resampling process to investigate the potential for premium rate dispersion within copulas as function of the factors that determine the premium rate level. Next we examine the finite sample performance of the Akaike Information Criterion (AIC), which provides one potential instrument by which an optimal copula can be selected based on its data fit performance. Finally, we use the AIC selection process to identify a unique copula for a subset of counties that

represent unique empirical yield and dependence characteristics and compared the premium rates generated by the AIC selected copula to those generated by the RMA's ratemaking process which relies on a Normal copula.

We find that the use of a particular copula for all ratemaking purposes, including the Normal copula, cannot be supported on statistical grounds, but a unique county-crop copula could be selected using the Aikake Information Criterion (AIC). Also, the potential for premium rates to diverge both when generated by alternative copulas and generated within a copula is a function of the factors that determine the premium rate level: coverage level, dispersion of the price and yield marginals and the strength of their relationship. An AIC based copula selection methodology indicates that the Normal copula is likely a sub-optimal alternative for many crops and counties across the MPCCI book of business. In light of the implications of copula modeling, a non-copula based alternative may be more desirable. Finally, the research reinforces previous studies' findings that the accuracy of copula models is dependent on accurately identifying the price and yield marginal distributions.

CHAPTER TWO

HISTORY OF CROP INSURANCE

Over its history, multiple peril crop insurance (MPCI) in the United States has expanded in the diversity of policies offered, the extent of crops and counties covered and subsidization of premium rates; ratemaking practices have become more complex in an effort to overcome small sample size. Currently, the federal MPCI program provides more subsidies to farmers than any other agricultural policy provisioned by the 2014 Farm Bill (Smith and Glauber, 2012). In characterizing the evolution of the program for the period between 1928 and 1994, Goodwin and Smith (1995) demarcate its history into four periods: 1938–1944, 1945–1973, 1973–1980, and 1981–1994; legislation in 1994 creates another natural break until 2000 and from then until the present day represents a final period in the history of the program.

A series of legislation leading to more active federal support of the agricultural sector, was begun with the 1933 establishment of the Commodity Credit Corporation, which provided some commodities price support (Kramer, 1983). Two droughts in the late 1930s and the political environment of the time increased interest in protecting farmers from low yields (Goodwin and Smith, 1995). These conditions created the context for the passage of the Agricultural Adjustments Act of 1938 and the Federal Crop Insurance Act (FCIA) of 1937, which offered yield based insurance for wheat farmers (Kramer, 1983). Globally there was no experience with government-provided MPCI programs. Hence this

legislation “represented a bold political and economic experiment” (Kramer, 1983, p. 181). In addition, the FCIA created the Federal Crop Insurance Corporation (FCIC), an agency of the United States Department of Agriculture (USDA), tasked with administering the insurance program (Kramer, 1983).

In part, the novel component of the FCIA was the government’s express interest in offering MPCCI. Named or single peril insurance had been offered by private companies for fire and hail prior to the 1930s and up until the present day, but private attempts to offer MPCCI had all ended in failure (Kramer, 1983). Throughout its history, critics of MPCCI have noted: 1) the failure of the private sector to successfully offer MPCCI, 2) the lack of any sound argument for market failure that would warrant government intervention, 3) the role systemic risks, 4) insufficient data for accurate rate-setting and 5) high monitoring costs associated with reducing adverse selection and moral hazard (Goodwin and Smith, 1995). Throughout its history “these criticisms fell on politically deaf ears” (Goodwin and Smith, 1995, p. 37). Over the last fifty years MPCCI has evolved by nominally addressing most of these criticisms through technical innovation and primarily addressing the more fundamental issues by drastically increasing participation through increased subsidization (Smith and Glauber, 2012).

From 1938–1944, the program was small both in its geographic scope and crops that it covered. The program was subject to moral hazard, premium rates were inaccurately rated on insufficient historical data and the insurance pool was comprised of a adversely selected producers who were allowed to let their policy

lapse upon realization of a non-indemnifiable yield outcome (Goodwin and Smith, 1995). Cotton was added to wheat in 1942, but the crop suffered such severe losses that the entire insurance program was discontinued in 1944 (Kramer, 1983).

The program did not remain dormant for long. An amendment to the Agricultural Appropriation Act reestablished the program later in 1944 (Kramer, 1983). The 1944 amendment authorized programs for several new commodities and stipulated a maximum number of counties and commodities that could be added each year (Kramer, 1983). The early half of the period from 1945–1973 was one of slow expansion in the program’s coverage, punctuated by several innovations and decisions. In 1946, a three year contract was developed to address adverse selection. In 1954, local field agents began to write and service policies. In 1956, policies were no longer sold in high risk marginal locations. Throughout the 1960s, the FCIC focused on increasing coverage, with the costs of the program increasing commensurately (Kramer, 1983). The FCIC came under new management and in 1969 a review cited the FCIC’s poor financial performance (Kramer, 1983). As a result of the use of short term yield trends for risk assessment, premiums had been set too low. In response adjustments were made to several programs, while the potato program was eliminated altogether (Kramer, 1983).

In 1970 a task force criticized the program for establishing premiums based on county level data and recommended that policies be developed based on farm level data instead (Kramer, 1983). After the Government Accounting Office released a report in 1977 concurring with the 1970 findings, the FCIC introduced

individualized crop insurance in 1978 (Kramer, 1983). Individual production based policies remain to this day; however, the period from 1973–1980 is primarily delineated by the interaction between the crop insurance program and disaster relief payments authorized by The Agriculture and Consumer Protection Act of 1973, the Rice Production Act of 1975 and The Food and Agriculture Act of 1977 (Goodwin and Smith, 1995; Kramer, 1983).

The rising costs of *ad-hoc* disaster relief programs became politically untenable and led to the 1980 Federal Crop Insurance Act, which expanded the scope of the crop insurance program at the expense of disaster aid (Goodwin and Smith, 1995). The bill established crop insurance as the major form of disaster protection by removing limits on annual expansion and allowing the FCIC to insure any commodity with sufficient actuarial data to base premium rates. In addition, the act subsidized premiums and stipulated that private insurance companies could sell crop insurance policies and develop their own MPCCI policies with government support (Kramer, 1983).

The period from 1981–1994 was one of expansion in participation because of the program’s increased geographic and commodity coverage as well as increased subsidization of premiums. Continuing the move towards more individualized programs and away from setting premium rates based upon county level data, in 1983 a new methodology for setting premium rates based on actual production history (*APH*) was implemented (Goodwin and Smith, 1995). During this time private insurance companies became more involved in marketing the program, not

only marketing the MPCCI policies, but servicing them as well under private reinsurance contracts (Goodwin and Smith, 1995). In 1990, in response to large loss ratios,² the FCIC became far more concerned with addressing tax payer costs (Goodwin and Smith, 1995). Therefore, premium rates were increased, but so too were subsidies and the method for assessing an individual farmer's *APH* was modified (Goodwin and Smith, 1995). Also, farms that “habitually” suffered losses were subject to separate underwriting procedures (Goodwin and Smith, 1995). The 1990 Food, Agriculture, Conservation and Trade Act also specified that new insurance contracts were to be developed in order to reduce loss ratios and subsidy costs. To this end, Area-Yield Crop Insurance contracts were introduced (Goodwin and Smith, 1995).

From 1995–2000, the program continued its rapid expansion as a result of increased subsidization and several new policies, including the now dominate revenue insurance (Glauber, 2012). The Crop Insurance Reform Act of 1994 required farmers to participate in catastrophic protection (CAT) insurance plans in order to participate in any other farm programs, as a result participation in MPCCI increased dramatically the following year (Glauber, 2012). In addition to fully subsidizing the CAT plans, subsidies were offered for any plan that offered higher levels of coverage (Glauber, 2012). Compulsory participation in CAT was removed in 1996, but farmers’ election of higher coverage levels increased (Glauber, 2012). In

²A loss ratio is the amount paid by an insurer divided by the amount of premiums collected. Loss ratios in excess of one indicate that premiums collected are insufficient to cover the costs of providing insurance.

1996, the Risk Management Agency (RMA) was established to administer MPCCI (USDA, 2014). The defining aspect of this period was the implementation in 1996 of revenue-based insurance plans, which have become the dominant policies selected by farmers to insure themselves (Glauber, 2012). These policies protect farmers from shortfalls in revenue caused by shortfalls in yields, prices or both. By 2013 three-quarters of all policies sold were revenue based and on a premium basis 83% were revenue based (Shields, 2012).

The most recent period, from 2000-2014, is marked by the passage of the Agricultural Risk Protection Act of 2000, which increased subsidization of the MPCCI program (Coble and Barnett, 2013). Then, despite the intent of the 1980 Act to replace disaster payments with crop insurance, the 2008 Farm Bill instituted two new disaster aid programs to supplement farmers' crop insurance payments and mitigate "shallow losses" (Glauber, 2012). In 2007, the relatively small coverage of the area yield program introduced in the early 1990s was expanded by the development of index-based policies (Glauber, 2012).

Much of history of MPCCI can be attributed to increased subsidization, expansion in scope, and innovation to overcome fundamentally limited data. The cost of each is born by the tax paying public. Revenue products represent the most total liability for MPCCI and this project focuses on the accuracy with which these premium rates are set because they are by far the most important mechanism by which public wealth is transferred to individual farmers.

CHAPTER THREE

RATEMAKING CONCEPTS, POLICIES AND PRACTICES

Multiple peril crop insurance (MPCI) policies that insure a farmer's revenue from short falls in yields, prices or both have been available since 1996. Revenue policies account for most of the crop insurance program's liability. In this chapter the major concepts used to estimate insurance premium rates are presented in the context of crop insurance. With these concepts, the formulas and ratemaking³ practices used to develop premium rates for revenue policies are then presented. In the following chapter, the scholarship analyzing the effectiveness of these practices is presented.

Concepts

Insurance allows one party to transfer risk to another party.⁴ The cost of this service is represented by a premium rate, which primarily represents the insurance writer's estimation of the expected value of the insured's future costs; the process of estimating this cost is called ratemaking (Society, 1988). Two critical concepts in ratemaking are liability and indemnity. For MPCI, liability ($L_{i,t}$) is the value of the insured crop and is expressed as:

$$L_{i,t} = AP_{i,t} \times \mathbb{E}(y_{i,t}) \times cvg \times \mathbb{E}(p_t) \quad (1)$$

³Ratemaking is the procedures used to develop premium rates for insurance policies.

⁴Much of this section is a summary of a report on MPCI performed by Josephson, Lord, and Mitchell (2000).

where $AP_{i,t}$ is the number of acres planted, and cvg is the selected coverage level, for producer i in insurance period t . Both the acres planted and coverage level are known as the producer either reports or selects them. However, the expected yield and expected price— $\mathbb{E}(y_{i,t})$ or APH^5 and $\mathbb{E}(p_t)$ or Projected Price (PP)—must be estimated. The estimation of PP at the time of harvest is based on futures contract data. The estimation of producer expected yields is based upon the producer's actual production history.

An indemnity is paid by the insurer to the insured, when a covered loss occurs, which is when the realized value of production ($VP_{i,t}$) at harvest time is less than the purchased $L_{i,t}$:

$$VP_{i,t} = AP_{i,t} \times y_{i,t} \times p_t \quad (2)$$

The harvest price p_t (HP) is also based on the futures market, while the producer reports the realized yields for a crop ($y_{i,t}$). The amount of the indemnity payment ($I_{i,t}$) is equal to $L_{i,t}$ minus $VP_{i,t}$:

$$I_{i,t} = \text{Max}(0, L_{i,t} - VP_{i,t}) \quad (3)$$

The price of this protection, part of which is paid by the federal government, is the

⁵Here the idea of expected yields is analogous to the concept of basing a producers premium rate on their actual production history.

premium ($R_{i,t}$) and is:

$$R_{i,t} = L_{i,t} \times \mathbb{E}(LCR_i) \quad (4)$$

An actuarially fair premium rests upon the accurate estimation of the appropriate pure-risk premium rate or $\mathbb{E}(LCR_i)$. The individual Lost Cost Ratio (LCR), which is the ratio of indemnity to liability, represents the insurer's estimation of expected future costs and identifies this premium rate:

$$LCR_{i,t} = \frac{I_{i,t}}{L_{i,t}} \quad (5)$$

The expected LCR ($\mathbb{E}(LCR_i)$) can then be calculated over the historical outcomes for the individual producer to generate their pure-risk premium rate:

$$\mathbb{E}(LCR_i) = \sum_t \frac{I_{i,t}}{L_{i,t}} \quad (6)$$

While indemnity is strictly identified upon the realization of producer's value of production, it also depends on the estimation of an appropriate liability.

Typically, an additional load is added to the $\mathbb{E}(LCR_i)$ or "pure risk premium rate" to account for the uncertainty in its estimation; however, these costs are separately provisioned for in crop insurance. The key to accurately estimating liability is the credibility or predictive power of the body of data upon which loss experience is identified (Society, 1988). However, the actuarial concept of credibility involves two

factors: premium rates are improved by increasing the data or by grouping the risk experience of producers into more homogeneous or similar units. The sample size of historical liability can be increased by aggregating producers into larger groups; however, this comes at the inherent expense of generating premium rates over a more heterogeneous pool.

These consideration are intersected with the heterogeneous nature of crop yields across producers and limited historical yield data. Many of the procedures in place to rate MPCCI arise from the ratemaking attempt to improve credibility by compromising between increasing sample size and maintaining the homogeneity of groups. At the heart of this issue is the estimation of $\mathbb{E}(y_{i,t})$ (i.e *APH*). To the extent that liability and therefore premium rates are inaccurately determined, the program exposes itself to adverse selection; like any insurance program monitoring procedures must be in place to check the incentive for moral hazard as well. Before addressing the criticism of MPCCI ratemaking procedures, the structure of the three major MPCCI products will be presented.

Policies

In 2014, more than two million policies were sold. With a total of 100 billion dollars in liability, 10 billion dollars in premiums paid by both the federal government and farmers and almost 300 million acres insured.⁶ While many policies cover a range of commodities and locations, there are two major types of insurance

⁶All information regarding the fiscal performance of MPCCI can be obtained at <http://www.rma.usda.gov/data/sob.html>.

products: yield or revenue. These products either cover losses based upon a producer's experience or the experience of a county or index. Most policies are sold to producers insuring against individual revenue short falls based on their own farms yield outcomes. Revenue insurance is important because it dominates the MPCI book of business; however, yield insurance, which the Risk Management Agency (RMA) calls APH, is important because the RMA elects to base revenue premium rates on the estimation of an APH yield premium rate. Two revenue policies are widely utilized by farmers: a traditional revenue product, called Revenue Protection with Harvest Price Exclusion, denoted RPHPE throughout this paper, and a more exotic policy, which the RMA calls Revenue Protection, but provides additional upside price protection and is denoted RP throughout this paper.⁷ RP is the most frequently selected policy, representing the most liability (Shields, 2012).

Actual Production History

For the 2014 crop year almost 12% of the liability in the federal crop insurance program was for Actual Production History (APH) yield contracts. The RMA also uses the APH premium rate as a base upon which revenue premium rates are adjusted. For each region and crop the projected price (PP) is calculated over a specified price discovery period and is the average settlement price for the commensurate commodity contract, if an active futures market exists as outlined in the Commodity Exchange Price Provisions (CEPP). An indemnity is triggered

⁷For a more detailed description of the taxonomy of MPCI, see Appendix C.

($T_{APH} \equiv 1$) at the end of the growing season when the realized yield ($y_{i,t}$) is less than the expected yield ($\mathbb{E}(y_{i,t})$) multiplied by the selected coverage level (cvg):

$$T_{APH} = \begin{cases} 1 & \text{if } y_{i,t} < \mathbb{E}(y_{i,t}) \times cvg \\ 0 & \text{if } y_{i,t} \geq \mathbb{E}(y_{i,t}) \times cvg \end{cases} \quad (7)$$

An indemnity is only paid on the basis of a yield short fall and the value of production is based upon the previously established Projected Price.

Traditional Revenue Protection

While a yield short fall may cause a fall in revenue, it is possible that yields do not fall, but prices do fall causing a shortfall in expected revenue per acre. Revenue products allow producers to protect themselves from a short-fall in revenue from either a price or yield short fall or both.⁸ For the 2014 crop year about 7% of all liability was for Revenue Products with Harvest Price Exclusion (RPHPE) policies. Insuring based upon revenue requires a similar projection and realization process for prices as defined for the APH yield policy. While the projected price ($\mathbb{E}(p_t)$) is already identified, a harvest price (p_t) must now be recorded. The harvest price is the average of the closing futures prices for the period around the time of harvest for a particular region and crop. An indemnity is triggered ($T_{RPHPE} \equiv 1$) as

⁸For most producers yields and prices have a negative relationship that varies by producer in strength. This natural hedge between prices and yields makes simultaneous shortfalls in both prices and yields less likely.

a function of both the harvest price and realized yield ($y_{i,t}$):

$$T_{RPHPE} = \begin{cases} 1 & \text{if } (y_{i,t} \times p_t) < \mathbb{E}(y_{i,t}) \times \mathbb{E}(p_t) \times cvg \\ 0 & \text{if } (y_{i,t} \times p_t) \geq \mathbb{E}(y_{i,t}) \times \mathbb{E}(p_t) \times cvg \end{cases} \quad (8)$$

Upside Revenue Protection

Revenue Protection (RP), which offers upside price protection, allows the producer to be indemnified at either the higher of the the harvest price (p_t) or the projected price ($\mathbb{E}(p_t)$). For 2014, 73% of the total program liability for MPCl was associated with RP policies. While producers have found the terms attractive, the potential to receive an indemnity based on an unrealized variable at sign-up time, implies that the liability for these programs cannot be known *ex-ante*.⁹ Instead potential liability is conditional upon two possible states:

$$L_{RP} = \begin{cases} AP_{i,t} \times (\mathbb{E}(y_{i,t}) \times cvg \times \mathbb{E}(p_t)) & \text{if } \mathbb{E}(p_t) > p_t \\ AP_{i,t} \times (\mathbb{E}(y_{i,t}) \times cvg \times p_t) & \text{if } p_t > \mathbb{E}(p_t) \end{cases} \quad (9)$$

The trigger for an indemnity payment becomes more complicated too:

$$T_{RP} = \begin{cases} 1 & \text{if } y_{i,t} \times \text{Max}(p_t, \mathbb{E}(p_t)) < (\mathbb{E}(y_{i,t}) \times \mathbb{E}(p_t)) \times cvg \\ 0 & \text{if } y_{i,t} \times \text{Max}(p_t, \mathbb{E}(p_t)) \geq (\mathbb{E}(y_{i,t}) \times \mathbb{E}(p_t)) \times cvg \end{cases} \quad (10)$$

⁹The RMA caps the liability from *HP* movements at two times the *PP*.

Ratemaking Practices

Ideally, the liability and indemnity at any coverage level would be calculated over the historical range of realizations and an $\mathbb{E}(LCR_i)$ would be estimated for each producer. However, limitations in available producer level data make this an impractical procedure to base premium rates. The RMA attempts to overcome the lack of credibility from small sample size through several procedures that augment producer data and the common liability and indemnity experience that premium rates are based upon. The most significant of these practices, is the use of assumed marginal distributions to generate price and yield realizations. The RMA uses the equivalent of a Normal copula to bind these independently simulated realizations together and estimate a particular county's probability of experiencing a revenue short fall.¹⁰

The RMA's lost cost ratio formulas for the actual production history yield (APH), Revenue Protection with Harvest Price Exclusion (RPHPE) and Revenue Protection (RP) policies are presented:

$$\begin{aligned}
 APH &= \sum_{i=1}^{500} \frac{Max(0, cvg*Y - Max(0, y_i))}{500*Y*cvg} \\
 RPHPE &= \sum_{i=1}^{500} \frac{Max(0, cvg*Y*P - Max(0, y_i)*Min(2*P, H_i))}{500*Y*cvg*P} \\
 RP &= \sum_{i=1}^{500} \frac{Max(0, cvg*Y*Min(2*P, Max(P, H_i)) - Max(0, y_i*Min(2*P, H_i)))}{500*Y*cvg*P}
 \end{aligned} \tag{11}$$

where Y is the producer's $\mathbb{E}(y_i)$, P is the projected price and H_i is a simulated

¹⁰For a more detailed description of the RMA's ratemaking practices refer to Appendix D.

harvest price (Coble et al., 2010).

Prices are assumed to be log-normally distributed, while yields are assumed to come from a standard normal distribution censored at zero. The central tendency of the price distribution is identified by the projected price and the dispersion by the implied volatility calculated from options contract data. The central tendency of the yield distribution is the farmer's normalized actual production history and the dispersion is a normalized measure of yield standard deviation. The simulated realizations are bound together using the Iman-Conover method, which under specific assumptions, is the equivalent of a Normal copula (Hungerford and Goodwin, 2014). The Iman-Conover method imposes the estimated linear correlation coefficient of the producer's county's yield and price dependence on the simulated realizations. Premium rates are estimated using the formulas above applied over the simulated price and yield realizations.¹¹

Summary

It is important to understand the rating process used by the RMA for three reasons. First, many of the procedures the RMA uses reflect an attempt to increase credibility along its countervailing dimensions. To the extent that these procedures affect credibility, they affect the accuracy of premium rates and therefore, the potential for adverse selection and invite criticism along these lines. Second, to the extent that premium rates are misspecified, the losses of the crop insurance program

¹¹This is a summary of the review performed by Coble et al. (2010); there are many additional steps and intricacies, which are outlined in more detail in Appendix D.

are more likely to increase, with tax-payer funds accounting for the difference. This study addresses the accuracy of the procedure used to estimate a producers expected loss cost ratio, which is the most fundamental ratemaking practice determining tax-payer costs. Third, the rating procedure used in this study is similar, but not identical to the RMA's process; therefore, the methods used in this paper and results are better understood in light of its similarities and differences to current practices.

CHAPTER FOUR

CRITICISM OF RATEMAKING PRACTICES

An extensive body of research has examined the lack of demand for crop insurance (Knight and Coble, 1997). The two dominant economic phenomena used to explain small inefficiencies or catastrophic failings of insurance markets are moral hazard and adverse selection. Crop insurance is no different. Moral hazard and adverse selection arise from asymmetric information (Knight and Coble, 1997).

When insured individuals behave differently because they possess insurance and the insurance provider cannot monitor their behavior, then moral hazard is in play. In practice, firms that offer insurance are constantly trying to ensure that the contracts they offer do not induce or allow for the insured to change their behavior and increase expected indemnities. In general, the crop insurance literature finds that moral hazard is present and that it contributes to the poor actuarial performance of MPCCI programs.

Adverse selection occurs because the insured generally knows more about the profile of their risk, than the insurance provider. The crop insurance literature finds adverse selection to be a reason for the poor actuarial performance of the multiple peril crop insurance (MPCCI) program. The focus of this study investigates whether premium rates can be accurately set and falls under the category of adverse selection. The literature pertaining to moral hazard will first be briefly discussed. Greater attention will be paid to the adverse selection literature and the issues

associated with accurately setting premium rates.¹²

Moral Hazard

When an agent does not bear the full cost of his actions, the potential for moral hazard exists. Insurance providers utilize deductibles and co-pays to mitigate the effects of moral hazard. The work of Borch (1960), Arrow (1971) and Stiglitz (1975) were foundational to the rigorous analysis of moral hazard in insurance markets (Raviv, 1979). Chambers (1989) was one of the earliest studies to utilize these frameworks to examine moral hazard in the MPCCI program.

Scholars have looked for and found evidence of moral hazard in the crop insurance program by examining input use, output production, and impacts on indemnities. Theoretical work by Knight and Coble (1997), Ramaswami (1993), Coble, Knight, Pope, and Williams (1997) and Horowitz and Lichtenberg (1993) has shown that the expected impact of moral hazard on the use of inputs and production can only be determined empirically.

Political interest in the effect of crop insurance on the environment has spurred a great deal of research on chemical use (Knight and Coble, 1997). Horowitz and Lichtenberg (1993) find the use of chemical inputs to increase, Walters, Shumway, Chouinard, and Wandschneider (2012) find no significant change, while Quiggin, Karagiannis, and Stanton (1993), Smith and Goodwin (1996), Babcock and Hennessy (1996), Goodwin (2001a), Goodwin (2001b), Goodwin and Smith (2003),

¹²For a review of the crop insurance literature up to 1997 please refer to Knight and Coble (1997).

Goodwin, Vandever, and Deal (2004), Mishra, Wesley Nimon, and El-Osta (2005) and Serra, Goodwin, and Featherstone (2003) find a decrease in chemical input use.

Quiggin et al. (1993) finds a reduction in output and therefore evidence of moral hazard, but Roberts, Key, and O'Donoghue (2006) finds no significant evidence of decreased output. Another approach to examine moral hazard, pioneered by Smith and Goodwin (1996), is to examine expected indemnities; using this approach Coble et al. (1997) find moral hazard effects in poor production years.

Moral hazard is certainly at work in the crop insurance program. Theoretically, its effect on input, output and expected indemnities is unknown (Knight and Coble, 1997). Glauber (2004) hypothesizes that changes to the program, notably lower deductibles, may have increased the effects of moral hazard. This study does not investigate moral hazard, though its effects on the crop insurance program are an important area of research.

Adverse Selection

The poor actuarial performance of and low participation in multiple peril crop insurance (MPCI) has, in part, been attributed to the effect of adverse selection (Knight and Coble, 1997). Adverse selection occurs because purchasers of insurance understand their risks better than insurance providers; therefore, premium rates may not accurately reflect the heterogeneous risks of the insured. To the extent that premium rates are accurately specified, they reflect “the distribution of indemnities relative to liabilities” (Atwood, Shaik, and Watts, 2002, p. 172).

For MPCI the primary impediment to assessing indemnities relative to liabilities is the determination of insurable yields (Smith and Goodwin, 1996). The majority of the crop insurance literature concerned with the direct estimation of adverse selection has examined the issue of insurable yields by investigating: 1) the distributional properties of yields, including the relationship between yield mean and variance, 2) the implications of not detrending yields, 3) smoothing techniques used to overcome data limitations and 4) the state dependent spatial correlation of yields or systemic risk.

The introduction of revenue products and their dominant role in the MPCI book of business warranted a more limited investigation into the distributional properties of prices. The dependency introduced between prices and yields also complicates the rating of revenue products. Recent studies have shown that assumptions regarding the underlying dependency, particularly those accounting for low-probability tail events, have a significant impact on the precision of estimated premium rates.

Indirect Evidence of Adverse Selection

As early as 1949, Halcrow, p. 419 observed that “[the] failure of [MPCI] to become self-supporting suggests that adverse selectivity must have been widespread, farmers with better than average expectations of collecting indemnities being the ones most frequently attracted to the program.” Subsequent studies of the demand for crop insurance have been interpreted as providing indirect evidence of adverse

selectivity in the MPCCI policies (Knight and Coble, 1997). Also, studies of the demand for MPCCI are studies of the demand for publicly provisioned and subsidized policies because there has never been a successful private market (Kramer, 1983; Wright and Hewitt, 1994).

Most studies of the aggregate demand for MPCCI examine rates of participation in per acres or per bushel terms (Knight and Coble, 1997). Hojjati and Bockstael (1988) and Nieuwoudt and Bullock (1985) found a positive and significant coefficient for expected indemnity. Goodwin (1993) reported that counties with a lower risk of loss have a lower elasticity of response to changes in premium rates, suggesting across the board rate increases would exacerbate adverse selection. Goodwin (2001b) and Goodwin et al. (2004) found indications of adverse selection because the interaction between the lost cost ratio (LCR) and premium rate indicated that for counties with higher LCRs the quantity of crop insurance demanded increased.

At the farm level, Coble, Knight, Pope, and Williams (1996) and Smith and Baquet (1996) find an significant coefficient for expected returns to insurance; but Smith and Baquet (1996) are also able to distinguish between the increased participation of farmers with positive expected returns to yield variability, versus decreased participation for farmers with negative expected returns. Smith and Goodwin (1996) and Goodwin (2001a) found an interaction between premium rates and the yield coefficient of variation to be positive and significant, indicating premium rates do not fully incorporate relative risk and adverse selection is evident.

However, Serra et al. (2003) find similar interactions to be statistically insignificant.

Direct Estimation of Adverse Selection

In contrast to studies of participation that indirectly indicate the effects of adverse selection, studies that identify the direct effect of adverse selection in the MPCCI program do so by analyzing contract provisions and rate setting procedures that allow for this behavior (Knight and Coble, 1997). Some of the earliest theoretical work on adverse selection and insurance by Rothschild and Stiglitz (1976) showed that a competitive equilibrium may not exist in insurance markets with adverse selection. Ahsan, Ali, and Kurian (1982) found that a separating equilibrium for high and low risk groups does not exist, but because their model allows farmers to specify the size of risk, they find that a pooled equilibrium may exist.

Before the 1980 Federal Crop Insurance Act, insurable yields were determined by county averages (Skees and Reed, 1986). In reference to the pre-1980 MPCCI policies, Halcrow (1949) identified that the assumptions behind this policy were that: 1) yields exhibited no time trends, 2) in a given county, all farms have the same average yield standard deviation and 3) the mean yield for all farms was the same. Botts and Boles (1958) examined the assumed normality of yields and were particularly troubled by the assumption that all farmers have the same standard deviation. Botts and Boles (1958, p. 736) speculated that “[it] would seem that the standard deviation of acre yields might be greater when the area yield is

high than in years when it is low.” It is striking that, almost 60 years ago, several of the major areas of scholarly investigation into the distributional properties of insurable yields were outlined: 1) the underlying distribution of yields, 2) the relationship between variation and yield size, and 3) the trend in yields.

The issue of how yields are distributed, whether they can be assumed normal or not, has been contested because alternative specifications have large impacts on the premium rates charged. Sherrick, Zanini, Schnitkey, and Irwin (2004) found that different underlying distributions had large impacts on the premium rate charged. Just and Weninger (1999) reexamined much of the previous literature, which had reported non-normal yield distributions, and argued that several methodological problems, including the failure to appropriately account for time trends, led to the conclusion that yields are non-normal. They concluded that normality cannot be rejected in the absence of a more tractable alternative. Atwood, Shaik, and Watts (2003) generally agreed with the criticisms of Just and Weninger (1999), but demonstrate that the alternative detrending specification of Just and Weninger (1999) increased the probability of committing a type-II error and failing to reject normality. Harri, Erdem, Coble, and Knight (2009) attempt to reconcile the conflict by examining many of the previous studies and commodities. They conclude that the distribution of yields is extremely heterogeneous. Thus, previous studies reflect the underlying heterogeneity of crops and locations.

Accounting for trends in yields is central to the debate over the normality of yields because it is fundamental to correctly identifying the underlying distribution

(Harri et al., 2009). Halcrow (1949) first identified the issue of assumed yields trend stationarity and its implications for adverse selection. Skees and Reed (1986) simulated and empirically examined the effects of yield trends. In their simulations the authors found farmers with positive yield trends could not purchase sufficient coverage. Empirically, the authors found that, while a yield trend existed for the producers in their study the individual producers did not have idiosyncratic yield trends. They conclude that not including yield trends induces adverse selection. Woodard, Sherrick, and Schnitkey (2011) show that the lost cost ratio (LCR) method of establishing premium rates will generate upward biased estimates if the liability growth rate is greater than the growth rate of expected indemnities, a condition they found for Illinois corn. The concern over yield trends led the RMA to develop a new policy endorsement that allows producers to elect a trend-adjusted actual production history (APH) as an option (Edwards and Plastina, 2014).

The relationship between mean yields and yield variation has also been closely examined. Botts and Boles (1958) hypothesized that the standard deviation would be greater when mean yields are higher. Skees and Reed (1986) examined this relationship and found no significant correlation between yield means and standard deviations; however, they found that the coefficient of variation tended to decline as mean yields rose. This result indicated that producers with higher than average yields were being charged premium rates that were too high and some form of discounting should be applied (Glauber, 2004). Goodwin (1994) confirmed this finding, but also noted that producers with higher than average standard deviations

are likely to be undercharged and producers with smaller than average deviations are likely to be overcharged. He reports evidence of adverse selection because producers with higher than average yield variation were more likely to insure. However, Schurle (1996) found no significant relationship between yield size and yield variability. The Risk Management Agency's (RMA) current practice is to individualize premium rates by loading the county unloaded base premium rate by an individual yield ratio that effectively decreases the premium rate as average producer yields increase (Coble et al., 2010).

In addition to criticism that traces its lineage to some of the earliest examinations of crop insurance, specific RMA practices for increasing credibility and overcoming limited producer yield and historical loss information have received considerable attention.¹³ These areas include the examination of how the RMA adjusts premium rates based upon: 1) coverage level differentials, 2) insurable units and 3) spatial smoothing practices.

The majority of loss histories occur at the 65% coverage level; therefore, the RMA adjusts all information to this level and then readjusts premium rates to higher or lower levels. Historically, a constant rate relativity was applied to adjust premium rates to higher coverage levels; however, higher risk producers were overcharged at higher coverage levels because they had a higher probability of realizing a yield above their coverage level (Coble et al., 2010). It was recommended that coverage levels should decrease as the county base premium rate increased

¹³A detailed description of these practices is outlined in Appendix D.

(Coble et al., 2010). Babcock, Hart, and Hayes (2004) showed that constant rate relativities caused inconsistent premium rates that could not be simultaneously actuarially fair. In response to this criticism the RMA introduced relative rates that depend on the coverage level and base premium rate (Coble et al., 2010), but the adjustment assumes a very specific functional form that may or may not accurately represent relative risks.

The RMA's practice of allowing producers to insure at different unit levels has been subject to similar critiques. Producers have the option of insuring fields in different crops and locations separately as optional units or pooling the same crops in the same location into basic units or pooling the same crop in any location into enterprise units (Plastina and Edwards, 2014). A discount is applied to the rate as units are pooled because the underlying asset is more diversified (Coble et al., 2010); however, the discount is a constant factor. Schurle (1996) found that yield variability and size of enterprise in acres were significantly related. Knight and Coble (1999) found similar evidence and noted that the ten percent discount for basic units as compared to optional units was justified. However, Shaik and Atwood (2003) found a disproportionate demand for optional units that belies premium rate misspecification and adverse selectivity. Shaik and Atwood (2003) noted that while premium rates are higher for optional units, the spatial correlation between units and risks is very high and the discount might be underestimating these correlated risks.

The RMA tries to increase the credibility of its rates by using aggregate data

to inform individual rates. Thus, the RMA employs several smoothing techniques to make loss cost estimates more credible: 1) set base rates at the county level, 2) cup and cap historical county losses at the 80th percentile 3) spatially weight the county base rate by experience of surrounding counties, 4) redistribute excess county losses across the state, 5) redistribute weighted state losses in excess of the maximum by county, and 6) load by miscellaneous factors (i.e. prevented planting, replant, quality). Premium rates are then individualized to the producer by: 1) the individual's yield ratio, 2) the type or practices of individual producers, 3) the insured unit, and 4) the selected coverage level. Furthermore, a farmer's expected yield is established on four to ten years of historical data and if there is missing data, the average for the county is scaled to form transitional yields (T-yields) that are then used in the premium rate calculation. Each of these techniques represents an effort to increase the credibility of the individual's rate by incorporating the information of similar producers. However, to the extent that aggregate information does not represent the individual's unique experience, rates will be misspecified and adverse selection will inevitably exist (Goodwin, 2001a).

Smoothing is used both to overcome and take advantage of the spatial dependence between individual producers. Proximate producers have similar outcomes because they experience similar weather events. Historically, the dependence scholarship has examined the impact of spatial and temporal correlation between yields. Miranda (1991) reconsidered the potential of area yield plans as a way to effectively mitigate systemic risk while reducing deductibles and premium

rates. Miranda and Glauber (1997) goes so far as to argue that systemic risk may be the more fundamental problem of providing MPCI. Wang and Zhang (2003) determined that while yields are not independent they are finite range positive dependent which is a weaker, but satisfactory condition for insurability.

Goodwin (2001a) evaluated the degree of spatial correlation in three major corn producing states and found that while correlation does decay rapidly under normal conditions, it is state dependent because under drought conditions spatial correlation is far more persistent. Hungerford and Goodwin (2014) derives spatial degrees of freedom for corn and soybean yields in McLean county Illinois with surrounding counties. The authors use these estimated degrees of freedom to calibrate a Student-t copula, and more accurately model the spatial correlation they find between county yields.

Proponents of multiple peril crop insurance (MPCI) have claimed that problems of systemic risk are severe enough to induce market failure and warrant government intervention; however, such claims are questionable given the ability of private markets to reinsure risk that is orders of magnitude larger than those of MPCI (Goodwin and Smith, 2013).

The advent of revenue products and their reliance on simulated price and yield realizations to generate premium rates, requires accurately modeling not only a county's unique price-yield dependence, but the multivariate distribution of their joint outcomes. The current ratemaking practice uses the equivalent of a Normal copula to impose a county's historical price-yield correlation. The use of a Normal

copula also imposes a multivariate distribution with zero tail dependence on the price and yield realizations. Analysis of revenue insurance ratemaking practices has focused on: 1) the estimation of county price-yield dependence, the assumed marginal distribution of prices, 2) the impact of calibrating an assumed Normal copula with imprecisely estimated county price-yield dependence estimates and 3) the impact of the unique distributional properties of the Normal copula on premium rates as compared to alternative copula specifications. This study falls directly within this vein of scholarship.

Goodwin, Roberts, and Coble (2000) investigated MPCIR revenue products and the assumed marginal distribution for prices. While alternative price specifications have been explored in the financial literature, the assumption of log-normality is generally accepted, especially given the tractability of the distribution (Coble et al., 2010). Goodwin et al. (2000) found strong support for log-normality and that premium rates estimated under an assumption of normality are significantly lower than those estimated with log-normality.¹⁴

A surprisingly sparse literature exists on the relationship between region-crop output and prices. Tew and Reid (1988) finds the expected negative correlation between cotton, corn and oat yields and their prices and that cotton has the strongest negative correlations because of its regional concentration. Paulson and Babcock (2008) calculated national correlations between percentage yield deviations and percentage price deviations for corn and soybeans between the periods:

¹⁴Under the “COMBO” consolation, revenue products assume a log-normal price distribution (Goodwin, 2014).

1975-2005, 1980-2005 and 1990-2005. The latest period's coefficient was stronger than the previous periods because of policy changes in 1995 (Lence and Hayes, 2002). The current ratemaking practice depends on the estimation of county-crop price-yield correlation estimates.

The issue of how prices and yields depend upon one another throughout their distributions, beyond correlations, has been addressed by Goodwin (2014), Goodwin and Hungerford (2014) and Hungerford and Goodwin (2014). Hungerford and Goodwin (2014) examined the precision of revenue premium rates when dependence is estimated based upon limited available data. The authors estimated revenue product premium rates at the 65% coverage level for McLean county Illinois, through a process similar to the RMA's and find that premium rates can vary substantially, as a function of the limited data upon which correlation estimates are generated.

Goodwin and Hungerford (2014) examined the impact of alternative copulas, when specifying price-yield dependence, on premium rates. The authors found that "premium rates vary significantly across the alternative copulas," as the estimation of the shape parameter also vary extensively (p.12). Furthermore, the authors find very large differences between the implied premium rates generated by different copulas as the coverage level decreases. Goodwin and Hungerford (2014) extend their work by using nested copulas to develop premium rates for whole farm units, which insure a producer's enterprise across space and crops. The authors do not find large differences between premium rates generated between the copulas they explored and attributed this result to the low correlation at the unit versus county

level.

Goodwin (2014) used dependence modeling as an example of the opportunities to apply advanced modeling capabilities to real-world problems. He found that: 1) several copulas might validly model the price-yield dependence for McLean county Illinois, 2) premium rates from alternative-equally supported copulas-differ, 3) and that they differ more at lower coverage levels..

Summary

A broad literature on the problems endemic to insurance markets in general, and crop insurance in particular, exists. This study falls within the category of direct estimation for the potential of adverse selection through the misspecification of premium rates. More specifically, it falls under the more recent literature examining the complications of rating revenue insurance products and the use of copulas to accomplish this task. This study extends the work of Hungerford and Goodwin (2014), Goodwin (2014) and Goodwin and Hungerford (2014). Each of these studies note the effect of coverage level on copula generated premium rates, Goodwin (2014) also report a convergence of premium rates as a function price-yield dependence.

In the experiments to follow, the effects of four factors that determine premium rate level: 1) coverage level, 2) price-yield dependence, 3) the dispersion of the price marginal distribution and 4) yield marginal distribution, are explored across a set of alternative copulas. Moreover, the sampling error investigated by

Hungerford and Goodwin (2014) is also explored across the variation in the four premium rate determining factors. Finally, the copula selection methods applied by Goodwin and Hungerford (2014) for one corn producing county is extended over nine counties that represent a range of empirical yield dispersion and price-yield dependence characteristics, in conjunction with variation in coverage level and price dispersion.

CHAPTER FIVE

COPULAS

When the observation of some event provides information as to the likelihood of another event occurring, these events depend on one another. Copulas are functions that characterize the dependence between random variables, while retaining the unique properties of these random variables underlying marginal distributions (Schmidt, 2007). Copulas originate from the work of Sklar (1959), who derived the following theorem.

Sklar's Theorem: *For a d -dimensional cdf F with marginals $F_1, \dots, F_d$ there exists a copula C , such that*

$$F(x_1 \dots x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (12)$$

$\forall x_i$ in $[-\infty, \infty], i = 1, \dots, d$ (Nelson, 1999).

Using the probability integral transformation the copula itself can be represented as

$$C(u) = F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d)) \quad (13)$$

where $F_i \circ F_i^{\leftarrow}(y) \geq y$ and $u_1, \dots, u_d$ are distributed uniformly (Schmidt, 2007).

Thus, if the cdf is differentiable, then $F_i^{\leftarrow} = F_i^{-1}$ and the density of a copula

can be expressed as

$$c(u) = \frac{f(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1)) \cdots f_d(F_d^{-1}(u_d))} \quad (14)$$

where f is the joint density and $f_i = 1, \dots, d$, are the marginal densities (Schmidt, 2007).

An important implication of Sklar's Theorem is that there is an infinite number of ways to represent the dependence between random variables. For this reason, it is not practical to consider all potential copulas. In this study the infinite set of copulas is culled and consists only of the best understood and utilized. Yet, even in this more limited set, it is not within the scope of this study to address all copulas for which there is a mathematical representation or even implementation in statistical software. This is why one parameter copulas,¹⁵ implemented in **R** that have both goodness-of-fit procedures and rotated representations are examined.¹⁶

Major Copula Families and Members of Interest

The two major branches of copulas are elliptical and Archimedean. They differ in derivation. Elliptical copulas are derived from distributions while Archimedean copulas can be stated explicitly (Schmidt, 2007). With one exception the dependence imposed upon the marginal distributions is determined by one

¹⁵Copulas for which only one parameter generates the dependence structure. The Student-t copula is examined in this study, but has two parameters.

¹⁶Price and yield observations are negatively related and require rotated representations of some copulas.

parameter.¹⁷ Because of their relative simplicity one parameter copulas have been more thoroughly studied and from a practical standpoint goodness-of-fit tests in the **copula** R package have been developed by Hofert, Kojadinovic, Maechler, and Yan (2014).

Elliptical Copulas

There are two major elliptical copulas: the Gaussian or Normal and the Student-t. A d -dimensional Normal copula has the form:

$$C_\rho(\mathbf{u}) = \phi_\rho(\phi^{-1}(u_1), \dots, \phi^{-1}(u_d)), \quad (15)$$

where ρ is an (d, d) dimensional linear correlation matrix of ρ and ϕ^{-1} is the inverse CDF of the standard normal distribution (Hungerford and Goodwin, 2014). In the Normal copula dependence is determined by one parameter, Pearson's linear correlation coefficient ρ between the marginal distributions, which can range from -1 to 1 and imposes linear and zero tail dependence (Schmidt, 2007).

The Student-t copula is very similar to the Normal and it has an analogous relationship as the univariate Student-t distribution has to the standard normal distribution. Like its univariate counterpart, it too is specified with a degrees of freedom parameter (df) and as $df \rightarrow \infty$ the copula approaches the Normal. The Student-t copula introduces a symmetric linear dependence in both tails and will

¹⁷The Student-t copula, like its univariate counterpart, is identified by both a correlation parameter and a degrees of freedom parameter; for this study the degrees of freedom are fixed at four, unless estimated by the Akaike Information Criterion.

even impose tail dependence at low degrees of freedom when $\rho = 0$. The Student-t copula is the only two parameter copula that will be considered in the subsequent experiments and only for one, relatively low, $df = 4$ in order to ensure that its shape is sufficiently different than the Normal copula. Student-t copulas of d-dimensions are of the form:

$$C_{df,\rho}(\mathbf{u}) = t_{df,\rho}(t_{df}^{-1}(u_1), \dots, t_{df}^{-1}(u_d)) \quad (16)$$

Archimedean Copulas

Archimedean copulas of d-dimensions can be generalized as:

$$C(\mathbf{u}) = \psi(\psi^{-1}(u_1) + \dots + \psi^{-1}(u_d)) \quad (17)$$

where $\psi(\mathbf{u})$ is called the generator function and $\psi^{-1}(u_i)$ is the inverse generator function (Schmidt, 2007). The best understood and most frequently implemented are the Clayton, Gumbel and Frank families.

Figure 1 shows the radically different higher order dependency structure these copulas can impose despite operating over the same $\phi(0, 1)$ standard normal marginals with their θ parameter calibrated to generate a $\tau = .5$ for the Archimedean copulas or $\rho \approx .71$ for the elliptical copulas. The Clayton copula introduces lower tail dependence, the data fans out from the origin, with a parameter that is a positive non-zero real number. Gumbel copulas introduce upper tail dependence,

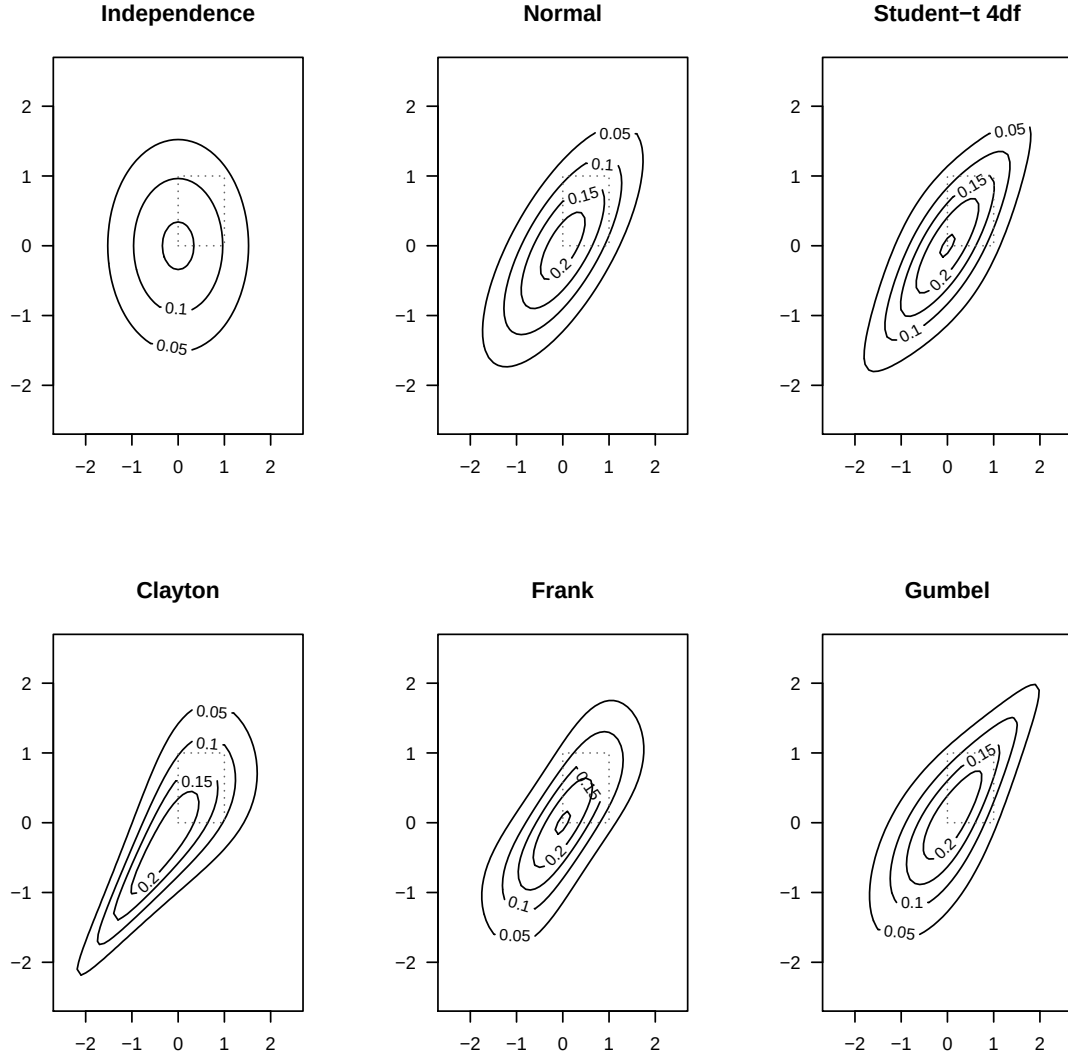


Figure 1: Contours of Copula Family Densities

Notes: These copulas are all calibrated to the same dependence of Kendall's $\tau = -0.5$. The Student-t copula is shown with a degrees of freedom parameter equal to four.

the data fans out towards the origin, with a parameter value ranging from one to infinity. Frank copulas introduce no tail dependence, the data are more square-shaped, and have positive non-zero parameter values. The Independence copula is shown because at weak dependence calibrations every copula in this study,

except the Student-t, converge to the Independence copula.¹⁸ A general overview of the important characteristics of the major copulas are summarized in Table 1.

Table 1: Summary of Copula Characteristics

Copula	Generator	Parameter(s)	Tail Dependence(lower,upper)
Normal	$\phi(\mathbf{u})$	$-1 \leq \rho \leq 1$	none
Student-t	$t_{df,\rho}(\mathbf{u})$	$-1 \leq \rho \leq 1$ $df \in [1, \infty)$	$2t_{df+1} \left(-\sqrt{df+1} \sqrt{\frac{1-\rho}{1+\rho}} \right)$
Clayton	$(1 + \mathbf{u})^{-1/\theta}$	$\theta \in (0, \infty)$	$(2^{-\frac{1}{\theta}}, 0)$
Gumbel	$e^{-\mathbf{u}^{1/\theta}}$	$\theta \in [1, \infty)$	$(0, 2 - 2^{\frac{1}{\theta}})$
Frank	$-\log(e^{\mathbf{u}-\log(e^\theta)})$	$\theta \in (0, \infty)$	$(0, 0)$

Rotated Copula Representations

The copulas used in the subsequent rating sections need to be able to represent negative dependence. While the Normal, Student-t and Frank can exhibit negative dependence, the Clayton and Gumbel cannot. However, negative dependence can be imposed through transformation. These rotated representations are implemented in the **R** package **CDVine** developed by Brechmann and Schepsmeier (2013) and represented in Figure 2. The packages allows copulas to be rotated 90°, 180° and 270°.¹⁹ The top three copulas in Figure 2 are all symmetric despite the Frank being an Archimedean copula. Also note that the eventual shape of the Clayton and Gumbel copulas for their cross rotations are similar, despite

¹⁸The finite sample performance of the available goodness-of-fit tests will be assessed over these copulas—excluding the Independence. Two of these copulas—the Clayton and the Frank—cannot strictly represent negative dependence; therefore, rotated representations of these copulas are required to generate premium rates for negatively dependent price and yield outcomes in later analysis.

¹⁹180° rotations retain the positive representation of the initial copula, but change its direction. Therefore, only the 90° and 270° representations are of interest for this study.

coming from different generating processes.²⁰

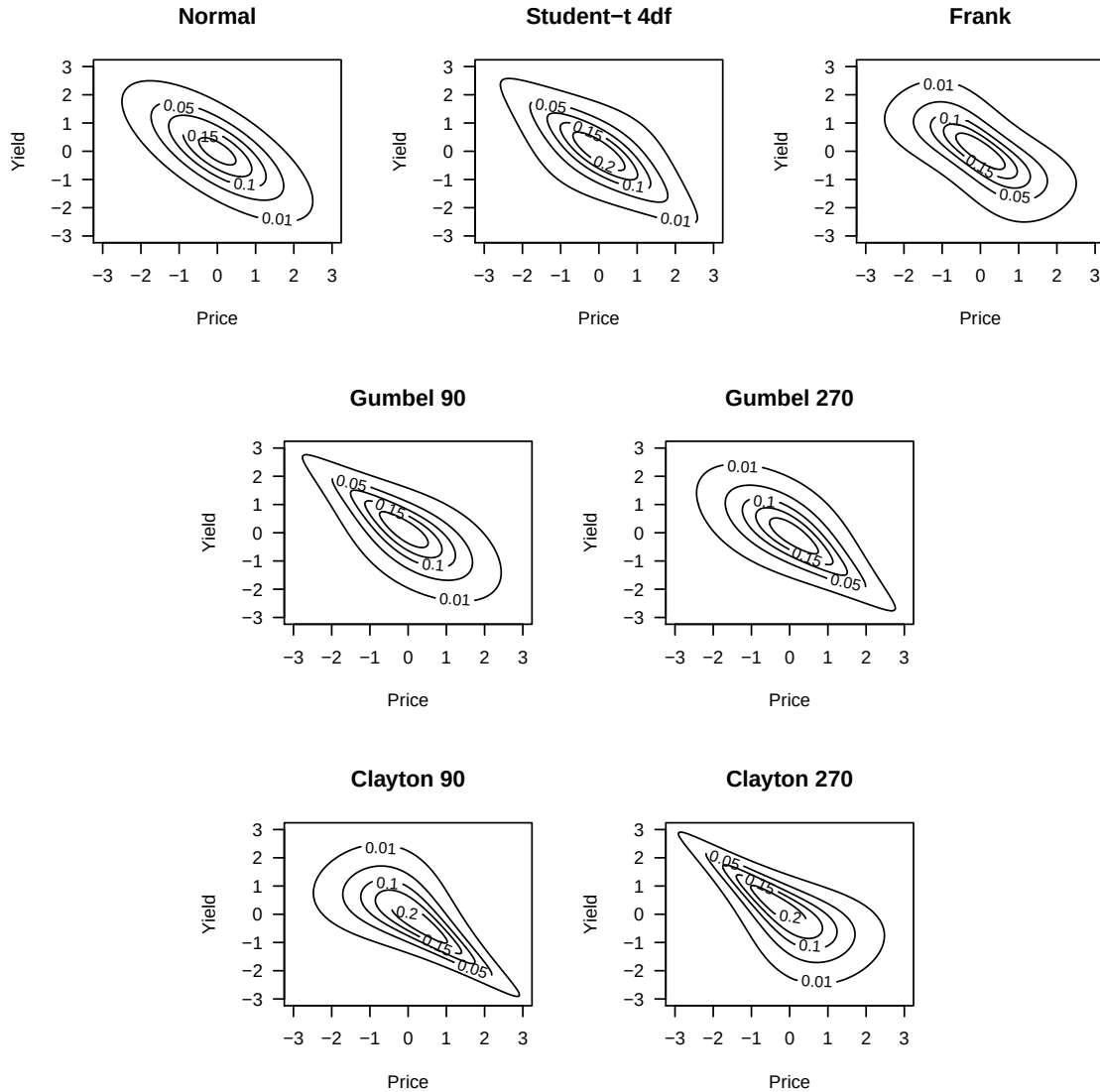


Figure 2: Rotated Copula Representations

Notes: These copulas are all calibrated to the same dependence of Kendall's $\tau = .5$. The Student-t copula is shown with a degrees of freedom parameter of 4. The numeric label for the asymmetric Archimedean copulas indicates the degrees of rotation.

²⁰The general shapes of the copulas, especially the asymmetric Archimedean, are important to the premium rate results because premium rates generated from similar copulas with similar shapes are more alike than those from the same copula family.

Theoretical Price-Yield Representations of Copulas

Different copulas represent different theoretical assumptions about the relationship between prices and yields. Particularly, the assumed dependence between prices and yields during extreme events. The Normal, Student-t and Frank copulas represent symmetric price and yield dependence in that the probability of extremely low-price and high-yield events are as likely as high-price and low-yield events. However, the Frank copula models a relationship in which cross-extremal outcomes like low-price and low-yield events are more likely to occur, which is reflected in its square shape. The Student-t copula, especially at low degrees of freedom, reflects a greater dependence between events in either tail, which is evidenced by its diamond shape.

The remaining Archimedean copulas: the Gumbel and Clayton, represents asymmetric assumptions about the relationship between prices and yields. The Gumbel rotated 90° (g90) and Clayton rotated 270° (c270) copulas represent a price-yield relationship where low-price and high-yield outcomes are more highly dependent upon one another. Thus, a low-price is highly likely to be associated with a high-yield event and the data fans out indicating that high prices could occur during low-yield events, but they could also occur during “other” market situations that aren’t related to yields. Conversely, the Gumbel rotated 270° (g270) and Clayton 90° (c90) portray the opposite state of the world, in which high-price outcomes are more likely to be associated with low yields and the data fans out as

low-price outcomes are more likely to be associated with low-yield outcomes.

The state of the world represented by the Clayton rotated 90° and Gumbel rotated 270° copulas, has the greatest potential to generate divergent premium rate outcomes, especially for the Revenue Protection (RP) product that offers upside price protection, because it represents a world where the natural hedge between prices and yields weakens for extreme outcomes and there is a higher likelihood of a low-yield and low-price event that is indemnified at a higher project price (PP).²¹

²¹An overview of the different tail dependence imposed by the Clayton and Gumbel Copulas is provided in Appendix E.

CHAPTER SIX

COPULA GOODNESS OF FIT STATISTICS PERFORMANCE

The first issue that must be addressed when using copulas is selecting the copula that best fits the data. There are a theoretically infinite number of copulas to choose from. This study, follows the design of previous studies performed by Genest, Rémillard, and Beaudoin (2009) and Kojadinovic, Yan, Holmes, et al. (2011) that have assessed the performance of available copula goodness-of-fit statistics to accurately identify the “true” data generating copula, while discriminating against false alternatives. However, this experiment is performed over the low sample sizes and weak dependence that currently defines the ratemaking environment for multiple peril crop insurance (MPCI) revenue products. The results of this study indicate that available goodness-of-fit tests cannot identify the underlying copula and the implications of this finding motivate the subsequent experiments.²²

Design

The Monte-Carlo experiments presented here repeat the design of the previous studies with some modifications due to the characteristic of MPCI data. This study will examine the performance of the fast multiplier method (*Mult*) of generating goodness-of-fit test statistics and p-values developed by Kojadinovic et al. (2011).²³ A set of copulas consisting of the Clayton (C), Frank (F), Gumbel (G),

²²For a more detailed description of the previous studies and goodness-of-fit attributes, refer to Appendix I.

²³Preliminary runs of the parametric bootstrap (*Pb*) method developed by Genest et al. (2009) confirmed the computational expense noted by the authors.

Normal (N) and Student-t with four degrees of freedom (t4) will serve as both the “true” copula (*genCop*), from which data calibrated with various degrees of dependence are generated, and the hypothesized copula (*gofCop*).²⁴ The Cramér-von Mises (S_n) test of goodness-of-fit is the only test statistic implemented by the fast multiplier approach, therefore it is the only goodness-of-fit test statistic examined. Following Kojadinovic et al. (2011), three different methods of estimating the hypothesized copula’s θ parameter were used: the inversion of Kendall’s τ (i- τ), the inversion of Spearman’s ρ (i- ρ), and maximum pseudo-likelihood (*mpl*).²⁵ Each method assumes that the data are pseudo-observations $\sim [0, 1]$ and that ties occur with probability zero.²⁶

The simulation repeats the process for each combination of “true” and hypothesized copula, for each method of estimating θ , three sample sizes (n) equal to 50, 75 or 150 and three levels of dependence. These sample sizes were selected because there are at most 45 years of paired price and yield data for corn,²⁷ a sample size of 50 allows the results for the *Mult* approach to be compared to the parametric bootstrap (*Pb*) results reported in by Beaudoin (2007); a sample size of 75 pairs provides a preview of future performance conditional upon 25 years of additional data and for the possibility of comparing these results to the unreported results of

²⁴Although the Plackett copula is implemented in **R** and studied by Kojadinovic et al. (2011), it does not have a rotated representation, which will become crucial when the copulas are used to model the negative dependence of prices and yields.

²⁵For this research project, the maximum likelihood estimation method failed over the pseudo-observations distributed $\sim [0, 1]$ generated in this experiment. Also, the estimation method i- ρ produced inconsistent results throughout and was not included in the results.

²⁶Ties were broken randomly.

²⁷There is county yield data for corn for most major producing counties going back almost 70 years; however, futures data are only available starting in 1969.

Kojadinovic et al. (2011); and 150 was selected, to compare these results with those of Kojadinovic et al. (2011) and test **R**'s implementation for this project.²⁸

Three different levels of dependence, as measured by Kendall's τ equal to 0.01, 0.25 or 0.50 are imposed on the various number of observations: empirically, dependence as measured by Kendall's τ over the major corn producing counties in the study, range from approximately -0.48 to -0.01, with an average Kendall's τ of -0.21. Thus, Kendall's τ of 0.25 is close to the absolute value of the mean empirical dependence while Kendall's τ of 0.50 is close to the absolute value of the empirical maximum, with the additional benefit for allowing these results to be compared with previous findings. The previous studies examined Kendall's τ of 0.75, this is a much stronger degree of dependence than is observed for corn; therefore, Kendall's τ of 0.01 was selected to better reflect the weak dependence exhibited between some counties' yields and national prices.

The Monte-Carlo specifications used by the three previous studies and the present one are presented in Table 26. In all, 225 unique tests are repeated 10,000 times with their performance measured by the numerical count of the percent number p-values less than or equal to a significance criterion of 5%. The process was separated on the basis of the two methods of estimating the copula σ parameter and run in parallel on separate CPUs of a high power computer over the 225 possible combinations.

²⁸Genest et al. (2009) also ran their experimenter at 150 observations so results are compared with the findings for the *Pb* method as well.

Results

Table 2 shows the results of the simulation study for data representing current multiple peril crop insurance (MPCI) number of observations, a sample size of 50 ($n = 50$) and the range of dependence. Tables 27 and 28 show the results of the simulations when performed at sample sizes of $n = 75$ and $n = 150$. At the comparable data levels (i.e. $n \in \{50, 150\}$ and $\tau \in \{.25, .50\}$), the results for this study are similar to those of previous studies for the maximum pseudo-likelihood (textitmpl) method of estimating the σ parameter for the hypothesized copula. The method $i\text{-}\tau$ generates results that are not similar, when the Frank copula is the “true” copula (*genCop*). Method *mpl* is the most reliable and is used in the subsequent experiments in which a goodness-of-fit test is applied for each alternative copula.

As either the sample size (n) or the imposed measure of Kendall’s τ (τ) increases, the tests approach their nominal levels²⁹ and the power of discrimination between false alternative copulas increases.³⁰ As seen in Table 28, at 150 observations and a high level of Kendall’s τ , the tests falsely reject the “true” copula at or below the 5% level and almost always correctly reject false alternatives.

However, at current levels of data, approximated by the results of Table 2 with $n = 50$, the tests appear to have no power of discrimination and do not hold their

²⁹The test does not reject a “true” copula more than 5%-the nominal level.

³⁰The tests more frequently reject false alternative copulas (i.e. the likelihood of a type-II error decreases).

Table 2: Percent Rejection of Null Hypothesis with $n = 50$

Study	genCop	Sim	Est	$\tau = .01$					$\tau = .25$					$\tau = .50$				
				C	F	G	N	t4	C	F	G	N	t4	C	F	G	N	t4
s2015	C	Mult	<i>mpl</i>	18	20	20	20	17	12	52	67	45	40	12	88	95	79	75
b2007		Pb	<i>i-τ</i>						5	25	47	23	27	6	57	86	54	59
s2015		Mult	<i>i-τ</i>	12	11	11	11	15	11	33	53	32	35	8	60	88	62	65
s2015	F	Mult	<i>mpl</i>	13	13	13	13	18	19	12	26	13	15	37	11	45	21	24
b2007		Pb	<i>i-τ</i>						8	5	16	6	9	21	5	27	8	12
s2015		Mult	<i>i-τ</i>	0	0	0	0	0	3	0	2	1	1	33	3	20	6	9
s2015	G	Mult	<i>mpl</i>	12	11	12	11	13	39	17	6	12	10	75	27	5	12	11
b2007		Pb	<i>i-τ</i>						15	4	6	4	5	49	8	5	4	5
s2015		Mult	<i>i-τ</i>	8	8	8	8	10	32	11	10	10	12	66	13	8	10	12
s2015	N	Mult	<i>mpl</i>	13	14	13	13	19	10	11	14	7	11	26	17	23	6	8
b2007		Pb	<i>i-τ</i>						7	6	14	5	8	18	8	20	5	7
s2015		Mult	<i>i-τ</i>	11	11	11	11	14	18	11	20	10	14	36	11	26	8	11
s2015	t4	Mult	<i>mpl</i>	11	11	11	10	12	12	15	17	11	8	26	29	28	13	8
b2007		Pb	<i>i-τ</i>						5	5	11	4	5	16	9	17	4	5
s2015		Mult	<i>i-τ</i>	10	11	10	11	11	16	10	18	10	11	33	11	23	7	8

Notes: These results are the percent rejection of the H_0 hypothesis that the data come the alternative copulas for goodness-of-fit tests at the 5% significance level from the simulations conducted at 50 observations ($n = 50$) for the study performed by Beaudoin (2007) (b2007) and the present study (s2015). The “true” copulas (*genCop*) and hypothesized copulas are the Clayton (C), Frank (F), Gumbel (G), Normal (N) and Student-t with 4 degrees of freedom (t4) and values are in bold font when the copula generating the data is the same as the hypothesized copula. The simulation methods (Sim), for generating the test statistics and p-values are the fast multiplier (*Mult*) and parametric bootstrap (*Pb*) methods. The estimation methods (*Est*) generating the hypothesized copula θ parameter are maximum-pseudo likelihood (*mpl*) and the inversion of Kendall’s τ (*i- τ*).

nominal level, regardless of the strength of dependence as measured by Kendall’s τ .

However, the test do perform better when the data generating copula is a Clayton, even at low sample size and weak dependence. Table 27 suggests that ratemaking procedures will not be able to confidently rely of copula goodness-of-fit tests, with 25 additional years of data. While improved, the tests still over reject the null hypothesis when true and fail to reject the null hypothesis when false. While the

parametric bootstrap approach appears to hold its nominal level better than fast multiplier approach, it also tends to have less power of discrimination and remains impractical to implement.

Summary

The data levels currently available for MPCCI premium ratemaking do not allow the goodness-of-fit test statistics to effectively discriminate between false alternative copulas. Moreover, the “true” copula is rejected at a frequency that exceeds a conventional level of statistical confidence. The current use of a Normal copula in the RMA’s ratemaking procedure for revenue insurance products cannot be statistically justified. The inability to confidently identify the best fitting copula *ex-ante* implies that the impact of definite copulas on estimated premium rates must be considered *ex-post* across a more extensive set of equally viable alternatives (Goodwin, 2014). If premium rates generated by different copulas do not deviate from one another then the use of copula modeling does not exacerbate premium rate dispersion and imprecision. However, if premium rates do differ between alternative copulas, then the use of copulas in the current ratemaking procedure for revenue products does induce additional implicit premium rate dispersion and imprecision. The first empirical experiment explores whether or not revenue product premium rates generated by alternative copulas for corn producing counties in Illinois, Indiana and Iowa diverge.

CHAPTER SEVEN

EMPIRICAL PREMIUM RATE DISPERSION BETWEEN COPULAS

Without the ability to discern between equally supported copulas, it becomes necessary to examine the effect of alternative copula specifications on premium rates as well as the conditions that cause dispersion or bias. Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates are generated for most counties in the major corn producing states of Illinois, Indiana and Iowa. Farm level premium rates are generated by county from alternative copulas and the pure empirical method. The results indicate that premium rates generated by alternative copulas diverge and there is implicit premium rate imprecision as a function of modeling price and yield dependence with copulas, which is not accounted for in the current ratemaking procedure.

Ratemaking Procedures

“Pure” empirical, parametric copula and empirical copula based approaches can be used to estimate premium rates using the lost cost ratio (LCR). In the subsequent analysis revenue policy premium rates are generated for empirical data and fully simulated data.³¹ The pure empirical method and the empirical copula method are only generated when the underlying price and yield realizations are based upon empirical data. Detrended county yields obtained from the National

³¹The ratemaking procedures used in this experiment as well as the subsequent simulation and empirical experiments is described in detail here and referenced subsequently.

Agricultural Statistics Service (NASS) and real price deviates from the futures market are used in ratemaking for the empirical experiments. Premium rates in the simulation experiments are generated by identifying the assumed marginal price and yield distributions from sets of parameters. For either experimental type, premium rates are estimated with 100,000 price-yield realizations.³² In the empirical experiments, a farm level shock is also added to county level yield deviations in order to estimate farm level premium rates at the county level.³³

The Pure Empirical Method

The pure empirical method uses the empirical price and yield observations to determine a farm specific county premium rate. Detrended yields and historical price deviates are estimated and used in the ratemaking procedure. The available empirical observations are replicated to obtain the total number of observations upon which premium rates are estimated—100,000 pairs.³⁴ Farm level yield deviations are randomly sampled from a producer database by county and the detrended county yields are shocked by this deviation. Adding farm level deviations increases dispersion in the realized yields and reflects the greater variability of yields at the farm level; *ceterus paribus* this increased dispersion increases the premium rate.

³²The industry standard is 10,000; the RMA uses 500. The use of 100,000 correlated realizations allows approximately “true” premium rates to be estimated, conditional on the underlying assumptions.

³³The RMA adjusts their county level estimates, to farm level estimates in an *ad-hoc* fashion (Coble et al., 2010).

³⁴With 53 years of data, the price-yield pairs are repeated $100000/53 \approx \lceil 1886.79 \rceil = 1887$ times, the eleven extra observations are truncated from the end ($11 = 53 * 1887 - 100000$).

General Copula Ratemaking

Unlike the pure empirical method, copula methods do not directly estimate premium rates from the empirical data. Rather the empirical data are used to calibrate the parameters for assumed marginal distributions. For parametric copulas, the copula shape parameter is also calibrated from an estimate of historical price-yield dependence. For the simulation experiments, these parameters are specified rather than estimated. The shape of copula's unique multivariate distribution is then imposed on the independently simulated price and yield realizations and premium rates are estimated from the simulated price and yield realizations structured by the copula's unique shape.

For premium rates estimated from the empirical data, the mean yields and dispersion parameters are estimated from the detrended county yield data. These parameter estimates identify the shape of an assumed normal yield distribution that is censored at zero. From the county-calibrated censored normal distribution, 100,000 yield realizations are generated and negative realizations are truncated to zero; that is,

$$y_j = \text{Max}(0, \phi(\mu_y, \sigma_y)) \tag{18}$$

where μ_y is the estimated county level detrended mean, and σ_y is the estimated county level detrended standard deviation. For empirical premium rates, producer yield deviations are sampled at the county level and these random shocks are added

to the simulated realizations. In the simulation experiments, the producer level shock is implicit to the parameter specification.

Prices are simulated from an Itô process , which imposes a log-normal distribution (Hull, 2012). 100,000 unique harvest price (HP_j) realization are generated as follows:

$$\log(HP_j) = \log(PP) - \left(\frac{(ivol)^2}{2}\right)dt + (ivol)dz \quad (19)$$

where dz is a generalized Wiener process such that $dz = \phi(0, 1)\sqrt{dt}$, $ivol$ is the implied volatility of the rating period and $dt = 2/3$, the difference in months between the forecast period (February) and the harvest period (October). The $\log(HP_j)$ realizations are distributed Normally; hence,

$$HP_j = e^{\log(HP_j)} \quad (20)$$

is distributed log-normally.³⁵

To identify this process and distribution, the mean, standard deviation and duration of time must all be determined. The mean is simply the $\mathbb{E}(p_t)$ or projected price, which in this study is: \$5.78, the average RMA calculated projected price between 2011, 2012, and 2013. Price dispersion is identified by the implied volatility in the options market on the underlying commodity contract.³⁶ Premium rates are

³⁵Appendix F provides an overview of this process and its assumptions.

³⁶Historical implied volatility for corn has ranged from .1 to .5. Recent implied volatility has been approximately .2.

calculated across a range of implied volatilities or “price dispersion” parameters to account for potential premium rate sensitivity to different price dispersion environments.

Parametric Copula Ratemaking

Parametric copula methods impose a copula’s unique dependence structure on simulated price and yield marginals by sorting the rank order dependence of the simulated price and yield marginal distributions by the copula’s equally dimensioned reference matrix. In theory, this process could be applied and premium rates generated under an infinite number of alternative copulas. Far fewer are actually used in the statistical software **R**. Amongst this more limited set the Clayton, Frank, Gumbel, Gaussian and Student-t have rotated representation that can reflect negative dependence between prices and yields. Therefore, these copulas are used to bind the simulated price and yield realizations.

The shape of the copula data are uniquely determined by the specific properties of the copula’s generator function and the specification of the copula parameter θ . Archimedean copula shape parameters are a function of the rank order dependence measure of Kendall’s τ , which is estimated from the county level dependence between log-price deviates and detrended county yields or pre-specified for the simulation experiments. For elliptical copulas, θ is equivalent to the linear correlation coefficient of Pearson’s ρ , which is also either estimated from the empirical data or pre-specified in simulation. Premium rates are estimated from the

100,000 simulated price and yield realizations, which have been reshaped by the copula's specific dependence structure and its θ parameter.

Empirical Copula Ratemaking

The empirical copula method, as the names suggests, possesses characteristics of both the pure empirical and copula techniques. Like the pure empirical method, detrended yields and scaled harvest price observations are repeated until a data-frame of 100,000 correlate pairs has been obtained. However, instead of directly estimating premium rates from the empirical observations, price and yield marginals are simulated from assumed marginal distributions (as in the parametric copula method) and these simulated values are then sorted by the rank order of the replicated empirical data. This process imposes an empirical copula on the simulated marginals. Premium rates are then generated from the simulated realizations, instead of the empirical observations.

Revenue Product Formulas

The formulas used in this study to generate Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates for a county i are similar, but not identical to those currently used in the ratemaking

process by the Risk Management Agency (RMA) and are given as:

$$\begin{aligned}
 RPHPE_i &= \sum_{j=1}^n \frac{Max(0, (cvg * \mathbb{E}(y_j) * \mathbb{E}(p_t)) - Max(0, y_j * p_j))}{n * \mathbb{E}(y_j) * cvg * \mathbb{E}(p_t)} \\
 RP_i &= \sum_{j=1}^n \frac{Max(0, (cvg * \mathbb{E}(y_j) * Max(\mathbb{E}(p_t), p_j)) - Max(0, y_j * p_j))}{n * \mathbb{E}(y_j) * cvg * Max(\mathbb{E}(p_t), p_j)}
 \end{aligned} \tag{21}$$

where the number of realizations (n) is equal to 100,000, cvg is the coverage level, $\mathbb{E}(y_i)$ average yield of the simulated realizations, $\mathbb{E}(p_t)$ is the projected price and always equal to \$5.78, y_j is a simulated yield realization, and p_j is the simulated harvest price realization.

Summary

A minor difference between these formulas and the RMA's, is that liability is not limited by realized prices greater than two times the projected price. Also, no attempt is made to augment a county's data with that of surrounding counties. Several substantial differences exist. The estimation of expected yields in the RMA's formulations is based on the $\mathbb{E}(y_i)$ of producer i 's yields or their actual production history. In the formulas here, $\mathbb{E}(y_j)$ is the average of the yield realizations from the assumed censored normal distribution. These values are similar in expectation, but they are not identical. This difference affects both the numerator and denominator of both insurance products.

Also, the RMA's formula for the denominator of the RP premium rate formula uses the projected price ($\mathbb{E}(p_t)$), which is invariant. Thus, the denominator is invariant. The premium rate denominator in this study is estimated based upon

the maximum of the realized or harvest price, which depends on the structure of the copula used to bind the data, making it a random variable. Therefore, the distributions of indemnities are standardized by the distribution of liabilities within a copula, which may better support later conclusions regarding copula modeling in general. Future research could investigate the impact of using a stable liability estimate, more similar to the current RMA practice, across different copulas on premium rates and test the sensitivity of the results of this study to such an alternative.

Data for Empirical Experiments

Price and yield data are required in order to set premium rates for multiple peril crop insurance (MPCI). Producers face the same effective price, which are represented by the futures markets for the commodity. Yields vary across counties, but individual producer yields within a county also vary. Both county level and producer level yield data are used in order estimate producer level premium rates by county. For the pure empirical method, premium rates are calculated based on empirical observations.

However, for copula methods the rank or linear dependence estimates between county level detrended yields and historical futures price deviates calibrates the copula's dependence parameter. Empirical parameter estimates identify an assumed censored normal yield distribution. Prices are simulated out of a process that generates a log-normal distribution. Thus, the data can consist of both

empirical data and simulated observations.³⁷

Prices

Empirical price data are obtained from futures data for corn, which were collected from “Financial Price Data” CDs purchased from R & C Research.³⁸ Price data for corn is available from 1969 and were collected until 2013, providing 45 years of price realizations. Historical projected prices and harvest prices are calculated from the futures data. For both calculations an average of the daily closing price for the December corn contract on the Chicago Board of Trade (CBOT) is calculated over the specified price discovery period. While these discovery periods can vary by crop and state, for Illinois, Iowa and Indiana, the projected price discovery period begins February 1st and ends February 28th.³⁹ The harvest price discovery period for these states begins October 1st and ends October 31st. Thus, historical projected prices are the February average of the December corn contract, while historical harvest prices are the November average of the December corn contract.⁴⁰ Pure empirical rates are based off of scaled harvest prices (HP_t). These values are calculated by multiplying the projected price (\$5.68) by the ratio of the historical

³⁷The data for the empirical experiments are described in greater detail here and referred to the next empirical experiment.

³⁸It appears that this company is no longer solvent. The raw data can be provided, upon request.

³⁹On leap years the price discovery period extends to February 29th

⁴⁰CEPP, the procedure for calculating the projected prices and harvest prices for all crops across all counties, can be found on-line at the RMA’s website: <http://www.rma.usda.gov/policies/cepp.html>.

harvest price (hHP_t) divided by the historical projected price (hPP_t) as follows:

$$HP_t = 5.78 * \frac{hHP_t}{hPP_t} \quad (22)$$

In order to establish dependence between prices and yields at the county level for calibrating the parametric copulas, price deviates (*Price Deviate*) are calculated instead of using nominal values. This is accomplished by dividing the scaled HP_t by the PP_t of the same period. Because prices are assumed to be log-normally distributed, the price deviate is transformed by taking the natural logarithm to generate $\log(\text{Price Deviate})$ calculations. Table 3 summarizes the price information in 2013 dollars.⁴¹

Table 3: Futures Price Data and Deviations from 1969 to 2013

	Min	Q1	Median	Mean	Q3	Max
Projected Price	4.10	4.60	5.48	5.62	6.15	10.03
Harvest Price	3.18	3.57	4.25	4.37	4.78	7.80
Price Deviate	0.73	0.81	0.97	0.99	1.09	1.78
$\log(\text{Price Deviate})$	-0.32	-0.21	-0.03	-0.03	0.08	0.57

Notes: Prices are in 2013 dollars. Projected prices and harvest prices are calculated from the futures data. The Price Deviate is calculated by dividing harvest price observations by projected price observations. The natural logarithm of the Price Deviate is the $\log(\text{Price Deviate})$.

Simulated harvest price realizations are generated using an Itô process. The mean, identifying this process, is the projected price used in this study which is \$5.78. The dispersion of this process is identified by the implied volatility. In order to encompass a range of historical implied volatilities, for corn, this parameter is

⁴¹The distributional properties of the price and price deviates are presented in Figure 33

either 0.20, 0.35, or 0.50.⁴²⁴³

Yields

County level corn yields, as measured by bushels per acre, were collected from National Agricultural Statistic Service (NASS) using the “Quick Stats” interface.⁴⁴ At the state level, Figure 3 shows that the data exhibits a strong positive trend after 1945, which is attributed to the increased use of hybrids and fertilizer after the WWII (Egli, 2008).

The data must be detrended and adjusted for heteroskedasticity before used in either the pure empirical rating procedure or to calibrate the copula’s θ parameter and the dispersion and central tendency of censored normal marginal yield distribution. Without adjusting for trend, the nominal trends in both prices and yields produce a spurious positive relationship.⁴⁵ Accounting for heteroskedasticity is critical for both the empirical rating procedure and the proper calibration of the dependence measure between the price and yield deviates.⁴⁶

The different post-war crop production function warrants excluding yield data prior to 1945. For this subset, only counties with the maximal 69 years of data

⁴²The RMA’s projected price and implied volatility calculations can be found for any crop-location at <https://ewebapp.rma.usda.gov/apps/costestimator/Estimates/QuickEstimate.aspx>.

⁴³The range of implied volatility is based of a historical corn implied volatility chart which can be found at [http://www.barchart.com/chart.php?sym=\\$CIV&t=BAR&size=M&v=0&g=1&p=M0&d=X&q=1&style=technical&template=](http://www.barchart.com/chart.php?sym=$CIV&t=BAR&size=M&v=0&g=1&p=M0&d=X&q=1&style=technical&template=).

⁴⁴These data can be found at <http://quickstats.nass.usda.gov/>.

⁴⁵Also, if producers exhibit positive yield trends and these are not accounted for in the ratemaking procedure, then they will not be able to purchase sufficient insurance because their realized value of production will always be greater than the retrospectively determined liability.

⁴⁶The method for detrending yields and accounting for heteroskedasticity follows that of the Income Protection program, outlined by Atwood, Baquet, and Watts (1996) and described in greater detail in Appendix G.

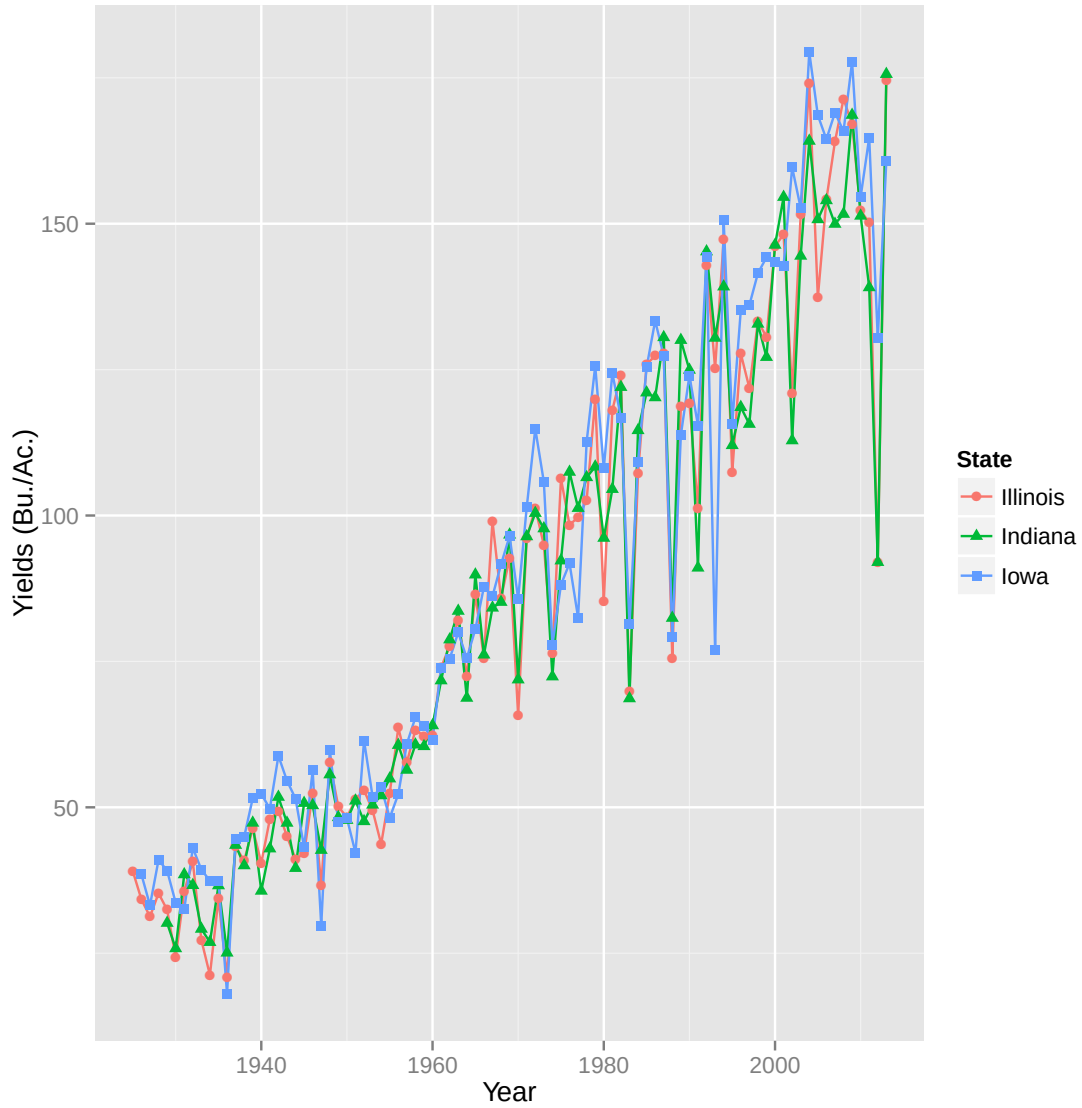


Figure 3: State Level Corn Yield Trends from 1925-2013

(1946-2013) were used in order retain the most robust trend estimates. From the 293 I-state counties with NASS data, 250 had 69 years of observations. A robust linear model is used to estimate the trend for each county. Table 4 illustrates general yield characteristics, dispersion and number of counties by state. Heteroskedasticity was found in 113 counties at the .015 significance level. The state

level variation in yields across time is shown in Figure 34, which indicates both that the large loss events in recent years and that nonlinear detrending methods are not required, which is consistent with the findings of Swanson and Nyankori (1979).

Table 4: Detrended County Yields (Bu./Ac.) from 1945 to 2013

State	#Counties	Min	Q1	Median	Mean	Std	Q2	Max
Illinois	82	19.47	149.76	169.23	165.25	29.48	185.42	269.51
Indiana	69	30.77	148.28	163.86	160.83	23.90	176.78	310.83
Iowa	99	28.17	160.05	175.07	169.93	24.93	185.91	231.23

Notes: This table summarizes the detrended values of all counties with the maximal 69 years of observation between 1945 and 2013. The number of counties is indicated by the column labeled #Counties. Q1 is the lower 25% of observations and Q2 is the upper 25%.

In addition, a county specific farm level deviation was introduced in order to generate farmer level premium rates.⁴⁷ Pulaski County, IL., Dearborn County IN., Lagrange County IN. and Monroe County IN. lacked farm level data; therefore, rates were not estimated for these counties. Moreover, the farm level data for Polk County IA. suffered from a scaling issue and this county was also removed from the rate setting experiment, bringing the total number of counties to 245.⁴⁸

Premium rates are based on farmer level deviations; Table 5 shows that farmer level deviations are greater than county level deviations. Mean farm level yields, which are not presented, are equal to mean county level yields by construction and these can vary by almost 100 bu/ac. County and farm standard deviations are presented; but, the farm level coefficient of variation, which ranges

⁴⁷These data were obtained from an RMA survey (USDA, 1999) and can be found at ftp://ftp.rma.usda.gov/pub/Miscellaneous_Files/yield98/.

⁴⁸It is assumed that producer level deviations have not changed markedly since the collection period for these data.

from 0.12 to 0.36, reflects the variation in the marginal yield distribution that affects premium rate levels. Farm level data are used directly by the pure empirical rating method, while estimates of mean and standard deviation identify the assumed censored-normal marginal yield distribution for the copula methods.

Table 5: Summary of County and Farm Level Corn Yields

	Min	Q1	Median	Mean	Q3	Max
Avg. Yield, County	117.20	157.00	167.82	165.49	177.44	192.79
Std. Yield, County	14.12	19.81	23.08	23.95	27.23	40.16
CV Yield, County	0.09	0.12	0.14	0.15	0.17	0.28
Std. Yield, Farm	23.34	27.31	31.16	31.69	34.63	47.58
CV Yield, Farm	0.12	0.16	0.19	0.20	0.22	0.36

Notes: The farmer and county level yield properties from 1945 to 2013 are summarized for the 245 counties in Illinois, Indiana, and Iowa. The standard deviation for these values is abbreviated Std. and the coefficient of variation is CV. Q1 is the lower 25% of observations and Q2 is the upper 25%.

Price and Yield Dependence

The estimated empirical dependence between prices and yields is another critical source of potential variation determining premium rates. Table 6 shows the rank order dependence measure of Kendall's τ between the county level yields and national prices. At the county level, this dependence is used to calibrate the copula's shape parameter. Archimedean copulas have shape parameters that are a function of Kendall's τ , while the elliptical copula shape parameters are a function of Pearson's ρ . Rank order measures of dependence are likely to prove more reliable, given the non-linearities between prices and yields. Both measures show overall variations from strongly dependent to virtually independent degrees of co-movement between national prices and county yields.

Table 6: Empirical County Yield and Price Dependence

	Min	Q1	Median	Mean	Q3	Max
<i>Kendall's τ</i>	-0.48	-0.33	-0.28	-0.27	-0.22	-0.04
<i>Pearson's ρ</i>	-0.60	-0.48	-0.42	-0.39	-0.33	-0.03

Notes: The county price-yield dependence is summarized for years between 1969 and 2013 for 245 counties in Illinois, Indiana and Iowa. Q1 is the lower 25% of observations and Q2 is the upper 25%.

Design

For this experiment Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates for 245 I-state counties are generated using either the pure empirical method (empP) or a set of copulas. This set of copulas includes the empirical copula (empC), or the parametric copula set, comprised of the Normal copula (N), the Student-t with four degrees of freedom (t4), Frank (F), Clayton with 90 (c90) and 270 (c270) degrees of rotation and Gumbel with 90 (g90) and 270 (g270) degrees of rotation. These copulas are used to bind the simulated price and yield realizations. The price marginal distribution is log-normal and identified by a constant projected price of \$5.78 and its dispersion varies over a set of historically plausible price dispersion of 0.20, 0.35 or 0.50. The yield marginal distribution is a censored normal distribution with the mean and standard deviation estimated from the county yields. The copula shape parameter is calibrated from historical county price and yield dependence. Premium rates for the complete set of coverage levels from 50% to 85% are generated. For each method and county, premium rates are estimated from 100,000 price and yield pairs. In total, this experiment yields 117,600 separate RP and RPHPE premium rates.

Results

First, premium rates generated by the pure empirical method are compared to the copula methods. The pure empirical method is not used in practice and the next section looks at how changes in the underlying factors determining premium rates affect the dispersion between premium rates generated between alternative parametric copulas. Finally, the tendency of premium rates generated by different copulas to deviate from one another is explored.

Copula Compared to Pure Empirical

There are many potential ways to parse the data by different categorical levels. Table 7 displays some of the tendencies of premium rates generated by the alternative methods and copulas at the expense of masking major sources of premium rate variation. Regardless, it is apparent that RP rates are higher than RPHPE rates, which is expected as RP has additional price risk exposure. The pure empirical method produces markedly lower premium rates than any copula method. At this coarse level, it does not appear that premium rates generated by any given copula diverge radically from one another. Of course, the difference between the Gumbel rotated 90° and Clayton rotated 270° — $.20\%$ —over \$20 billion of RP I-state corn, is \$40 million dollars in premiums collected.

Looking more closely at the variation in premium rates by type and across the copula set, the highest premium rates for RP are generated from the Clayton

Table 7: Empirical Revenue Policy Premium Rates

Policy	Method	Min	Q1	Mean	Median	Q2	Max	Std	CV
RP	empP	0.04	0.60	2.31	1.65	3.54	12.12	2.12	91.75
RP	c270	0.03	0.90	3.44	2.37	5.13	16.71	3.21	93.27
RP	g90	0.03	0.94	3.49	2.43	5.18	16.80	3.22	92.12
RP	empC	0.03	0.89	3.55	2.49	5.34	17.03	3.30	92.88
RP	F	0.03	1.03	3.58	2.57	5.30	16.89	3.21	89.74
RP	t4	0.04	1.05	3.59	2.59	5.34	16.85	3.22	89.51
RP	N	0.03	1.01	3.60	2.58	5.38	17.00	3.26	90.49
RP	g270	0.03	1.05	3.65	2.63	5.43	17.07	3.27	89.40
RP	c90	0.03	1.08	3.68	2.67	5.48	17.13	3.27	88.88
RPHPE	empP	0.02	0.38	1.72	1.15	2.62	10.07	1.69	98.36
RPHPE	empC	0.01	0.71	3.03	2.07	4.60	14.23	2.87	94.88
RPHPE	c270	0.02	0.77	3.06	2.06	4.62	14.53	2.89	94.57
RPHPE	g90	0.02	0.80	3.06	2.09	4.61	14.45	2.87	93.57
RPHPE	t4	0.03	0.88	3.11	2.20	4.63	14.32	2.82	90.86
RPHPE	c90	0.01	0.85	3.11	2.18	4.70	14.30	2.82	90.80
RPHPE	g270	0.02	0.85	3.12	2.20	4.71	14.37	2.84	91.19
RPHPE	N	0.02	0.84	3.13	2.20	4.71	14.52	2.88	91.84
RPHPE	F	0.02	0.88	3.14	2.23	4.72	14.44	2.84	90.53

Notes: All values are scaled by 100 and can be interpreted in terms of percentages. The methods used to generate premium rates are the pure empirical (empP), empirical copula (empC) and the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270) copulas. The standard deviation of the premium rates is Std and the coefficient of variation is CV. The lower 25th percentile is Q1 and the upper 75th percentile is Q2.

copula rotated 90° or Gumbel copula rotated 270°. This result is in accord with the general shape of the underlying structure each copula imposes: both copulas fan out towards low prices, which would tend to increase the probability of joint low yield and low price events. The exposure of RP to additional price risk would tend to make it more sensitive to these joint outcomes, which is indicated by the fact that the highest rates for RPHPE do not come out of these copulas, rather they occur for the Normal and the Frank copulas. Conversely, the shape of both the Clayton rotated 270° or Gumbel rotated 90° imposes a tight association of low-yield events

with an off-setting high-price event. Indeed, for both RP and RPHPE, the lowest premium rates come from these two copulas.

These results present some important initial findings and illuminate some avenues to explore for potential premium rate divergence. First, the pure empirical method generates fundamentally different and lower premium rates than copula based methods—including the empirical copula. Second, between any copula average premium rates do not appear to vary radically across all implicit sources of variation. Third, the highest premium rates correspond to copula shapes that increase the probability of low yield-price outcomes. Fourth, the relative order of premium rate central tendency by copula for RP and RPHPE corresponds to the price risk each program faces and the shape of the copula.

When compared to pure empirical premium rates, the decision to use any copula has greater effect on premium rates than the selection of any particular copula; however, several important sources of variation in premium rates have been lumped together: 1) the effects of coverage level, which is expected to increase premium rates and could interact uniquely with different copulas, 2) the unexamined variation in the marginal price and yield distributions and 3) the dependence between the marginal distributions.

Linear regression allows some of these complex possibilities to be distilled and presented. Tables 8 and 9 present the results of a linear regression of RP and RPHPE premium rates regressed onto the major independent sources of variation: the copula or rating procedure, the coverage level, the dependence as measured by

Kendall's τ , the variation in prices, and the coefficient of variation in farmer yields.

The reference condition, as represented by the intercept, is a Normal copula at the median coverage level of 65% and the median price dispersion of 0.35.⁴⁹

Table 8: LRM of RP Premium Rates

	Estimate	Std. Error	t value	Pr(> t)
Intercept	-0.01	2.99×10^4	-43.57	$< 2 \times 10^{16}$
Pure Empirical Method	-0.01	2.03×10^4	-61.92	$< 2 \times 10^{16}$
Empirical Copula Method	-4.33×10^4	2.03×10^4	-2.13	0.03
Student-t (df=4) Copula Method	-8.09×10^5	2.03×10^4	-0.40	0.69
Frank Copula Method	-2.11×10^4	2.03×10^4	-1.04	0.30
Gumbel 90° Copula Method	-1.16×10^3	2.03×10^4	-5.69	1.29×10^8
Gumbel 270° Copula Method	5.64×10^4	2.03×10^4	2.78	0.01
Clayton 90° Copula Method	9.02×10^4	2.03×10^4	4.44	9.10×10^6
Clayton 270° Copula Method	-1.54×10^3	2.03×10^4	-7.59	3.37×10^{14}
50% Coverage Level	-0.02	1.92×10^4	-84.02	$< 2 \times 10^{16}$
55% Coverage Level	-0.01	1.92×10^4	-64.34	$< 2 \times 10^{16}$
60% Coverage Level	-0.01	1.92×10^4	-36.93	$< 2 \times 10^{16}$
70% Coverage Level	0.01	1.92×10^4	48.46	$< 2 \times 10^{16}$
75% Coverage Level	0.02	1.92×10^4	110.63	$< 2 \times 10^{16}$
80% Coverage Level	0.04	1.92×10^4	188.70	$< 2 \times 10^{16}$
85% Coverage Level	0.05	1.92×10^4	284.29	$< 2 \times 10^{16}$
0.20 Price Dispersion	-0.01	1.17×10^4	-104.43	$< 2 \times 10^{16}$
0.50 Price Dispersion	0.02	1.17×10^4	197.20	$< 2 \times 10^{16}$
Kendall's τ	0.03	6.00×10^4	56.45	$< 2 \times 10^{16}$
Farm Yield CV	0.22	1.05×10^3	213.71	$< 2 \times 10^{16}$

Notes: This tables provides the results of a simple linear regression model (LRM) of Revenue Protection (RP) premium rates regressed onto the generating copula or pure empirical method, the complete set of coverage levels, price dispersion, dependence and yield variation. The reference condition is for a Normal copula, rated at the 65% coverage level and with a price variation of .35. The coefficient of variation is abbreviated as CV. The Student-t copula is calibrated with 4 degrees of freedom.

Besides a negative intercept term, which can be explained as an artifice of the linear estimation process, the signs of all the coefficients are in accord with theory. The coefficient for the measure of farm level yield variation is positive, indicating that as yield variation increases so too do premium rates and the coefficient is larger

⁴⁹The Normal copula was selected because the RMA uses it to establish price-yield dependence.

Table 9: LRM of RPHPE Premium Rates

	Estimate	Std. Error	t value	Pr(> t)
Intercept	-3.18×10^{03}	2.84×10^{04}	-11.19	$< 2 \times 10^{16}$
Pure Empirical Method	-0.01	1.94×10^{04}	-71.63	$< 2 \times 10^{16}$
Empirical Copula Method	-1.02×10^{03}	1.94×10^{04}	-5.26	$< 2 \times 10^{16}$
Student-t (df=4) Copula Method	-2.55×10^{04}	1.94×10^{04}	-1.32	$< 2 \times 10^{16}$
Frank Copula Method	6.93×10^{05}	1.94×10^{04}	0.36	0.72
Gumbel 90° Copula Method	-7.59×10^{04}	1.94×10^{04}	-3.92	$< 2 \times 10^{16}$
Gumbel 270° Copula Method	-1.10×10^{04}	1.94×10^{04}	-0.57	0.57
Clayton 90° Copula Method	-1.78×10^{04}	1.94×10^{04}	-0.92	0.36
Clayton 270° Copula Method	-7.42×10^{04}	1.94×10^{04}	-3.83	1.26×10^{04}
50% Coverage Level	-0.01	1.82×10^{04}	-76.96	$< 2 \times 10^{16}$
55% Coverage Level	-0.01	1.82×10^{04}	-58.83	$< 2 \times 10^{16}$
60% Coverage Level	-0.01	1.82×10^{04}	-33.65	$< 2 \times 10^{16}$
70% Coverage Level	0.01	1.82×10^{04}	43.65	$< 2 \times 10^{16}$
75% Coverage Level	0.02	1.82×10^{04}	98.83	$< 2 \times 10^{16}$
80% Coverage Level	0.03	1.82×10^{04}	166.96	$< 2 \times 10^{16}$
85% Coverage Level	0.05	1.82×10^{04}	248.96	$< 2 \times 10^{16}$
0.20 Price Dispersion	-0.01	1.12×10^{04}	-101.43	$< 2 \times 10^{16}$
0.50 Price Dispersion	0.02	1.12×10^{04}	203.35	$< 2 \times 10^{16}$
Kendall's τ	0.04	5.71×10^{04}	68.33	$< 2 \times 10^{16}$
Farm Yield CV	0.17	9.99×10^{04}	165.66	$< 2 \times 10^{16}$

Notes: This tables provides the results of a simple linear regression model (LRM) of Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates regressed onto the generating copula or pure empirical method, the complete set of coverage levels, price dispersion, dependence and yield variation. The reference condition is for a Normal copula, rated at the 65% coverage level and with a price variation of .35. The coefficient of variation is abbreviated as CV. The Student-t copula is calibrated with 4 degrees of freedom.

in magnitude for RP than RPHPE. The coefficient for dependence, as measured by Kendall's τ , has a positive sign indicating that as the relationship between prices and yields becomes more positive or weaker, premium rates increase. As coverage level increases premium rates increase, but the magnitude of this change is larger for RP than RPHPE. As the dispersion in the price marginal distribution increases, premium rates increase, but the magnitude is larger for the RP than RPHPE. This difference in magnitude seems logical because the structure of the RP program would make it more sensitive to changes in the price marginal distribution.

Comparing between copulas, for both RP and RPHPE, the premium rates generated out of the Student-t with 4 degrees of freedom and Frank copulas are not significantly different than those of the Normal copula, which may corresponds to their symmetric shapes. For both revenue policies, at the reference coverage level and price dispersion, the pure empirical and empirical copula methods generate lower premium rates. The magnitude of these coefficients is greater for RPHPE than RP. The difference in premium rates from the Normal copula when Archimedean copulas are considered increases. Both coefficients are negative for the Gumbel rotated 90° and Clayton rotated 270° copulas, but the magnitude of the RP coefficient is larger than the RPHPE. More extreme differences between the two insurance types appear for Gumbel rotated 270° and Clayton rotated 90° copulas which have negative and significant RP coefficients, but these coefficients are positive and insignificant for RPHPE.

This analysis is limited to a comparison of the Normal copula to the alternatives; however, it appears that RP and RPHPE premium rates diverge if a copula reflective of prices and yields becoming less dependent as prices decrease is imposed (i.e. the Gumbel rotated 90° and Clayton rotated 270° copulas). That the copulas interact singularly with the marginal distributions prompts an investigation of this interaction, disregarding their assumed distribution and focusing solely on variation in their parameter identification.

In order to better see how premium rates diverge across all methods, Figure 4 shows premium rates for RP and RPHPE for each copula against all others using

a pairs plot. The dashed 45° line represents equality between premium rates between generated by the paired alternative methods.

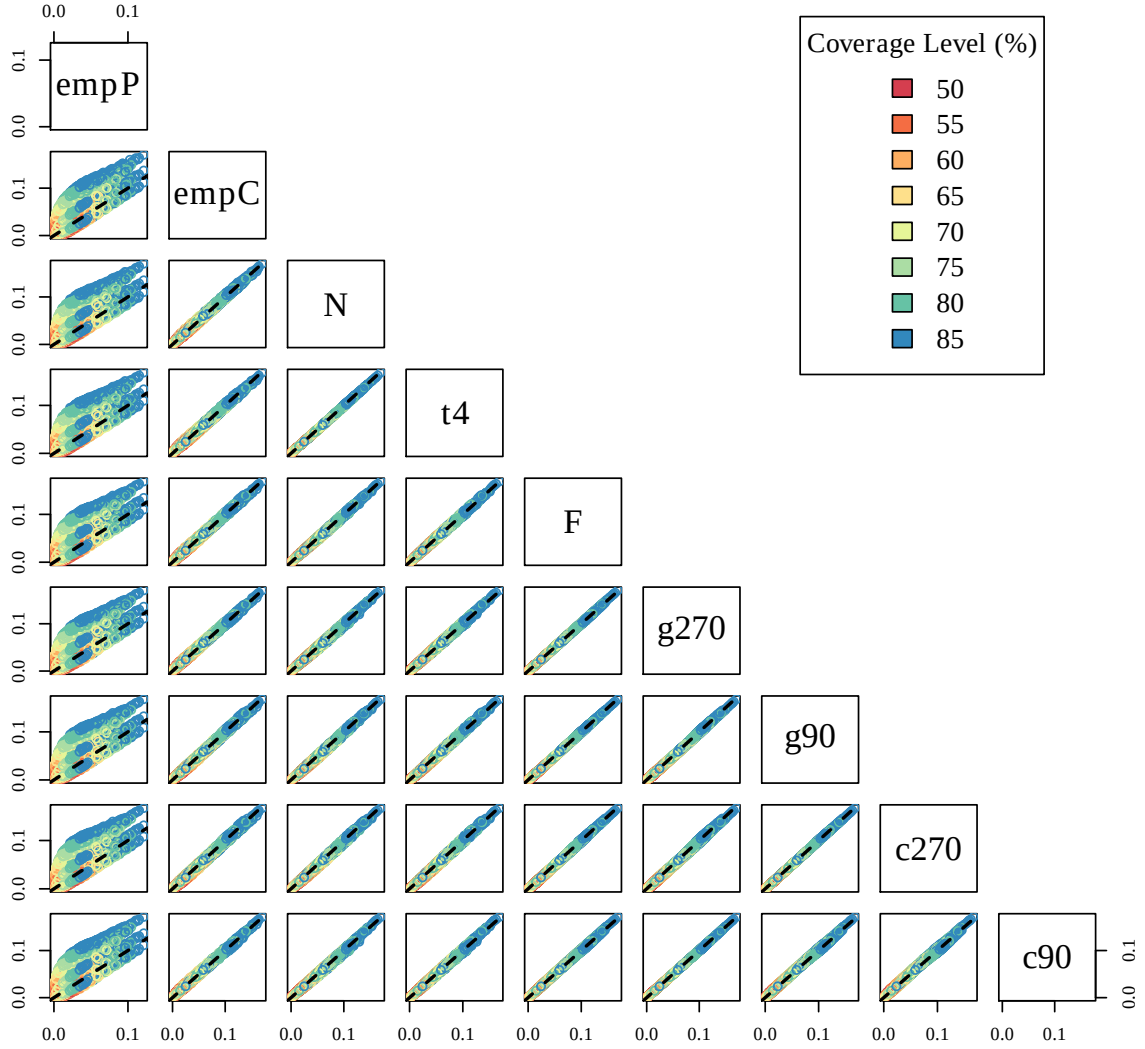


Figure 4: Empirical Revenue Premium Rate Pairs

Notes: This figure includes both Revenue Protection and Revenue Protection with Harvest Price Exclusion premium rates calculated for most Illinois, Indiana and Iowa corn producing counties: when separately identified the figures are not perceptibly different at this scale. Premium rates generated are plotted by method of generation for each county. The methods used to generate premium rates are the pure empirical (empP), empirical copula (empC) and the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270) copulas. The dashed 45° line represents equality between premium rates pairs.

In general, pure empirical premium rates are lower than the those generated by any copula over the same set of parameters; at this scale, there does not appear to be significant deviation or dispersion between copula rates. The figure also indicates that while the pure empirical premium rates tend to be lower than premium rates for all copulas, there are systematic instances across all copulas—including the empirical copula—in which copulas will produce lower premium rates. Possibly, price dispersion drives this divergence because it is the only source of variation that does not directly affect, both pure empirical and copula premium rates.

In order to explore the effect of price dispersion, Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates generated by copulas were collapsed to an average copula rate for each county, at each coverage level and the three levels of price dispersion. Figures 5 and 6, show the impact of variation in the global estimate of price dispersion on RP and RPHPE copula premium rates relative to the rates of the pure empirical method, using a pairs-like plot. Variation in price dispersion does appear to drive the divergence between premium rates. Low price dispersion induces copula generated premium rates to be lower than the premium rates of the pure empirical method. However, at sufficiently low coverage levels copula premium rates can be lower—even at higher levels of price dispersion. This implies that combinations of the remaining empirical variation in yield dispersion and price-yield dependence can induce lower copula generated premium rates. This result indicates the complexity with which all

premium rate driving factors: the insurance policy, coverage level, dispersion in the price and yield marginal distributions and their dependence can interact to induce premium rate divergence.

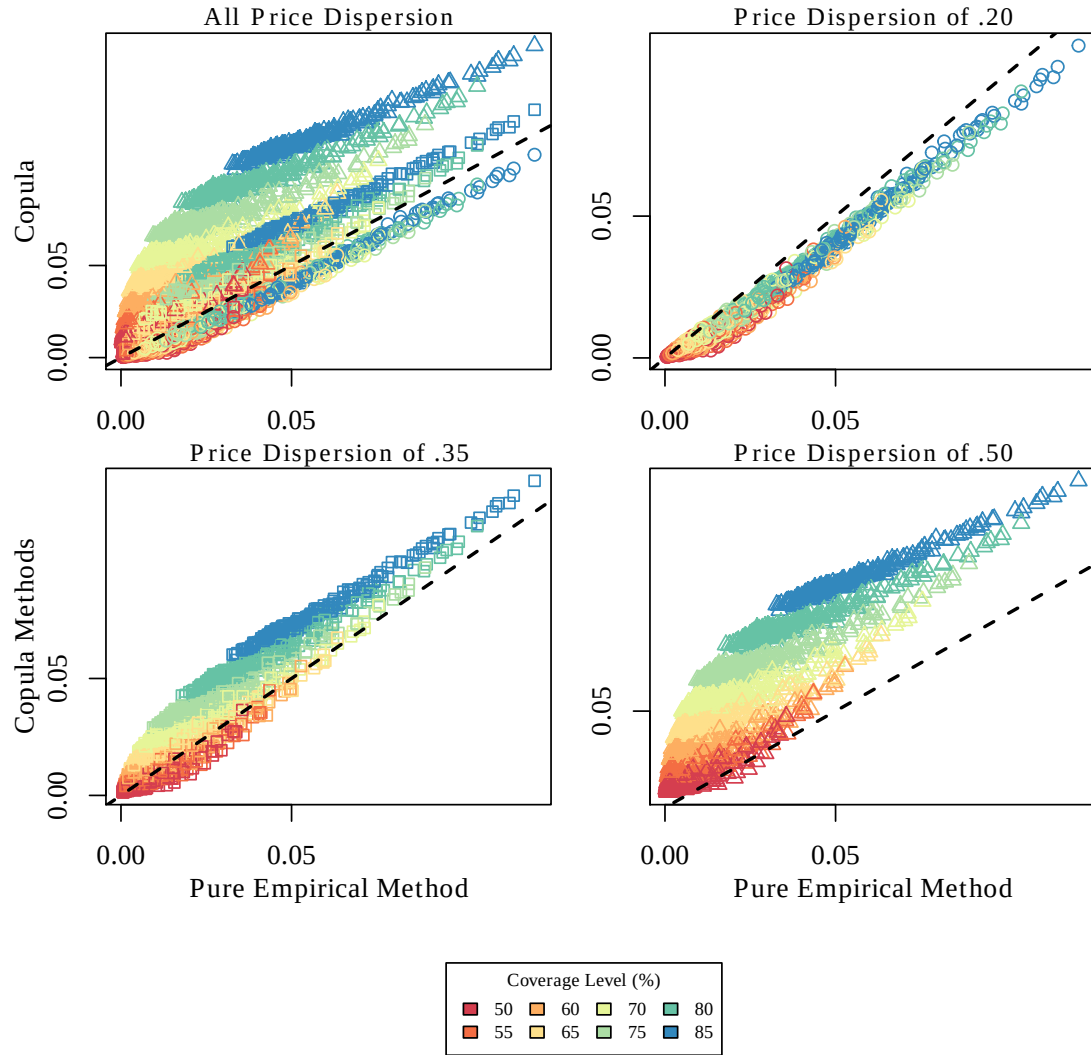


Figure 5: Copula and Pure Empirical RP Premium Rates

Notes: This figure shows the Revenue Protection (RP) premium rates calculated for most Illinois, Indiana, and Iowa corn producing counties generated by the pure empirical method against the average copula premium rate at the three levels of price dispersion. The dashed 45° line represents equality between premium rates pairs.

Tables 29 and 30 numerically summarize the results. During low price

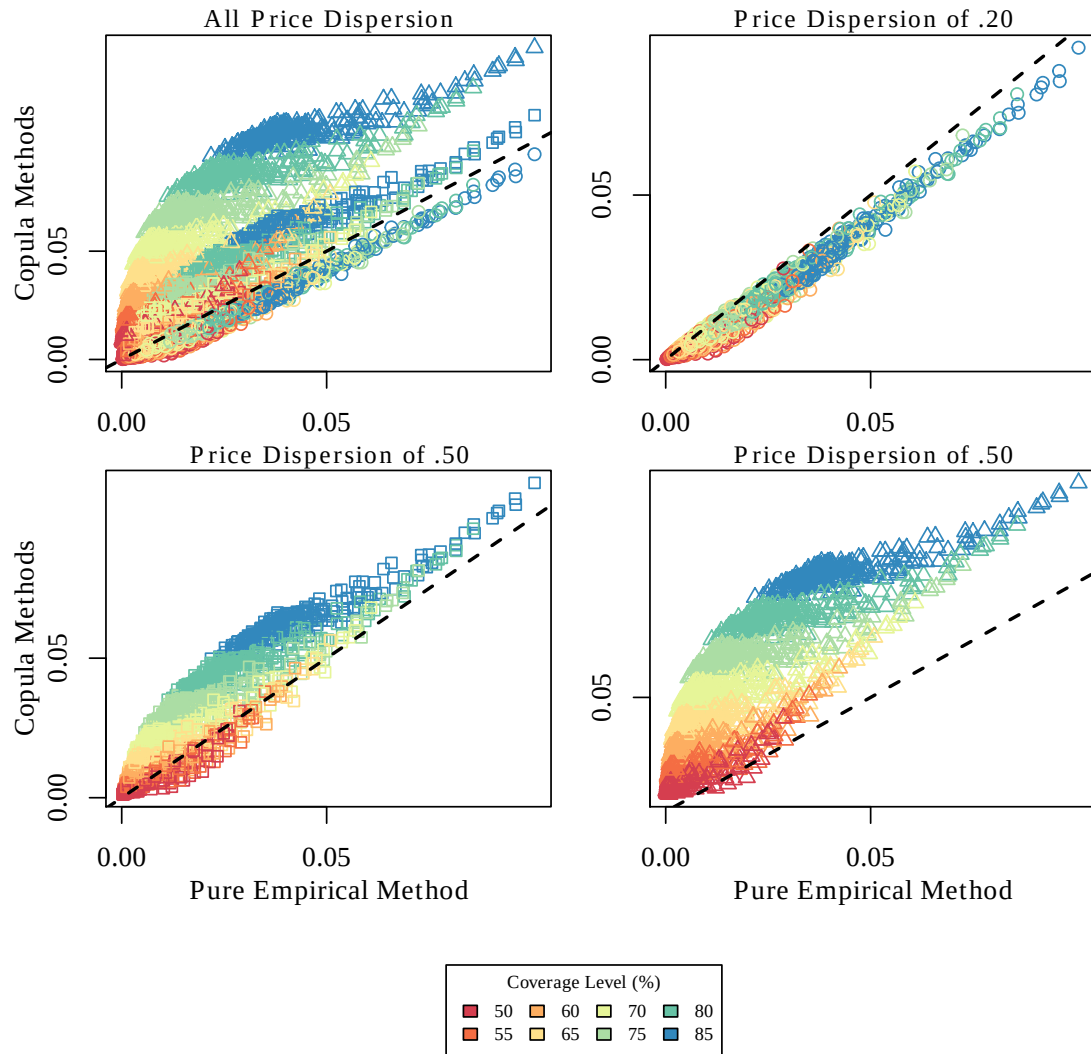


Figure 6: Copula and Pure Empirical RPHPE Premium Rates

Notes: This figure shows the Revenue Protection with Harvest Price Protection (RPHPE) premium rates calculated for most Illinois, Indiana, and Iowa corn producing counties generated by the pure empirical method against the average copula premium rate at the three levels of price dispersion. The dashed 45° line represents equality between premium rates pairs.

dispersion rating periods, the average copula premium rate will be lower than the pure empirical premium rate rate for all coverage levels. The difference premium rates generated by the methods during high price dispersion indicates that the average copula premium rate will be higher and at the highest coverage level it is

6% percentage points greater than the pure empirical premiums, which would amount to a \$120 million dollar difference in premiums collected for the \$20 billion dollars of RP liability across Illinois, Indiana and Iowa in 2013.

Given the limited empirical variability in the dispersion of yields and price-yield dependence, it is difficult to identify how these sources of variation affect copula premium rates relative to the premium rates generated by the pure empirical method. The results from this investigation suggest that high yield variation counties might generate higher average copula premium rates than the pure empirical method. Also, low dependence counties might tend to generate higher average copula premium rates than the pure empirical method because at low levels of dependence, except for the Student-t, all copulas in this study approach the Independence copula, which may be less structured than the empirical data. Future research might explore multiple crops to increase the empirical variation and examine these conjectures; however, the subsequent simulation and final empirical experiment provide some insight into several of these issues.

Dispersion Between Copulas

While interesting to investigate, the pure empirical method is not currently used for rating multiple peril crop insurance (MPCI). Instead, dependence is imposed on price-yield pairs by the equivalent of a specific copula—the Normal. The previous findings indicate that the dispersion of the price marginal distribution can affect the average bias of all copula generated premium rates, at least relative to the

pure empirical method. Thus, it is logical to investigate the effect of price dispersion on premium rate dispersion between alternative copulas, as well as the remaining sources of variation in price-yield dependence, as measured by Kendall's τ and the coefficient of variation for farmer yields.

A simple way to represent the potential for premium rate dispersion is the coefficient of variation between premium rates generated across the alternative copulas for each county, by coverage level and price dispersion. This represents the variation in premium rates between the copulas in the copula set for a county at a given coverage level and price dispersion rating environment. The coefficient of variation proxies for potential premium rate misspecification by representing the dispersion between premium rates generated by equally viable alternative copulas.

Figures 7 and 8 plot the coefficient of variation for RP and RPHPE premium rates generated by alternative copulas for each county by grouping the data into five quantiles of farm level yield variation along the horizontal axis.⁵⁰ The three levels of price dispersion divide the figure into subplots vertically. Within each of these subplots, the coefficient of premium rate variation is mapped against Kendall's τ .

RP and RPHPE premium rates converge as coverage level increases and dependence becomes weaker. It appears that as price dispersion increases, premium rates tend to converge, but this pattern does not strictly hold across each yield variation group. Similarly, as yield variation increases premium rates tend to converge, but this does not hold across all levels of price dispersion. Moreover, for

⁵⁰These quantiles range from the 49 counties with the lowest variation to the 49 counties with the highest.

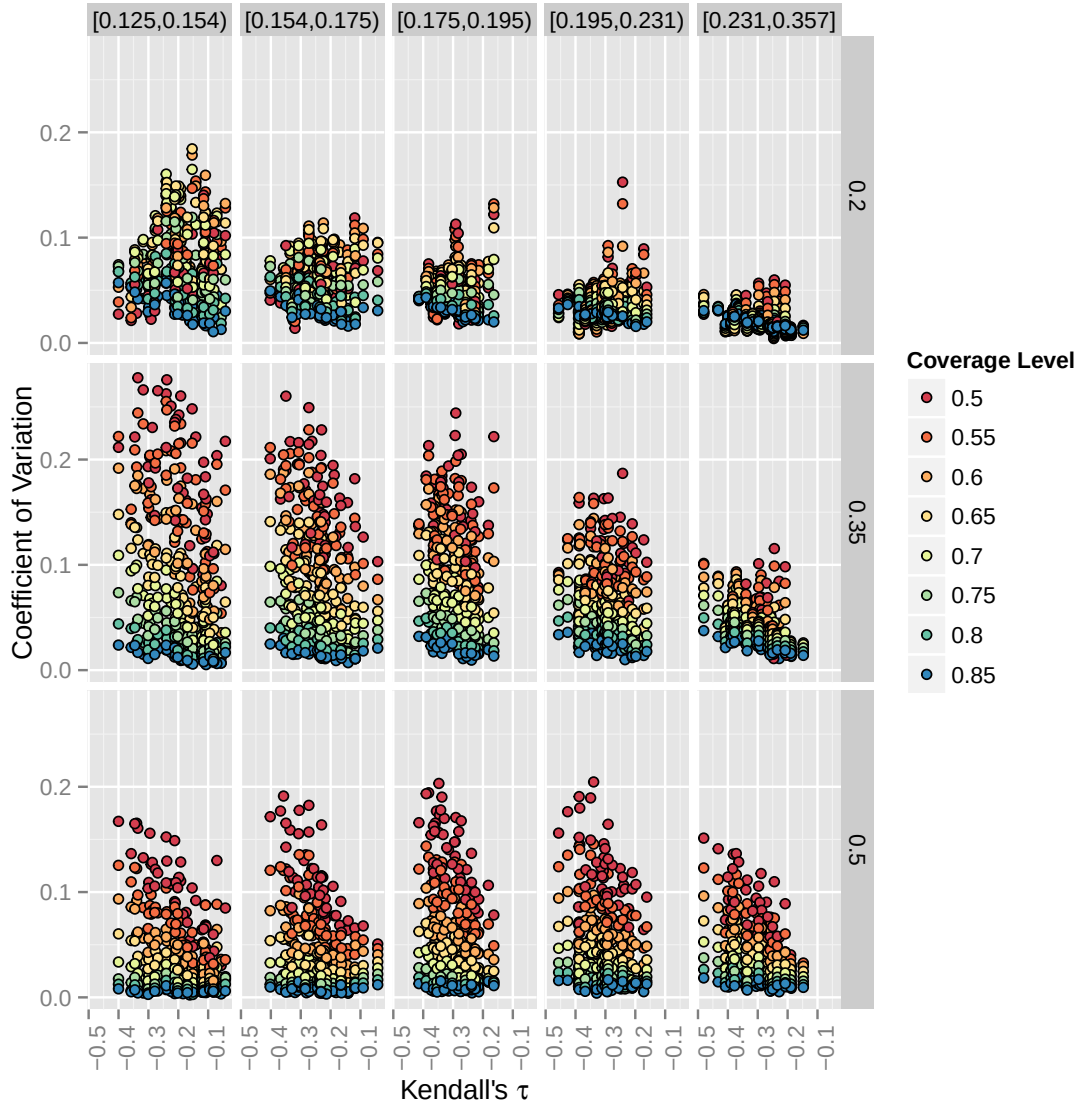


Figure 7: Empirical RP Premium Rate Dispersion

Notes: This figure maps the coefficient of variation between Revenue Protection (RP) premium rates generated by alternative copulas for most Illinois, Indiana and Iowa corn producing counties onto Kendall's τ . The figure is divided into subplots by grouping the data into five quantiles of the coefficient of variation for farm level yields along the horizontal axis and the three levels of price dispersion on the vertical axis, which creates nine unique categorical combinations.

both insurance types, the coefficient of variation increases as a function of

dependence for the lowest yield variation counties. Of course, the results presented

by these plots are somewhat muddled given the range of potential variation in yields

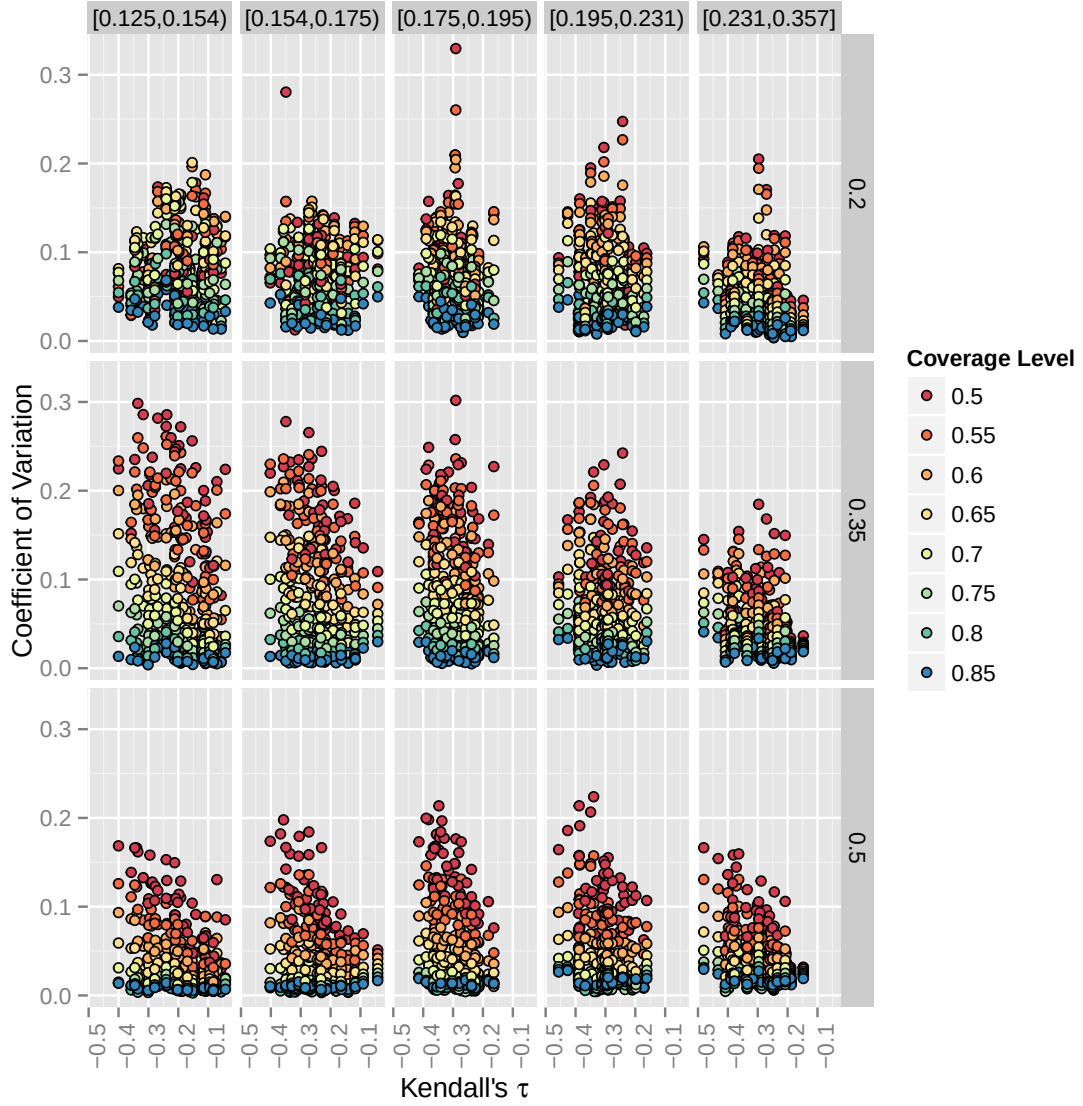


Figure 8: Empirical RPHPE Premium Rate Dispersion

Notes: This figure maps the coefficient of variation between premium rates generated by alternative copulas for most Illinois, Indiana and Iowa corn producing counties onto Kendall's τ . The figure is divided into subplots by grouping the data into five quantiles of the coefficient of variation for farm level yields along the horizontal axis and the three levels of price dispersion on the vertical axis, which creates nine unique categorical combinations.

possible at each level of dependence.

Clearly, the dispersion of the marginal price and yield distributions, their dependence and the coverage level can interact in complex ways that are obfuscated

by the empirical data. Greater clarity is achieved in a subsequent simulation experiment, which uniquely identifies all rate determining factors. Moreover, these figures indicate the limitations of the empirical data to effectively populate some combinations of the rate determining factors. For example, there are few counties characterized by strong dependence and low yield variation or weak dependence and high yield variation. These limitations are addressed by the simulation experiment, which support broader conclusions and provide context for a copula based rating procedure for other crop-county revenue insurance policies.

Deviations Between Copulas

While the coefficient of variation proxies for premium rate dispersion and the potential for premium rate imprecision, it comes at the expense of identifying the direction of deviation between alternative copulas. In order to eliminate scale from the premium rates and examine both the direction and magnitude of premium rate divergence across coverage level and the alternative copulas, a deviation metric is calculated.⁵¹ The premium rates generated by each copula for a given county, at each level of price dispersion and coverage level were divided by the mean premium rate generated by the alternative copulas. Given the current use of the Normal copula it is perhaps most important to investigate the direction of rate bias relative to that copula. It is also informative to examine alternative copula combinations. If, for example, the rating procedure might be improved by assuming a copula that

⁵¹The motivation behind the development of this deviation metric is discussed in Appendix H.

reflects an alternative theoretical understanding of the dependence between prices and yields, then understanding the potential for bias would provide context to this decision.

Figures 9 and 10 show the deviation between RP and RPHPE premium rate pairs from the mean premium rate for a county by common price dispersion. The most striking result is that dispersion and bias from the mean copula rate between pairs increases as coverage level decreases. Premium rate deviation is greatest at a coverage level of 50%, but non-trivial at coverage level of 65%. Conversely, as coverage level increases premium rates converge. Examining the figures more closely, the fatter tails of the Student-t copula⁵² appear to cause significant deviations from the mean rate, relative to the Normal copula at low coverage levels, but at high coverage levels deviations appear systematically biased below the average premium rate for both insurance policies.⁵³

It is also interesting to observe the relationship between premium rate deviation amongst the asymmetric Archimedean copulas because these copulas could be used to impose a radically different form of the dependence between prices and yields. For example, between the two copulas that represent a price-yield relationship where dependence becomes less structured as prices fall (i.e. the Gumbel rotated 270° and Clayton rotated 90° copulas) the overall deviation from the mean premium rate tends to be quite similar; however, the Clayton rotated 90°

⁵²The Student-t copula is calibrated with 4 degrees of freedom.

⁵³A similar type of bias switching is evident in comparison of the Student-t copula with the Gumbel rotated 270° and Clayton rotated 90° copulas. The Frank copula also appears to show this behavior, especially for RP premium rate deviations.

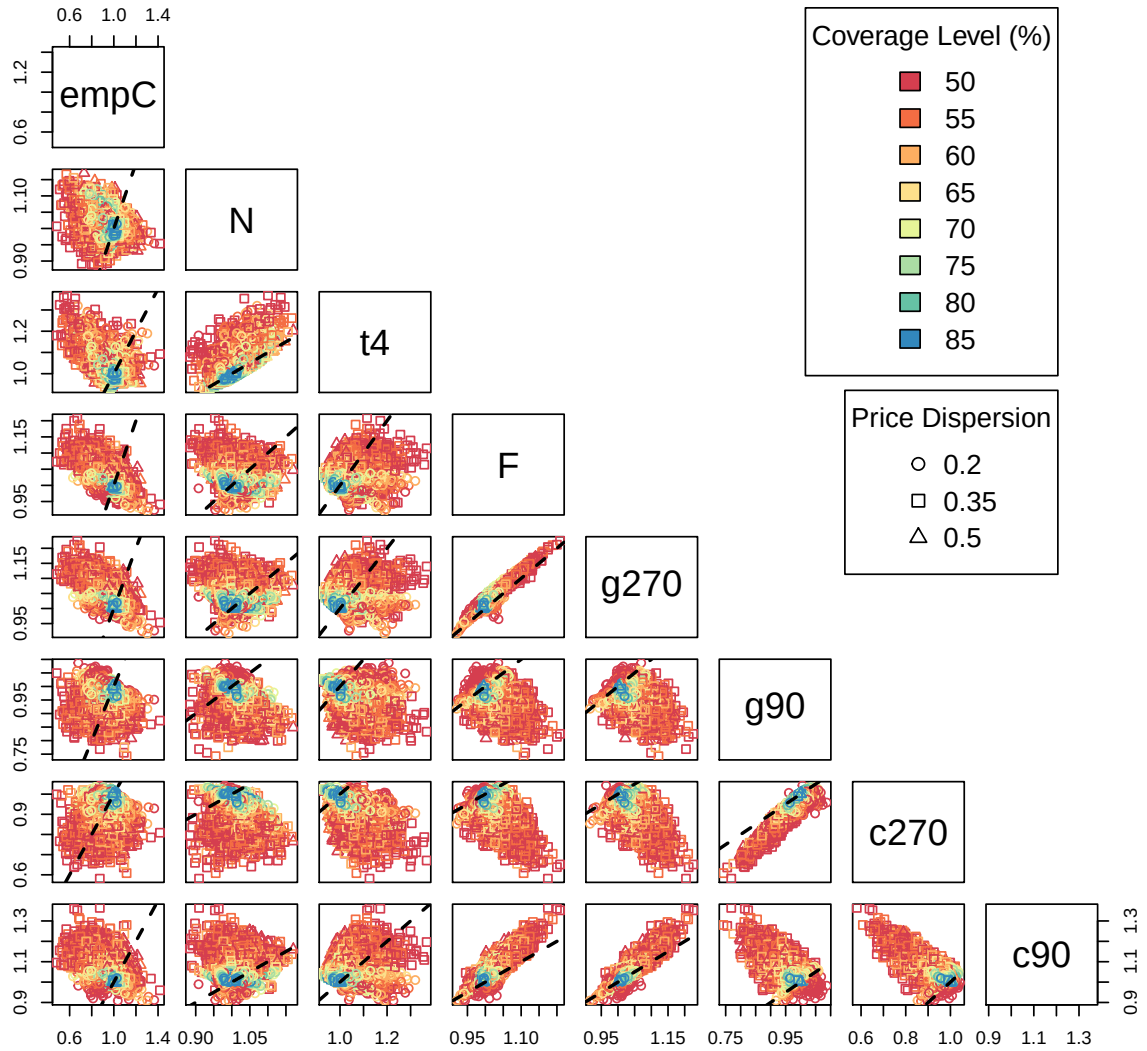


Figure 9: Between Copulas Copula RP Deviations

Notes: This figure shows Revenue Protection (RP) premium rate deviations, which are calculated by dividing premium rates by the average premium rate for a county at each combination of coverage level and price dispersion, between premium rate generation methods for most Illinois, Indiana and Iowa corn producing counties. The methods used to generate premium rates are the pure empirical (empP), empirical copula (empC) and the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270) copulas. The dashed 45° line represents no deviation from the county average.

premium rates will tend to deviate more from the the mean premium rate than the Gumbel rotated 270° rates. This may be due to the fact that for most counties, the Clayton copula will impose a stronger tail dependence than the Gumbel.

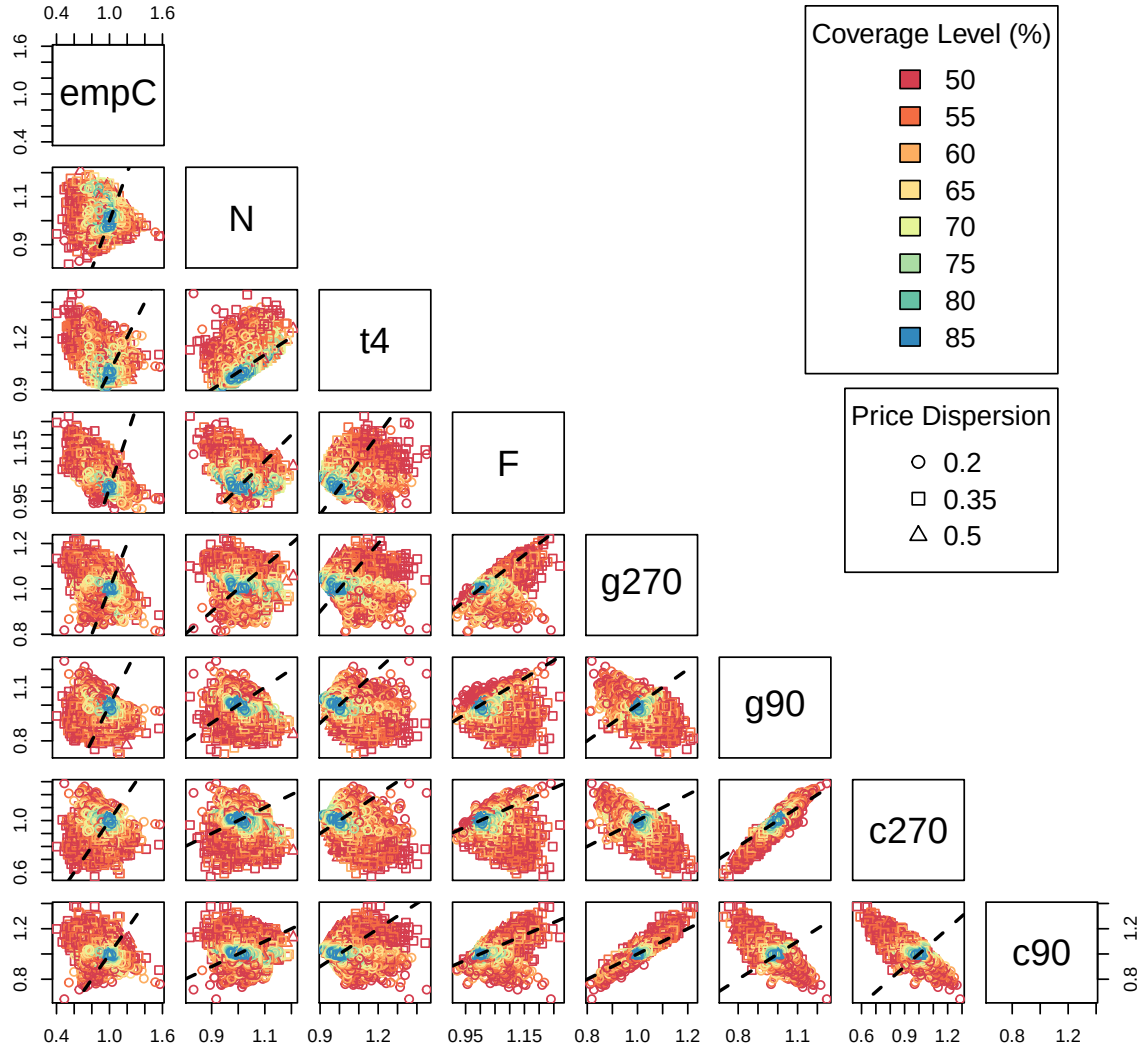


Figure 10: Between Copulas Copula RPHPE Deviations

Notes: This figure shows Revenue Protection with Harvest Price Exclusion (RPHPE) premium rate deviations, which are calculated by dividing premium rates by the average premium rate for a county at each combination of coverage level and price dispersion, between premium rate generation methods for most Illinois, Indiana and Iowa corn producing counties. The methods used to generate premium rates are the pure empirical (empP), empirical copula (empC) and the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270) copulas. The dashed 45° line represents no deviation from the county average.

Also it is notable that for RP deviations there is “tab” of premium rate deviations at the lowest coverage level, which are the only observations above or below the dashed line of premium rate equality. These deviations tend to come from

the lowest price dispersion rating environment. Thus, behavior identified for premium rates generated by copulas at the lowest coverage level are not constant across variation in the price marginal distribution. Also, while difficult to see, the symmetric deviations between the asymmetric Archimedean copulas for RPHPE are generally comprised of 0.35 and 0.50 levels of price dispersion on one side of the dashed line and premium rates generated at a price dispersion of 0.20 on the other; however, unlike RP this applies to higher coverage levels not just the lowest.

When comparing the Gumbel rotated 270° premium rates against those copulas that impose the opposite shape, the Gumbel rotated 90° and Clayton rotated 270° , the deviation between these premium rates from the mean is virtually perpendicular. Moreover, the direction of bias switches when comparing the Gumbel rotated 270° to the Gumbel rotated 90° versus the Clayton rotated 270° .⁵⁴ Imposing an asymmetric shape on the multivariate price-yield distribution through an asymmetric Archimedean copula, would induce significant additional potential premium rate bias, if the true structure is its opposite and the direction of this deviation would change between high and low price dispersion rating periods.

Summary

The first result from this analysis is that premium rates generated by copulas differ from those generate by the pure empirical method. On average, they are higher, but during low price dispersion rating periods they will be lower. An

⁵⁴A similar analysis can be performed for the remaining asymmetric Archimedean copulas.

interesting course of future research would be to incorporate price dispersion into the pure empirical ratemaking method and examine the stability of premium rates generated by this process compared to copula generated premium rates. Across the copula set, estimated premium rates tend to be more similar than different on average. However, closer inspection showed that dispersion between copula premium rates increases at low coverage levels, high price-yield dependence, and low yield variation. Premium rate dispersion was greatest between copulas at the median level of price dispersion.

An attempt to consider the underlying direction of this dispersion showed that coverage level, the copulas of comparison and price dispersion can all drive deviation in a non-linear fashion. However, low levels of price dispersion and coverage levels can produce radically different directions of premium rate deviation from the mean premium rate between alternative copula pairs, when compared to higher levels of price dispersion and coverage levels. This fact, in conjunction with the general better behavior of the Normal copula, in comparison to the alternatives, perhaps lends some support to its selection over the alternatives (unless a compelling theoretical case can be made for an alternative higher order structure).

The copula, the level of dependence identifying the copula and the parameters identifying the marginal distributions affect the potential for premium rate dispersion and deviation between copulas. The potential impact of the parameters identifying the dispersion of the underlying distributions must be placed in context. The fact is these distributions are themselves assumed. Perhaps the

most important implication of these results is that the marginal price and yield distributions matter and the once lively debate over their properties may need to be revisited in light of a copula based rating procedure. Even simple variation in the shape of assumed distributions interacts with the overarching copula and coverage level to generate premium rate deviations. Variation along a more expansive set of dispersion and dependence characteristics will be explored in the next experiment. The potential for additional premium rate misspecification would likely increase, if a more expansive set of marginal distributions were considered in conjunction with the alternative copulas. The impact of alternative marginal assumptions—particularly the assumed censored normal yield distribution—on copula generated premium rates might prove a fruitful avenue of exploration.

CHAPTER EIGHT

SIMULATED PREMIUM RATE DISPERSION BETWEEN COPULAS

Limitations in the empirical data warranted a simulation experiment to more thoroughly explore the interaction of the independent sources of premium rate variation across a broader spectrum.⁵⁵ In this experiment, Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates vary as a function of the the copula binding structure, the coverage level, the dispersion parameters for both the price and yield marginal distributions and the imposed dependence between the two marginal distributions. The potential for premium rate dispersion is represented by the coefficient of variation between premium rates generated between alternative copulas at each categorically unique intersection of the sources of premium rate variation.

The cost of summarizing the distributional properties of premium rates with the coefficient of variation is that it does not allow the direction and magnitude of the dispersion to be examined between copulas, nor across coverage level. A deviation measure is used to facilitate the investigation of premium rate bias and its evolution. While initial investigations of the empirical data motivated this experiment, the simulation results are more general and allow the empirical corn results to be framed within a broader context that sheds light on the potential impact of copula rating across the multiple peril crop insurance book of business.

⁵⁵There was insufficient variation in empirical yield dispersion and price-yield dependence to adequately identify the effect of interactions between these two sources of variation.

Design and Data Generation

This experiment examines the impact on RP and RPHPE premium rates generated by seven parametric copulas⁵⁶ of four unique sources premium rate variation.⁵⁷ Like the previous empirical experiment premium rates are estimated from 100,000 correlate price-yield pairs at eight coverage levels between 50% and 85%. Price realizations are generated using an Itô process. The mean identifying this process is \$5.78. The dispersion in prices is identified by nine the implied volatilities, which range from 0.10 to 0.50. Yield realizations are generated from an assumed censored normal marginal distribution. The mean of this distribution can be considered an arbitrary scaling factor and is held constant at 100 bu./ac., making interpretation of the relative yield variation more tractable. The dispersion of the yield distribution is identified by 9 yield standard deviations, which varies from 5 to 45 bu./ac. Yield variation is easily interpreted as a coefficient of variation ranging from 5% to 45% of the mean. Dependence between prices and yields is identified by eleven measures of Kendall's τ , which ranges between -0.55 and -0.05. RP and RPHPE premium rates are generated by each of the seven parametric copulas at all of the 7,128 unique combinations of these premium rate determining factors, providing a broader range of possible variation that is precisely identified.⁵⁸

⁵⁶The parametric copulas in this experiment are the same as the previous experiment and consist of the the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270) copulas.

⁵⁷Greater detail of the ratemaking process is provided in the Ratemaking Procedures section of the first empirical experiment

⁵⁸This experiment generates 49,896 RP and RPHPE premium rates.

Results

As in the previous empirical experiment, results for dispersion between premium rates generated by alternative copulas is presented first. The data is fully simulated and not empirical, therefore the pure empirical or empirical copula methods are not explored. Next the potential for deviation in premium rates estimated from the alternative copulas is examined.⁵⁹

Dispersion Between Copulas

The coefficient of variation between premium rates generated for the copula set is calculated at each combination of yield variation, price dispersion and price-yield dependence. Figures 11 and 12 show RP and RPHPE premium rate coefficients of variation mapped onto Kendall's τ that are separated into subplots for unique combinations of yield variation along the horizontal axis and price dispersion along the vertical axis.

While it is clear that dispersion can evolve in a non-linear manner as a function of the underlying combination of the premium rate determining factors, it is equally clear that the potential for premium rate imprecision is greatest for low yield variation groups with strong negative price-yield dependence, during low price dispersion periods. Moreover, premium rate dispersion between copulas is exacerbated by low coverage level election. High coverage level election is the single

⁵⁹The results of the empirical study for corn are compared over the common range of rate determining factors in the simulation study and presented in Appendix J.

most powerful factor inducing convergence across all remaining combinations of premium rate determining factors.

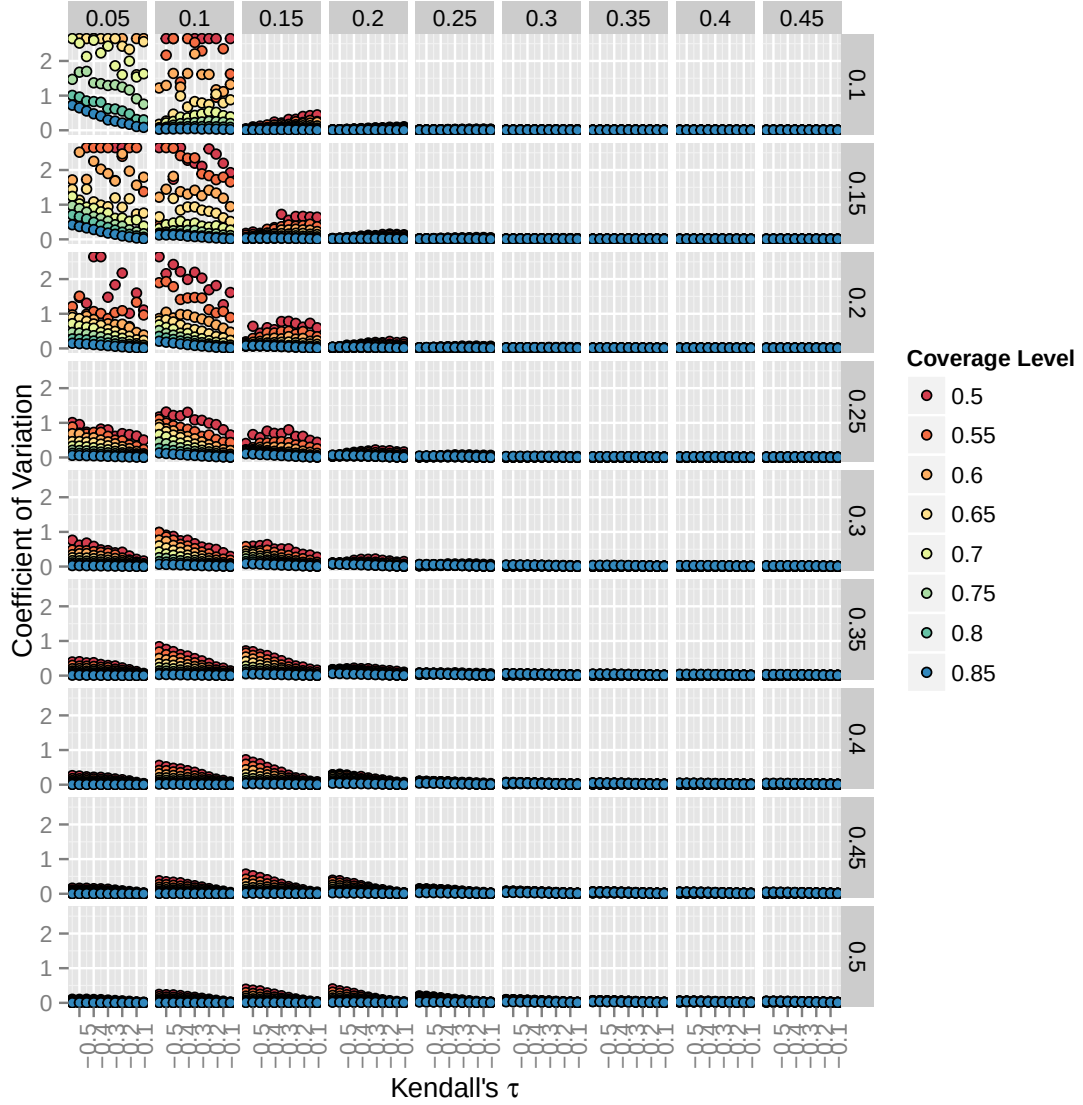


Figure 11: Between Copulas Simulated RP Dispersion

Notes: This figure maps the coefficient of variation between Revenue Protection (RP) premium rates generated by alternative copulas across an extensive set of variation in price dispersion and yield variation onto Kendall's τ . The figure is divided into subplots by grouping the data by the coefficient of variation for farm level yields along the horizontal axis and price dispersion on the vertical axis, which creates 81 unique categorical combinations.

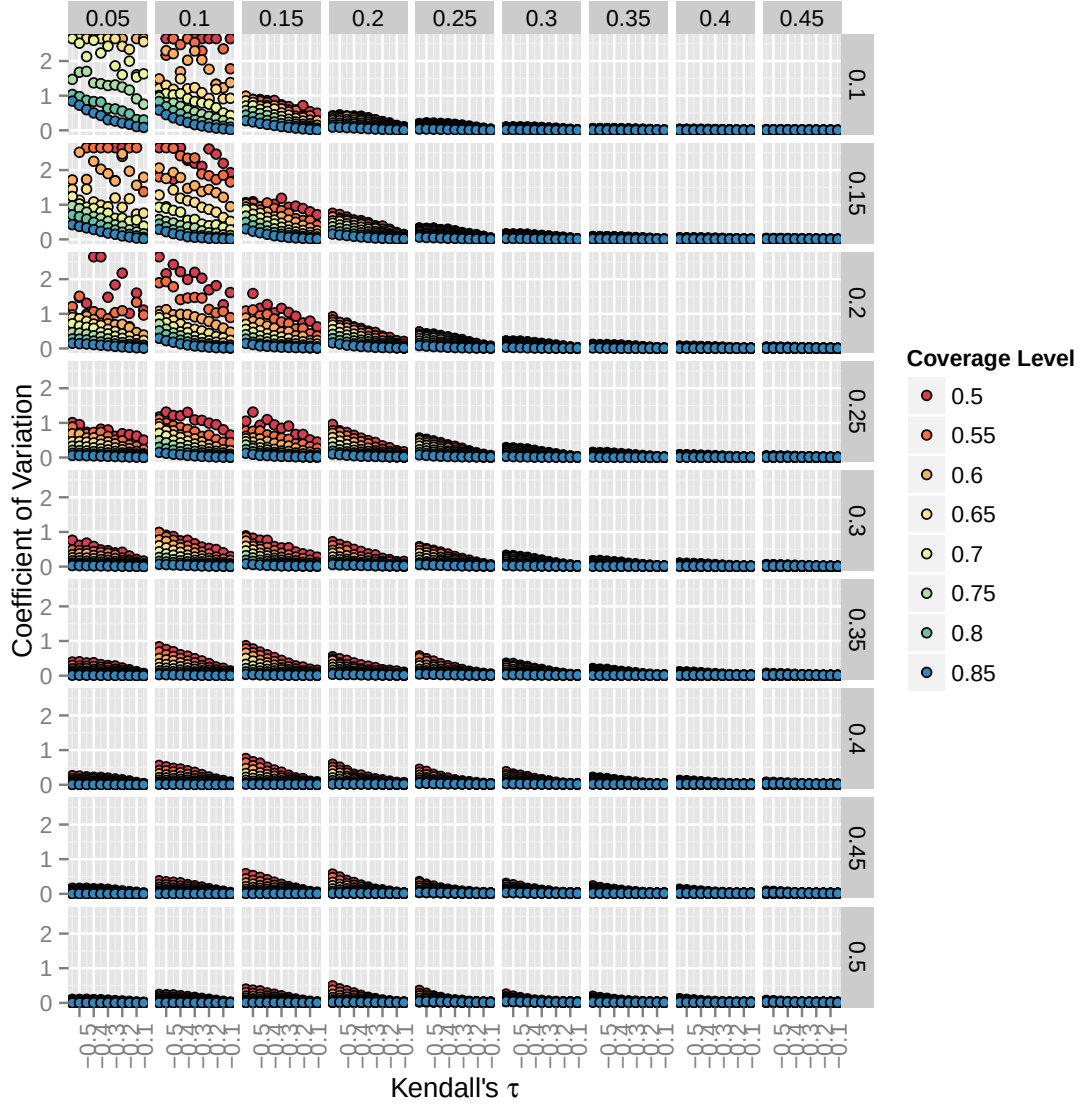


Figure 12: Between Copulas Simulated RPHPE Dispersion

Notes: This figure maps the coefficient of variation between Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates generated by alternative copulas across an extensive set of variation in price dispersion and yield variation onto Kendall's τ . The figure is divided into subplots by grouping the data by the coefficient of variation for farm level yields along the horizontal axis and price dispersion on the vertical axis, which creates 81 unique categorical combinations.

For neither revenue product are there coefficient of variation estimates for the lowest combination of price dispersion, yield variation and coverage level because the average premium rate is zero. Broadly, for RPHPE, premium rate

dispersion increases as coverage level decreases, price-yield dependence strengthens, yield variation decreases and price dispersion decreases. However, closer inspection reveals a consistent non-linear pattern. Starting at a yield variation of 0.15 and price dispersion of 0.25, the maximum level of variation actually occurs away from the minimum yield variation or price dispersion groups. A similar, but more complex pattern is evidenced in the RP premium rates. In particular, along yield variations of 0.15, the coefficient of variation in premium rates actually increases as a function of dependence for low price dispersion regions, is concave with respect to dependence for price dispersion levels of 0.20 and 0.25 and then decreases thereafter. Also, for both insurance types, a yield variation of 0.15 at low price dispersion exhibits smaller coefficient of variation values than the lower yield variation groups.

Deviation Between Copulas

The greater variation in the simulated data along with no-overlapping identification present a more comprehensive picture of the potential for rate deviation between the copula set as a function of parameter variation and coverage level. Figures 13 and 14 show RP and RPHPE premium deviation between the parametric copulas in the simulation.⁶⁰ A significant difference between these results and those of the previous empirical experiments is that mean premium rate is calculated over each unique categorical combination of rate determining factors. Thus, deviations are from the mean premium rate generated by the alternative

⁶⁰The motivation for this deviation metric is presented in Appendix F.

copulas at each combination of coverage level, price dispersion, yield variation and level of Kendall's τ .⁶¹

In general, premium rate deviations converge to the mean premium rate, as coverage level increases. However, the broader underlying base of parameter variation shows that high coverage level can also possess biased premium rates deviations between alternative copulas. It is hypothesized, from the dispersion experiment that the low yield variation, strong negative price-yield dependence and low price dispersion groups account for this lack of convergence because these groups have the most potential for premium rate dispersion. Future research could explore the functional form of this deviation. For both rate types, the general orientation of the deviation pairs between copulas match the empirical results: the RPHPE deviations are similarly symmetric while the RP deviations are all biased in alignment with the empirical results. Also, the simulated data shows that the Student-t copula calibrated with four degrees of freedom has a large range of potential premium rate deviation from the average premium rate as compared to all other copulas for both RP and RPHPE policies. This result reflects the fact that, of all the copulas examined in this study, the Student-t copula is the only one that does not approximate the Independence copula as dependence weakens.⁶²

For RP it is apparent that even at high coverage levels, for almost every combination except the Normal copula and its pairs, deviations among copulas from

⁶¹The variation metric for the empirical data did not normalize over the empirical variation in yields and dependence.

⁶²This result provides some support to the conclusion that low dependence counties are likely accounting for the lack of premium rate convergence, even at high coverage levels.

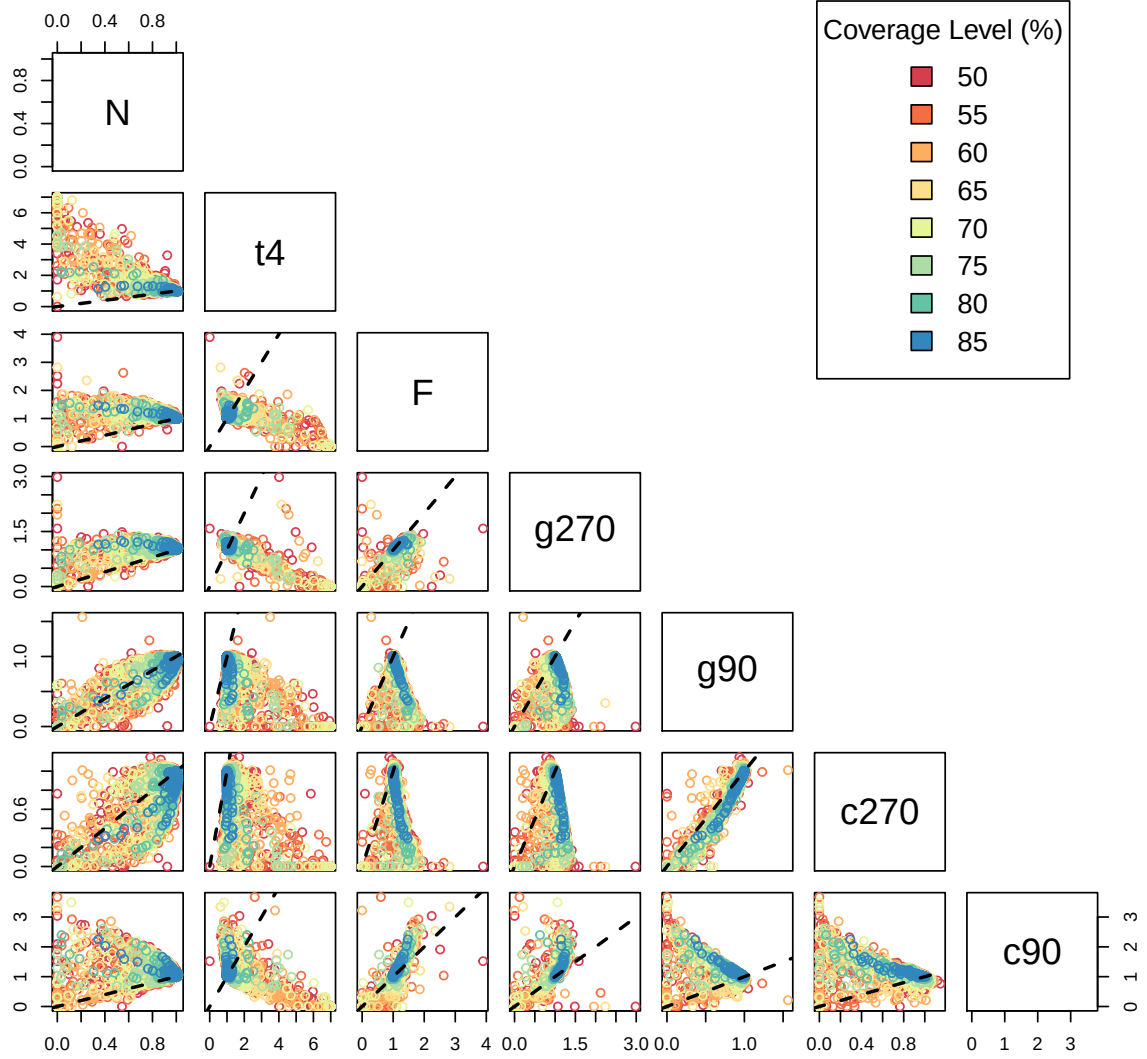


Figure 13: Between Copulas Simulated RP Deviations

Notes: This figure shows the simulated Revenue Protection (RP) premium rate deviations, which are calculated by dividing premium rates generated for each copula by the average premium rate generated by all copulas at each categorically unique combination of coverage level, price dispersion and yield variation and dependence, across an extensive set of factors determining premium rate levels. The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. Deviations greater than six are truncated.

the mean rate, for some underlying categorical combinations of the factors

determining premium rates, will not converge to the mean premium rate. A similar

conclusion can be made for RPHPE; however, there is an axis of symmetry, most

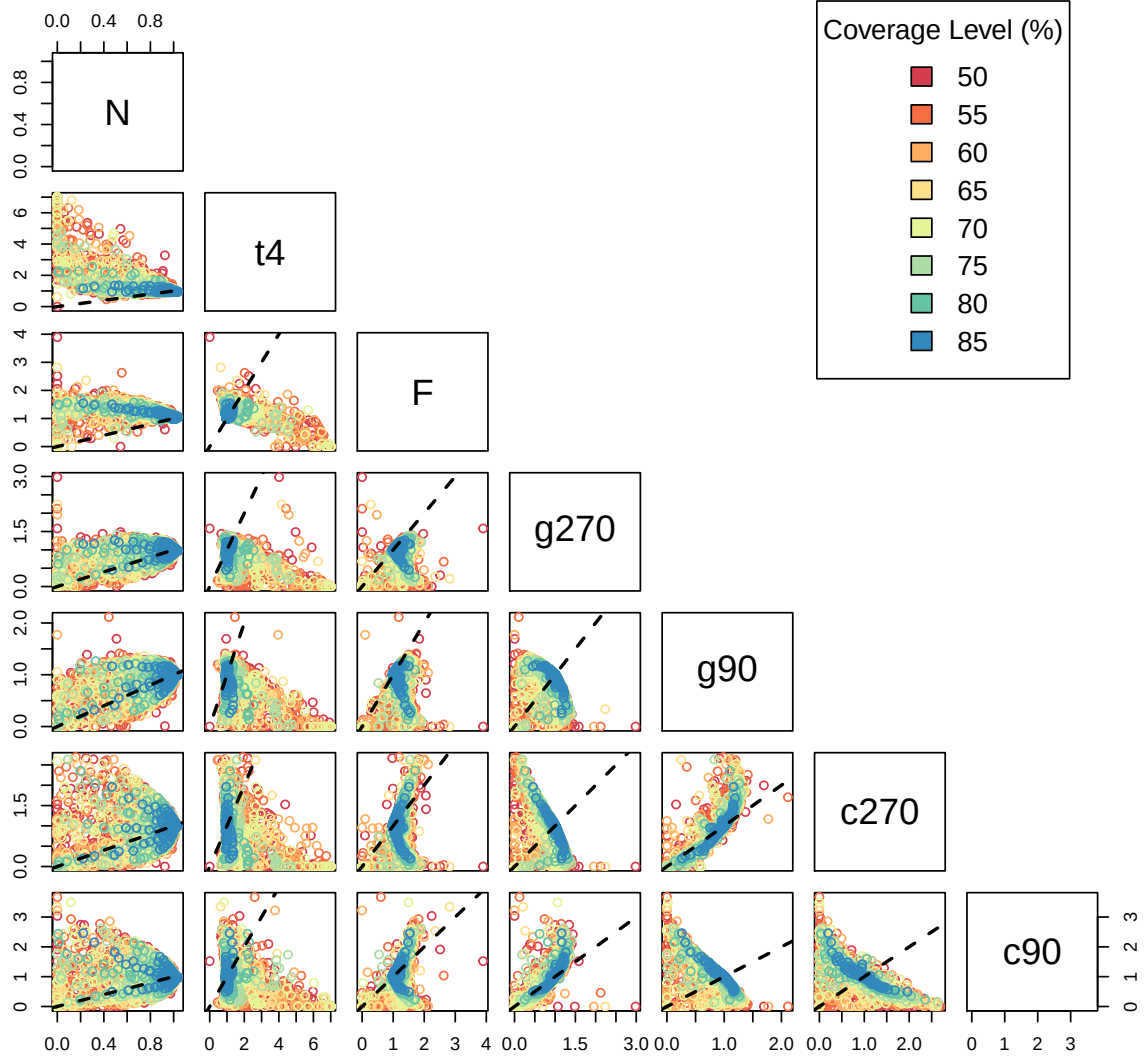


Figure 14: Between Copulas Simulated RPHPE Deviations

Notes: This figure shows the simulated Revenue Protection with Harvest Price Exclusion (RPHPE) premium rate deviations, which are calculated by dividing premium rates generated for each copula by the average premium rate generated by all copulas at each categorically unique combination of coverage level, price dispersion and yield variation and dependence, across an extensive set of factors determining premium rate levels. The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. Deviations greater than six are truncated.

noticeable for the highest coverage levels, among almost all copula deviation pairs,

which belies some underlying bifurcation as a function of the underlying

parameter identification and combinations. Perhaps the most important implication

of these results is that the Normal copula's premium rate deviations against the deviations of all other copulas most consistently converge to the mean premium rate—at least at high coverage levels.

Summary

The simulation results indicate that dispersion amongst RP and RPHPE premium rates generated by alternative copulas increase as coverage level decreases, yield variation decreases, price dispersion decreases and negative dependence becomes stronger. The results mirror the findings of Goodwin (2014), who found that premium rates generated by different copulas diverged more at lower coverage levels. Premium rates for low coverage levels are based upon fewer price and yield realizations, which are located in the tails of the multivariate distribution imposed by the copula. These results indicate that the effect of a copula ratemaking strategy has a differential impact on the how the relationship between extreme price and yield events are modeled, which is most apparent at the low coverage levels that protect against such outcomes. While the affect of the price and yield marginal distributions are non-linear, dispersion between premium rates tends to be greatest when these marginal distributions have low levels of dispersion. When the marginal distributions exhibit low levels of dispersion even fewer observation are determining premium rates and the effect of the copula has an even larger impact on premium rate dispersion. These results indicate that the identification of the marginal distributions matter when generating revenue premium rates with a copula based

method. Presumably, both premium rate dispersion and deviation would increase if alternative marginal distributions were considered.

If revenue product premium rates continue to be generated with a copula based approach, without a method for selecting between alternative copulas, the Normal copula may prove the most viable alternative, as it appears to produce the least potential for premium rate bias when compared to the alternative copulas of this study. However, any copula—including the Normal—introduces implicit premium rate imprecision because alternative copulas cannot be validly excluded. This imprecision is not currently incorporated into revenue product premium rates. Therefore, all revenue product premium rates generated by a copula based ratemaking procedure are currently downward biased and the extent of this bias depends on the empirical characteristics of the county and crop. Incorporating the additional imprecision induced by copula modeling into premium rates would cause across the board premium rate increases. As the MPCCI program is heavily subsidized, such action could greatly increase the burden on tax payers. In light of this fact, using the pure empirical ratemaking method and perhaps incorporating price dispersion, might prove not only a more defensible ratemaking method, but would likely generate lower premium rates that better meet the policy objectives of MPCCI.

CHAPTER NINE

SIMULATED PREMIUM RATE DISPERSION WITHIN COPULAS

This study investigated the potential for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion (RPHPE) premium rate dispersion within copulas by generating distributions for the premium rates produced by alternative copulas. The affects of the factors that determine premium rates: coverage level, price-yield dependence as measured by Kendall's τ , and the dispersion in the price and yield marginal distributions are examined. "True" premium rate distributions are generated by calibrating the copula by known measures of dependence. Sample premium rate distributions are generated by estimating the level of dependence based on a sub-sample of price and yield realizations bound by a known copula and know level of dependence. The distributions of sampled premium rates model the imprecision implicit to the current ratemaking process. By examining the distributional properties of the "true" premium rates along side the sample premium rates, generated from an estimation process, the impact of sampling error on copula based ratemaking is isolated. The results demonstrate the conditions that induce copula based premium rate imprecision, regardless of the copula selected and apply to the current ratemaking practice, which uses a Normal copula.

Design and Data Generation

As in the previous experiments, premium rates are generated from 100,000 correlate price-yield realizations. One thousand Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates⁶³ are estimated for all parametric copulas examined in this study.⁶⁴ To reduce computation time, the underlying sources of premium rate variation are limited. Premium rates are only calculated for the 50%, 65% and 85% coverage levels. Dependence is measured by Kendall's τ and is limited to -0.55, -0.35 and -0.05. The dispersion of the yield marginal distribution is identified by standard deviations of 5 bu./ac., 25 bu./ac. or 45 bu./ac., with the mean yield fixed at 100 bu./ac. Therefore, the coefficient of yield variation is either 5%, 25% or 45% of the mean yield. Finally, the dispersion in prices is either 0.20, 0.35 or 0.50 and the mean of the process is expected price of \$5.78.

In this experiment 50 price-yield realizations are sampled from the 100,000 price-yield pairs joined by a copula calibrated at a specified level of dependence. These data serve as simulated "empirical" data, from which a hypothetical ratemaking practitioner would need to not only estimate dependence, but also identify the appropriate copula to generate premium rates. Using the simulated empirical data a dependence estimate is calculated. The estimated measure of

⁶³Greater detail of the ratemaking process is provided in the Ratemaking Procedures section of the first empirical experiment

⁶⁴The parametric copulas in this experiment are the same as the previous experiment and consist of the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270) copulas.

dependence is then used to calibrate the shape parameter of the copula and premium rates are then established from 100,000 price-yield realizations bound by a copula that is calibrated by the estimated level of dependence.⁶⁵ These data have the unique property of being generated and bound by a known processes. Therefore, the distribution of the sample premium rates generated by a hypothetical practitioner can then be compared to the distribution of the “true” premium rates.

For either “true” or sampled premium rates, 100,000 price and yield realization, are generated from the assumed marginal yield distribution calibrated by the true underlying parameters, upon which premium rates are based. Thus, additional error is not introduced by estimating the properties of the yield distribution from the simulated empirical data. This allows the effect of the copulas and their parameter estimate to drive any subsequent variation in premium rates.⁶⁶

Results

The dispersion of both “true” and sample generated premium rates within copulas is first examined using the coefficient of variation metric. Next, the potential sources causing within copula premium rate dispersion are isolated and explored.

⁶⁵Sampling error in the estimation of the price and yield dependence, creates the possibility of positively dependent price-yield data. While this poses no problem for the Normal, Student-t and Frank copulas, the rotated Clayton and Gumbel copulas cannot be calibrated using a positive dependence measure. In both cases the shape of the rotated 90° copula as it passes from a negative representation to positive is reflected in the rotated 180° degree or “survival” copula and the rotated 270° copula by the untransformed (i.e. rotated 0°) copula. In instances where the data are positively dependent, Clayton and Gumbel premium rates were generated by these equivalently shaped positive representations.

⁶⁶Variation was further reduced by not utilizing the parameter estimate generated by the goodness-of-fit tests, but by one based directly from the initial estimate of dependence. Thus, the parameter estimate across rotations of the Gumbel and Clayton copulas are identical in absolute value.

True and Sample Premium Rates

In order to preliminary examine the impact of the premium rate determining sources of variation, linear regression analysis was performed on the coefficient of variation for premium rates generated by the “true” calibration of the copulas. Table 10 shows that the same functional relationship that drives dispersion between premium rates generated by alternative copulas, also drives the dispersion of premium rates generated within copulas. On average, as coverage level increases premium rate dispersion decreases, as price dispersion or yield variation increases premium rate dispersion decreases and as the negative price-yield dependence weakens rate dispersion decreases.

This result reinforces that effect of the particular relationship copulas impose on the tails of their multivariate distributions. The only difference between the one thousand “true” premium rates, is the resampling of the 100,000 price and yield realizations. Therefore, even slight differences in the population of these marginal distributions is magnified at coverage levels that only protect against outcomes that occur deep in the tails of the multivariate distribution imposed by the copula (Hungerford and Goodwin, 2014). Premium rate dispersion is amplified when the copula is more uniquely identified and operates over tightly dispersed marginal distributions, with even fewer price or yield realizations populating the tails.

Table 11 shows that, at least on average, the dispersion of sample premium rates follows a similar pattern as that of the true rates. However, the limits of linear

Table 10: Within Copula True Premium Rate Dispersion

	Estimate	Std. Error	t value	Pr(> t)
Intercept	-0.11	0.16	-0.67	0.50
RPHPE	2.60×10^{-2}	0.08	0.03	0.97
Student-t (df=4) Copula	-0.12	0.15	-0.84	0.40
Frank Copula	-0.11	0.15	-0.74	0.46
Gumbel 90° Copula	0.15	0.15	1.03	0.30
Gumbel 270° Copula	-0.09	0.15	-0.60	0.55
Clayton 90° Copula	-0.11	0.15	-0.78	0.44
Clayton 270° Copula	0.25	0.15	1.68	0.09
50% Coverage Level	0.45	0.10	4.65	3.68×10^{-6}
85% Coverage Level	-0.02	0.10	-0.17	0.87
0.20 Price Dispersion	0.46	0.10	4.74	2.39×10^{-6}
0.50 Price Dispersion	-0.01	0.10	-0.07	0.95
Yield CV 0.05	0.46	0.10	4.79	1.85×10^{-6}
Yield CV 0.45	-0.01	0.10	-0.08	0.94
Kendall's τ of -0.30	-0.21	0.10	-2.22	0.03
Kendall's τ of -0.05	-0.30	0.10	-3.08	2.10×10^{-3}

Notes: This table provides the results of a linear regression model of the coefficient of variation for one thousand "true" premium rates generated by each copula at each unique combination of coverage level, price dispersion, yield coefficient of variation and level of dependence. The premium rates are generated from a known copula and level of dependence. The reference condition is for Revenue Protection (RP) premium rates generated by a Normal copula, rated at the 65% coverage level and with a price variation of 0.35, yield variation of 0.25 and dependence as measured by Kendall's τ of -0.30. The Student-t copula is calibrated with 4 degrees of freedom. The Revenue Protection with Harvest Price Exclusion policy is abbreviated RPHPE.

regression analysis mask the non-linearities of within copula generated premium rate dispersion, as well as the different relationship that the dispersion of sample premium rates and "true" premium rates have to the sources determining premium rate variation.

Figures 15 through 25, show the coefficient of variation for RP and RPHPE premium rates generated within the copulas at 50%, 65% and 85% coverage levels. The figures are divided in subplots by yield variation along the horizontal axis and price dispersion along the vertical axis. The coefficient of variation is calculated at

Table 11: Within Copula Sample Premium Rate Dispersion

	Estimate	Std. Error	t value	Pr(> t)
Intercept	0.08	0.05	1.65	0.10
RPHPE	0.06	0.02	2.51	0.01
Student-t (df=4) Copula	0.01	0.04	0.27	0.78
Frank Copula	-0.01	0.04	-0.12	0.91
Gumbel 90° Copula	3.90×10^{-3}	0.04	0.09	0.93
Gumbel 270° Copula	3.80×10^{-3}	0.04	0.09	0.93
Clayton 90° Copula	3.50×10^{-3}	0.04	0.08	0.93
Clayton 270° Copula	3.20×10^{-3}	0.04	0.07	0.94
50% Coverage Level	0.25	0.03	8.86	$< 2 \times 10^{-16}$
85% Coverage Level	-0.11	0.03	-3.74	1.91×10^{-4}
0.20 Price Dispersion	0.28	0.03	9.70	$< 2 \times 10^{-16}$
0.50 Price Dispersion	-0.04	0.03	-1.43	0.15
Yield CV 0.05	0.24	0.03	8.51	$< 2 \times 10^{-16}$
Yield CV 0.45	-0.13	0.03	-4.57	5.42×10^{-6}
Kendall's τ of -0.30	-0.07	0.03	-2.29	0.02
Kendall's τ of -0.05	-0.10	0.03	-3.42	6.41×10^{-4}

Notes: This table provides the results of a linear regression model of the coefficient of variation for one thousand sample premium rates generated by each copula at each unique combination of coverage level, price dispersion and level of dependence. The dependence used to calibrate the premium rate generating copula is estimated from a sample of 50 price and yield pairs that come from a known copula and level of dependence. The reference condition is for Revenue Protection (RP) premium rates generated by a Normal copula, rated at the 65% coverage level and with a price variation of 0.35, yield variation of 0.25 and dependence as measured by Kendall's τ of -0.30. The Student-t copula is calibrated with 4 degrees of freedom. The Revenue Protection with Harvest Price Exclusion policy is abbreviated RPHPE.

each unique combination of coverage level, price dispersion and yield variation, for each copula and insurance policy. This simple metric gives some indication of how premium rate dispersion and therefore imprecision evolves, when standardized by a measure of central tendency, across the sources determining premium rates. The coefficient of variation for the sample premium rates tend to be greater than the coefficient of variation for the “true” premium rates.⁶⁷

⁶⁷ A notable exception occurs at the lowest coverage level, price dispersion and yield variation combination wherein sample rates actually tend to be almost an order of magnitude less. This results, at least accounts, for the regression results where the intercept for the sample data are actually less than the true data.

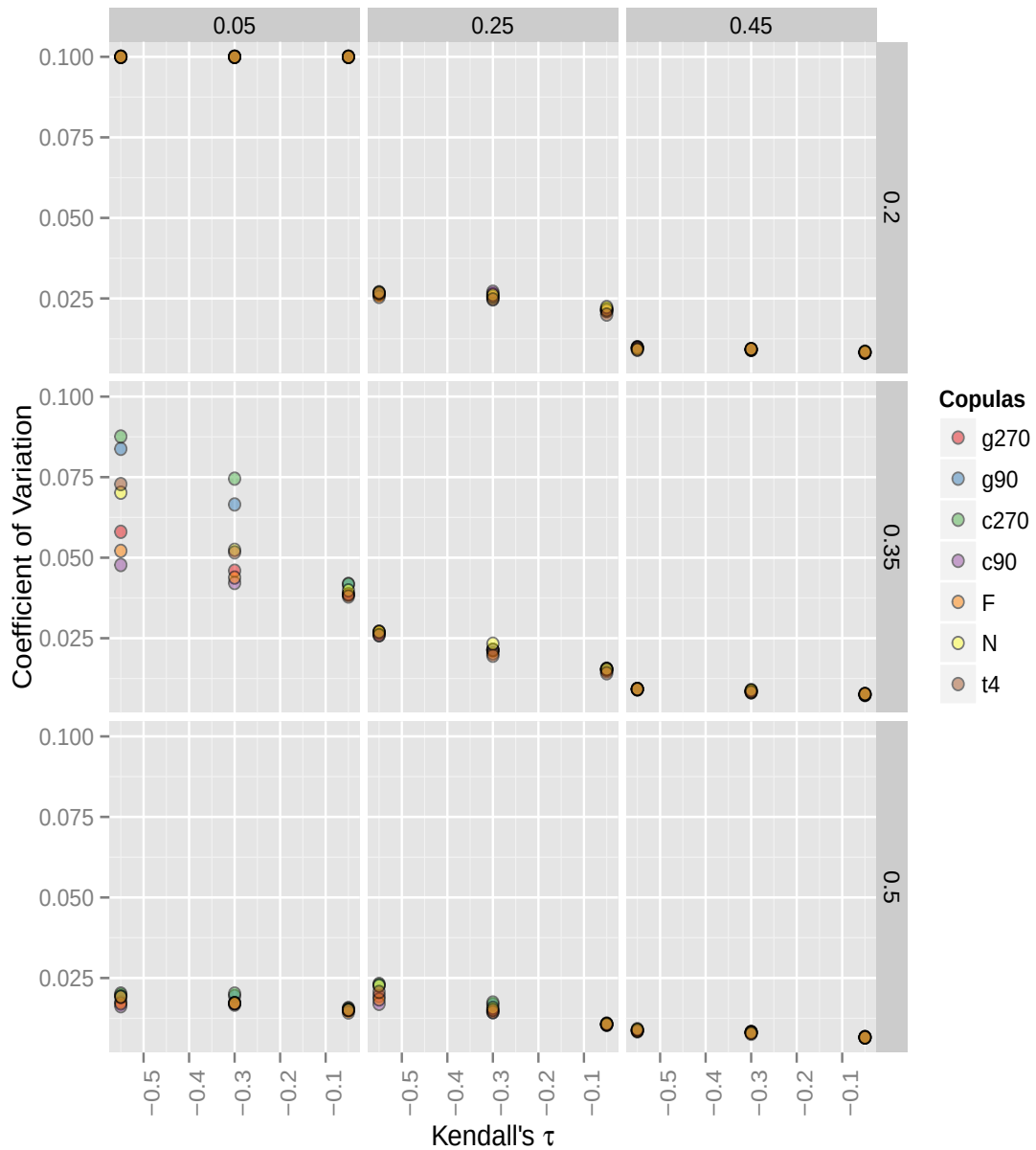


Figure 15: True RP Dispersion Within Copulas 50% Coverage

Notes: This figure maps the coefficient of variation of one thousand simulated “true” Revenue Protection (RP) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall’s τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270), and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are precisely identified by the all underlying parameters. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .1 are truncated.

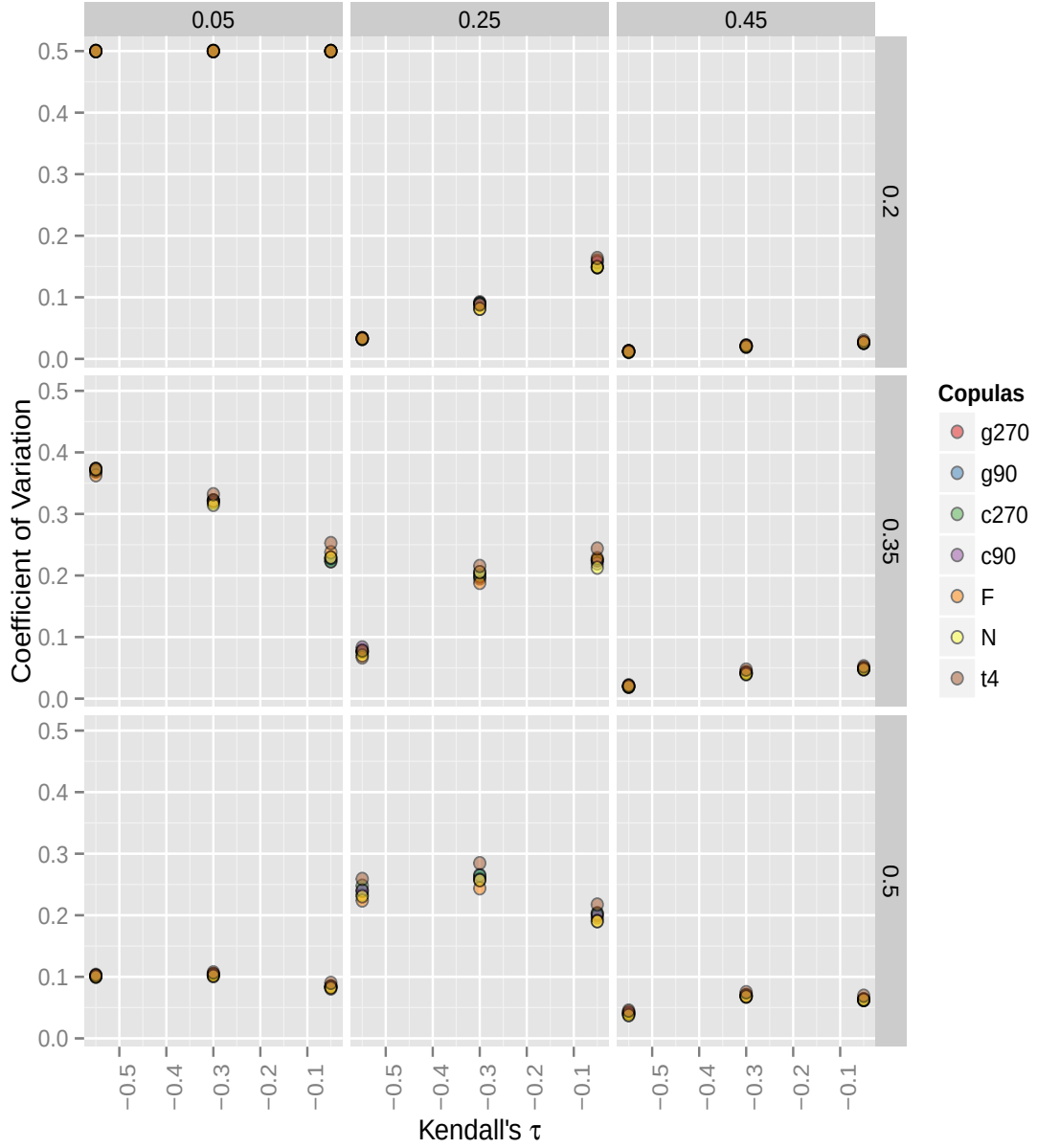


Figure 16: Sample RP Dispersion Within Copulas 50% Coverage

Notes: This figure maps the coefficient of variation of one thousand sample Revenue Protection (RP) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall's τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are generated from copulas calibrated by a price-yield dependence that is estimated from a sample of 50 price-yield pairs that come from a known copula and dependence. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .5 are truncated.

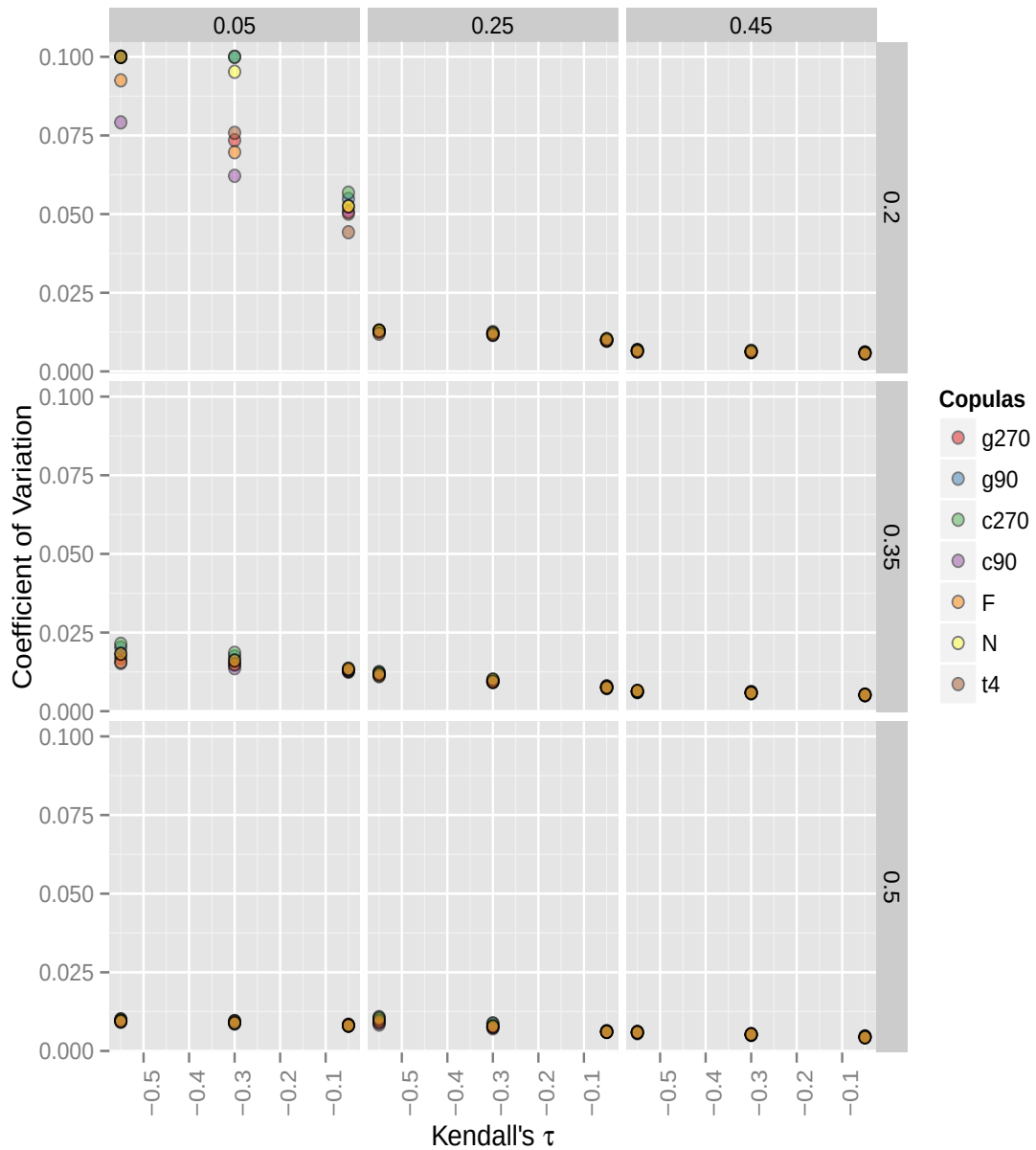


Figure 17: True RP Dispersion Within Copulas 65% Coverage

Notes: This figure maps the coefficient of variation of one thousand simulated “true” Revenue Protection (RP) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall’s τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are precisely identified by the all underlying parameters. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .1 are truncated.

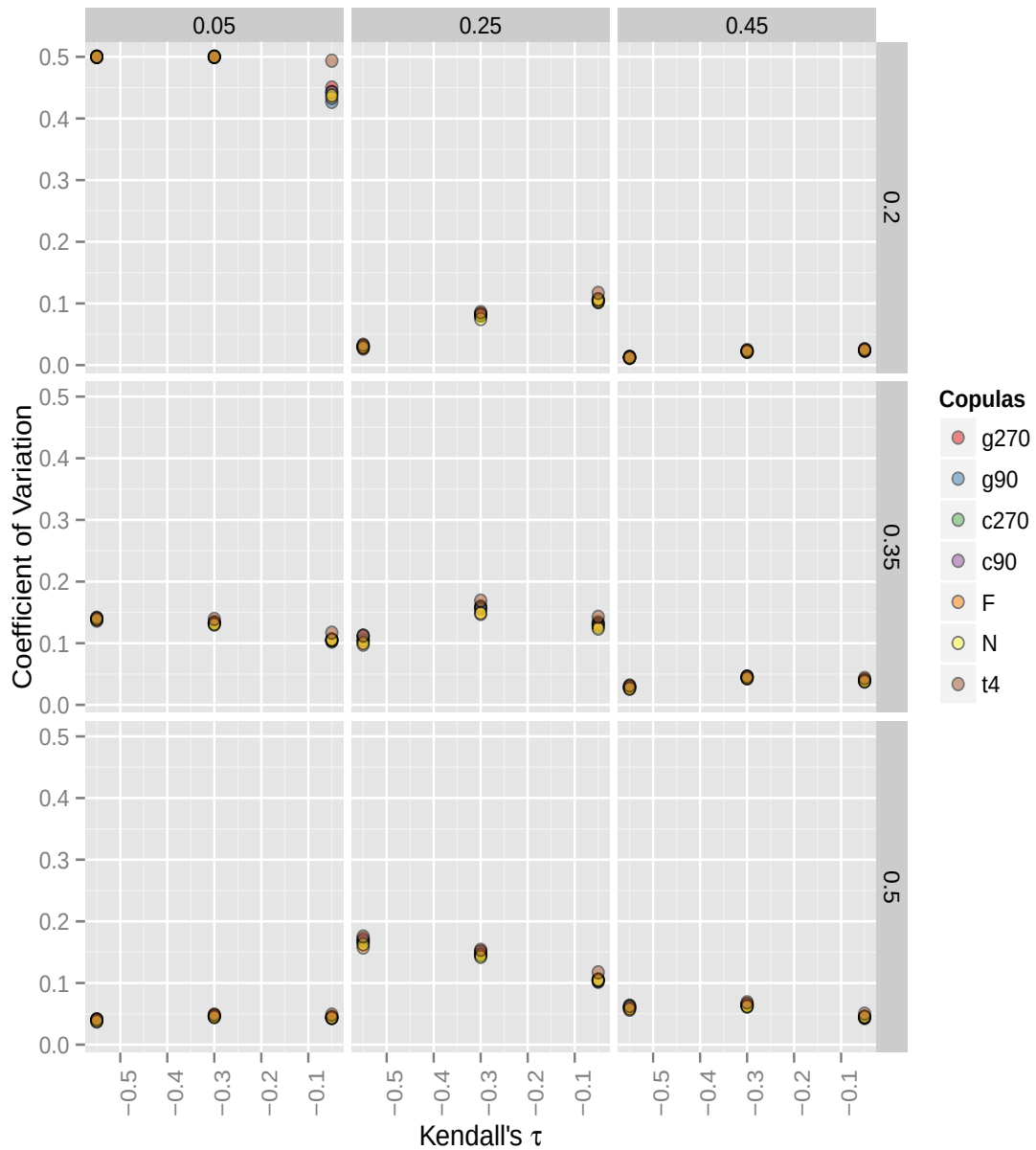


Figure 18: Sample RP Dispersion Within Copulas 65% Coverage

Notes: This figure maps the coefficient of variation of one thousand sample Revenue Protection (RP) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall's τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are generated from copulas calibrated by a price-yield dependence that is estimated from a sample of 50 price-yield pairs that come from a known copula and dependence. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .5 are truncated.

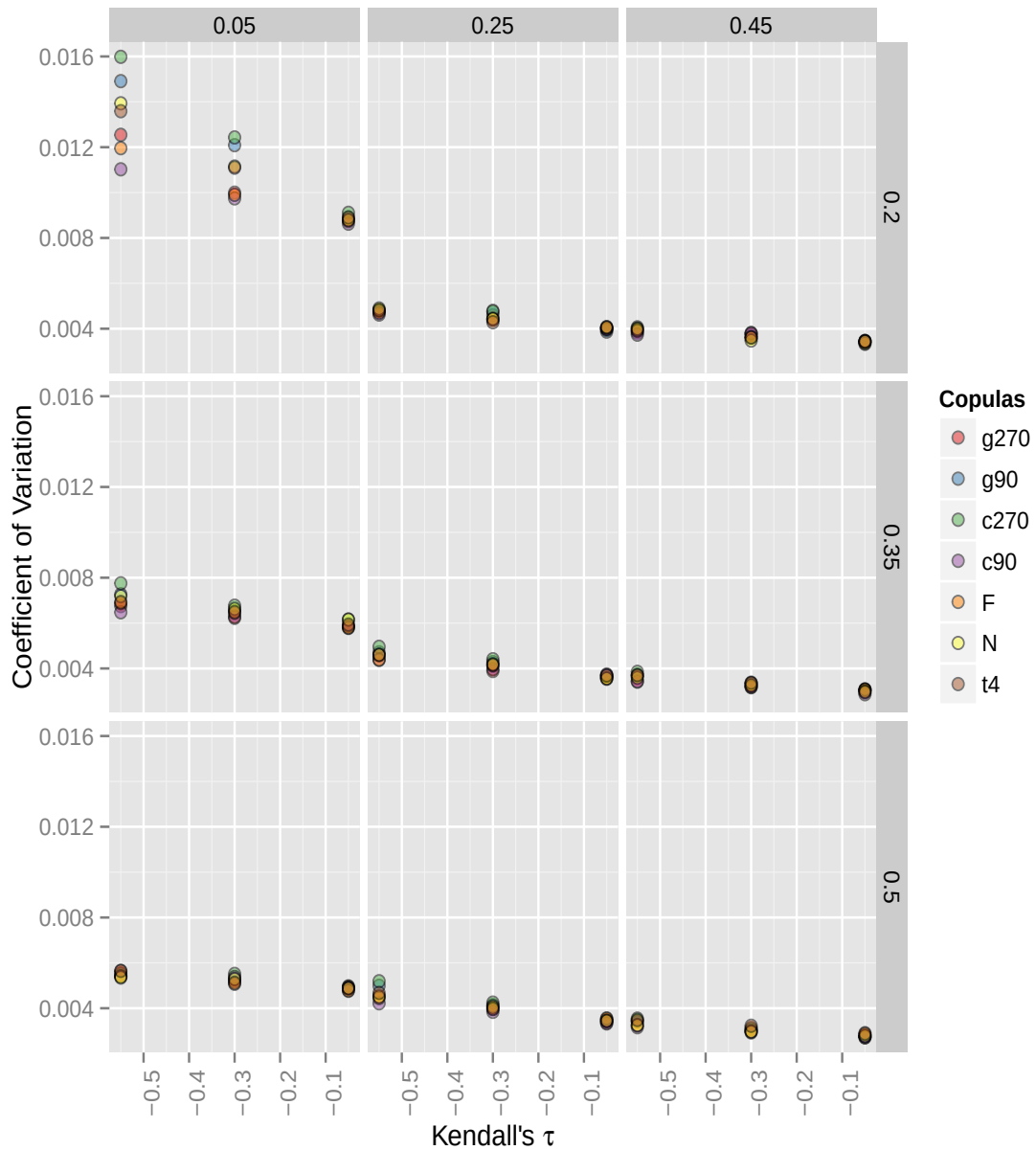


Figure 19: True RP Dispersion Within Copulas 85% Coverage

Notes: This figure maps the coefficient of variation of one thousand simulated “true” Revenue Protection (RP) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall’s τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are precisely identified by the all underlying parameters. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .1 are truncated.

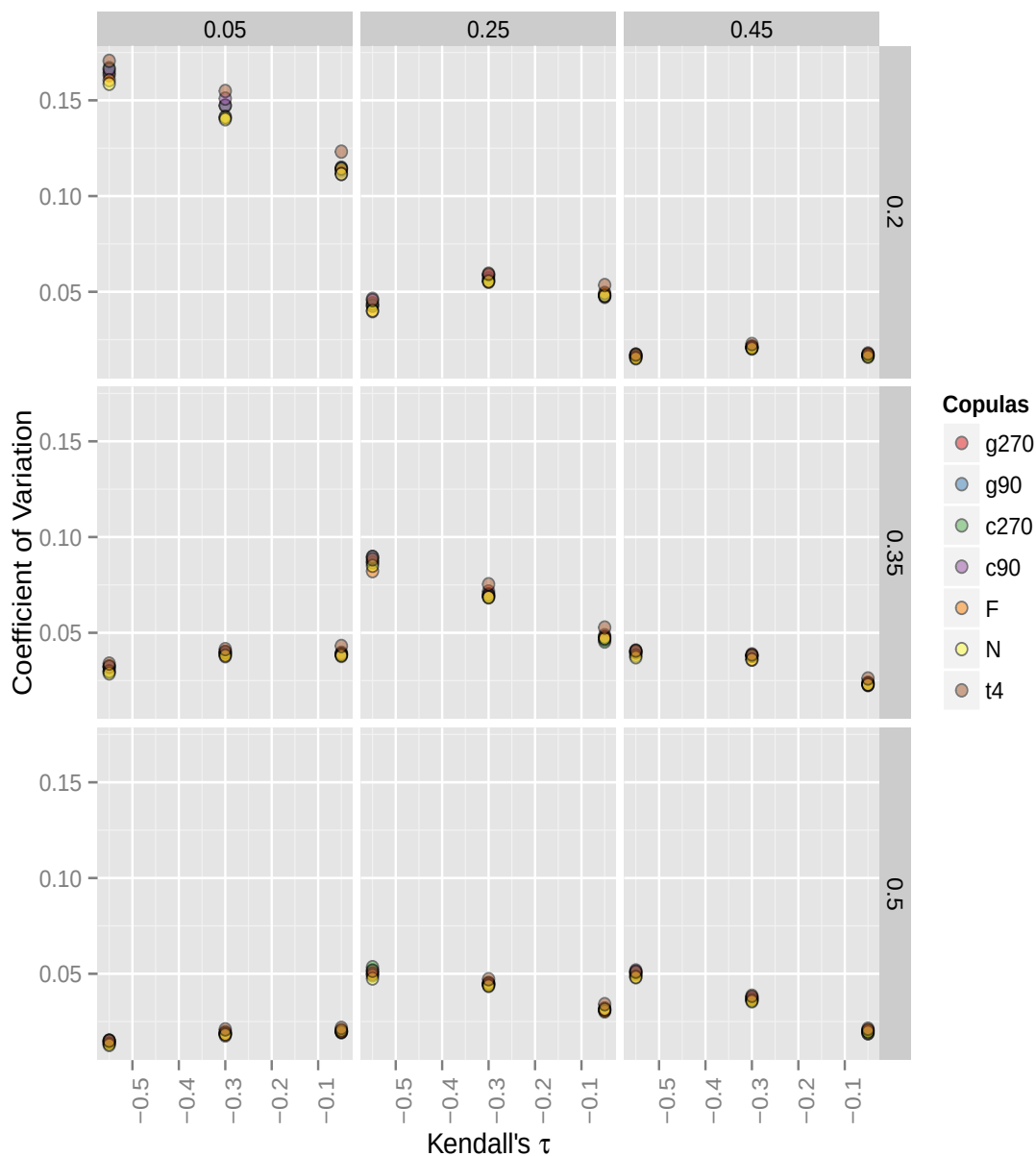


Figure 20: Sample RP Dispersion Within Copulas 85% Coverage

Notes: This figure maps the coefficient of variation of one thousand sample Revenue Protection (RP) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall's τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are generated from copulas calibrated by a price-yield dependence that is estimated from a sample of 50 price-yield pairs that come from a known copula and dependence. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .5 are truncated.

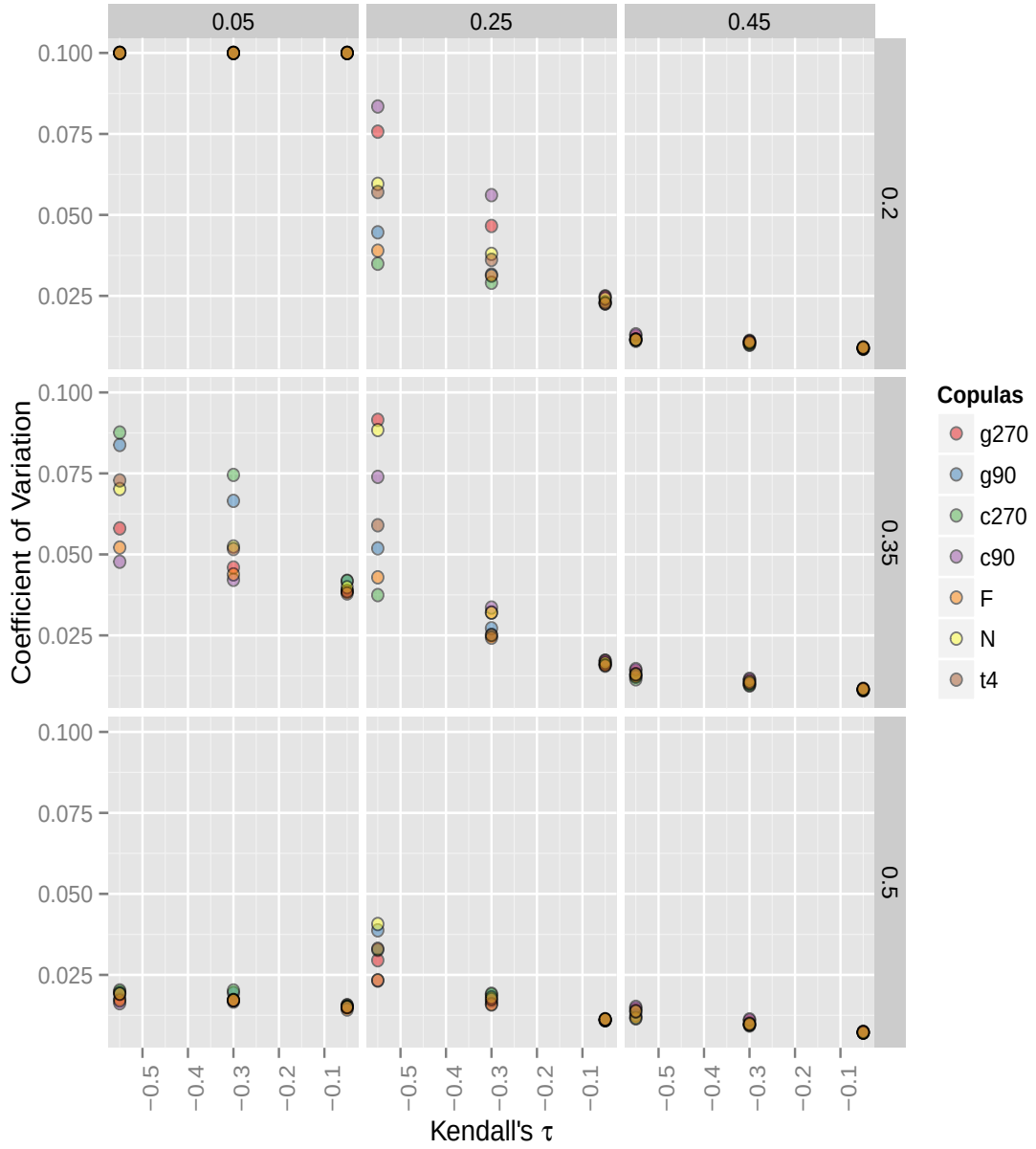


Figure 21: True RPHPE Dispersion Within Copulas 50% Coverage

Notes: This figure maps the coefficient of variation of one thousand simulated “true” Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall’s τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are precisely identified by the all underlying parameters. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .1 are truncated.

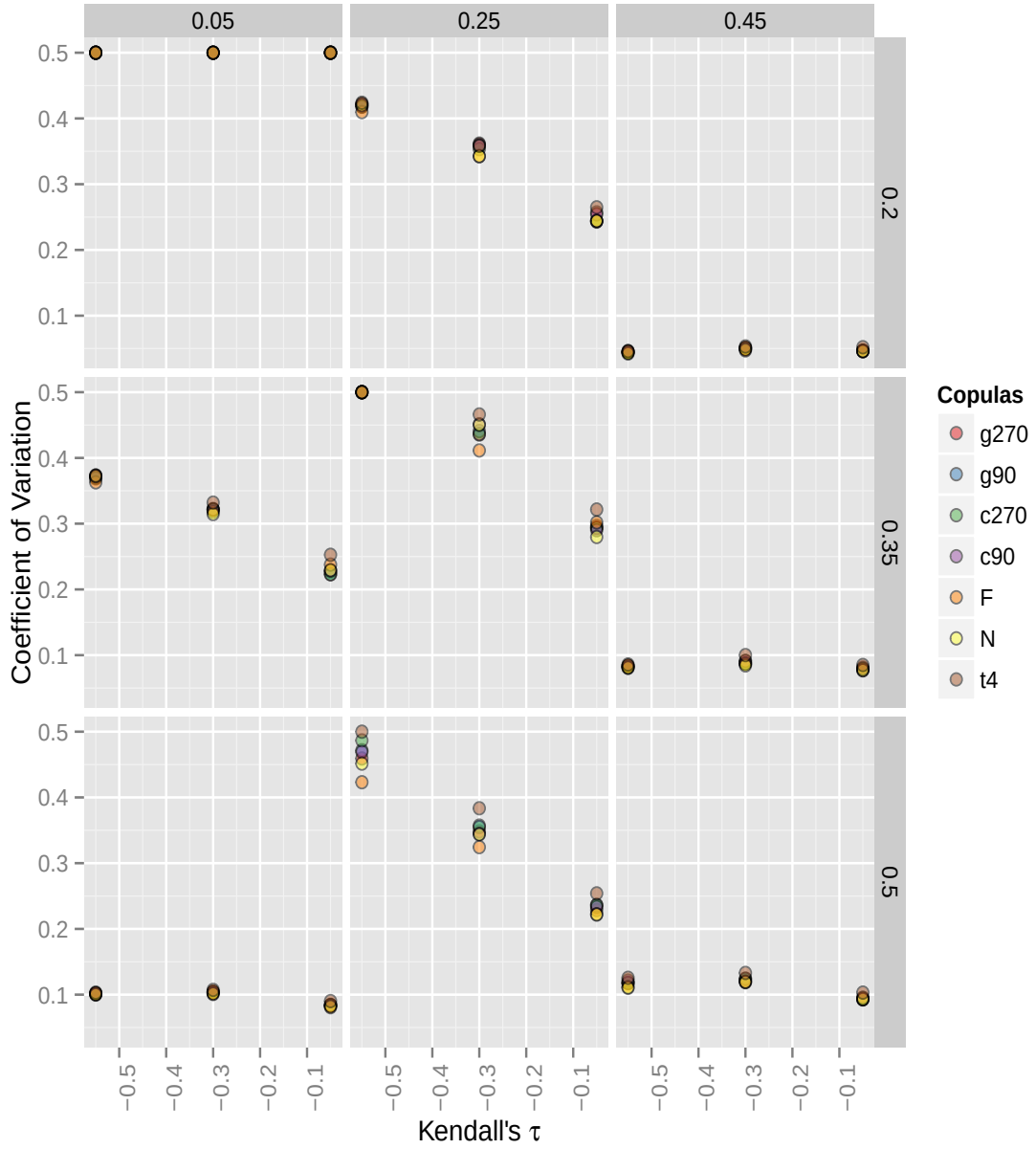


Figure 22: Sample RPHPE Dispersion Within Copulas 50% Coverage

Notes: This figure maps the coefficient of variation of one thousand sample Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall's τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are generated from copulas calibrated by a price-yield dependence that is estimated from a sample of 50 price-yield pairs that come from a known copula and dependence. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .5 are truncated.

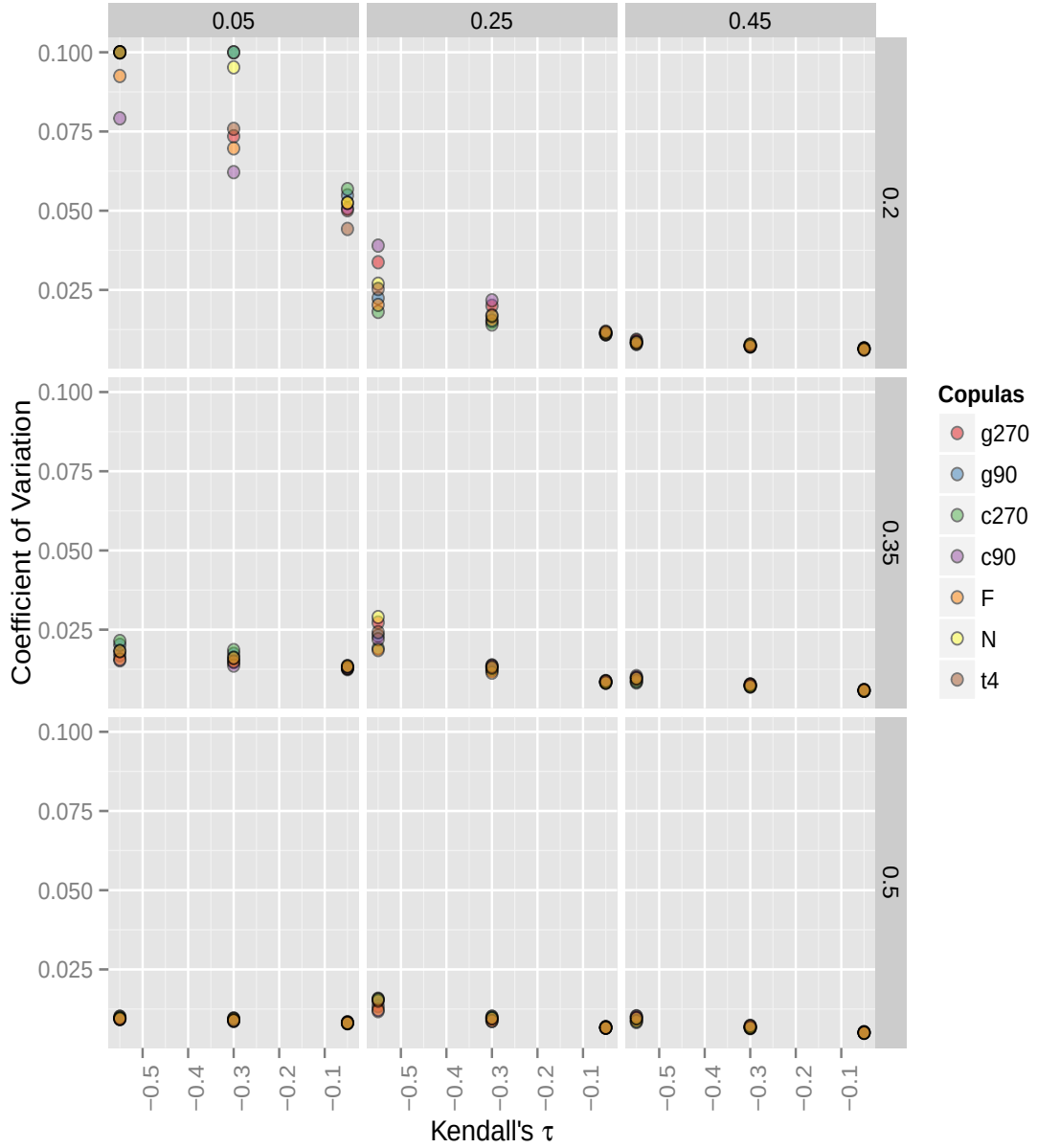


Figure 23: True RPHPE Dispersion Within Copulas 65% Coverage

Notes: This figure maps the coefficient of variation of one thousand simulated “true” Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall’s τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are precisely identified by the all underlying parameters. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .1 are truncated.

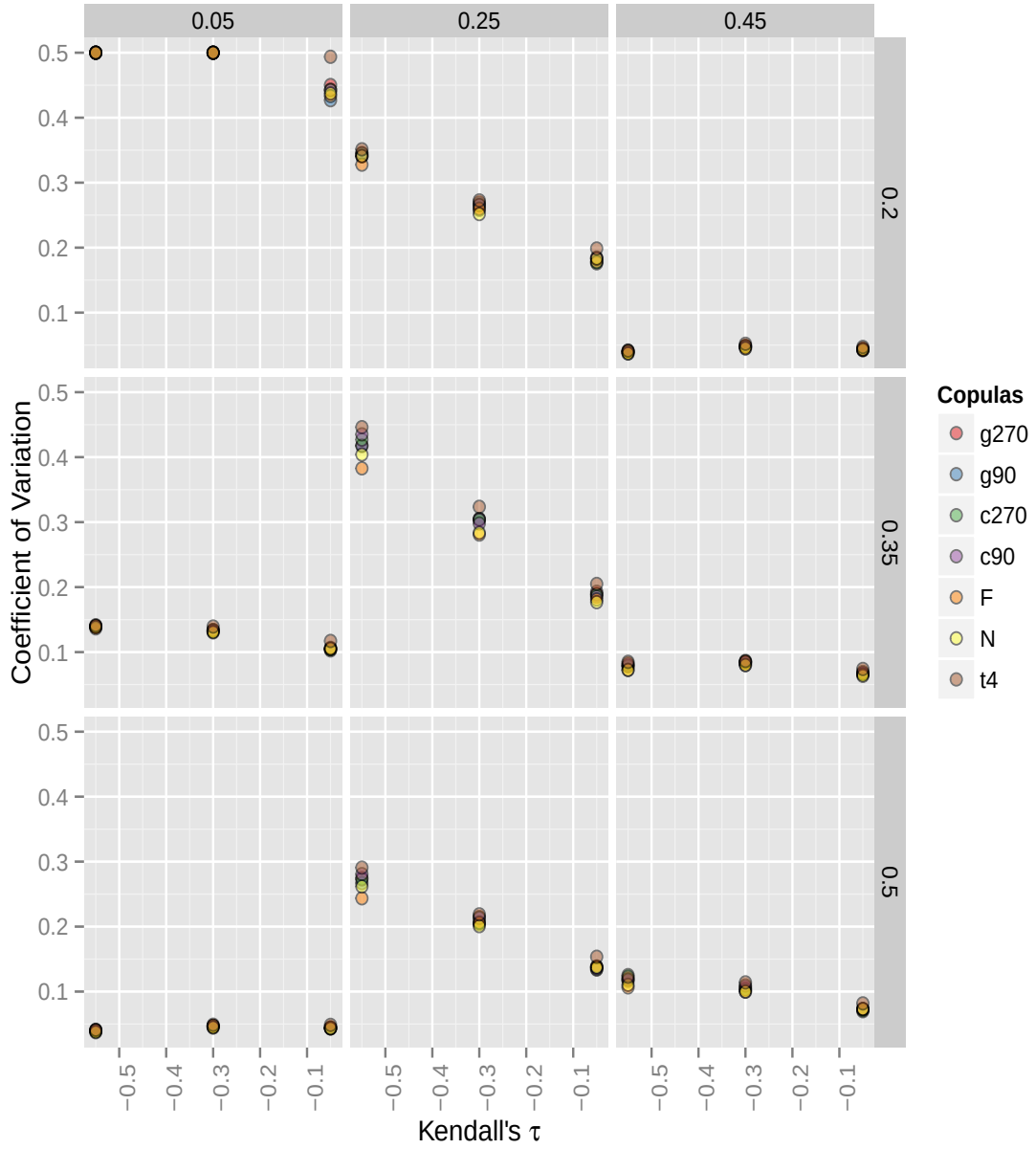


Figure 24: Sample RPHPE Dispersion Within Copulas 65% Coverage

Notes: This figure maps the coefficient of variation of one thousand sample Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall's τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are generated from copulas calibrated by a price-yield dependence that is estimated from a sample of 50 price-yield pairs that come from a known copula and dependence. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .5 are truncated.

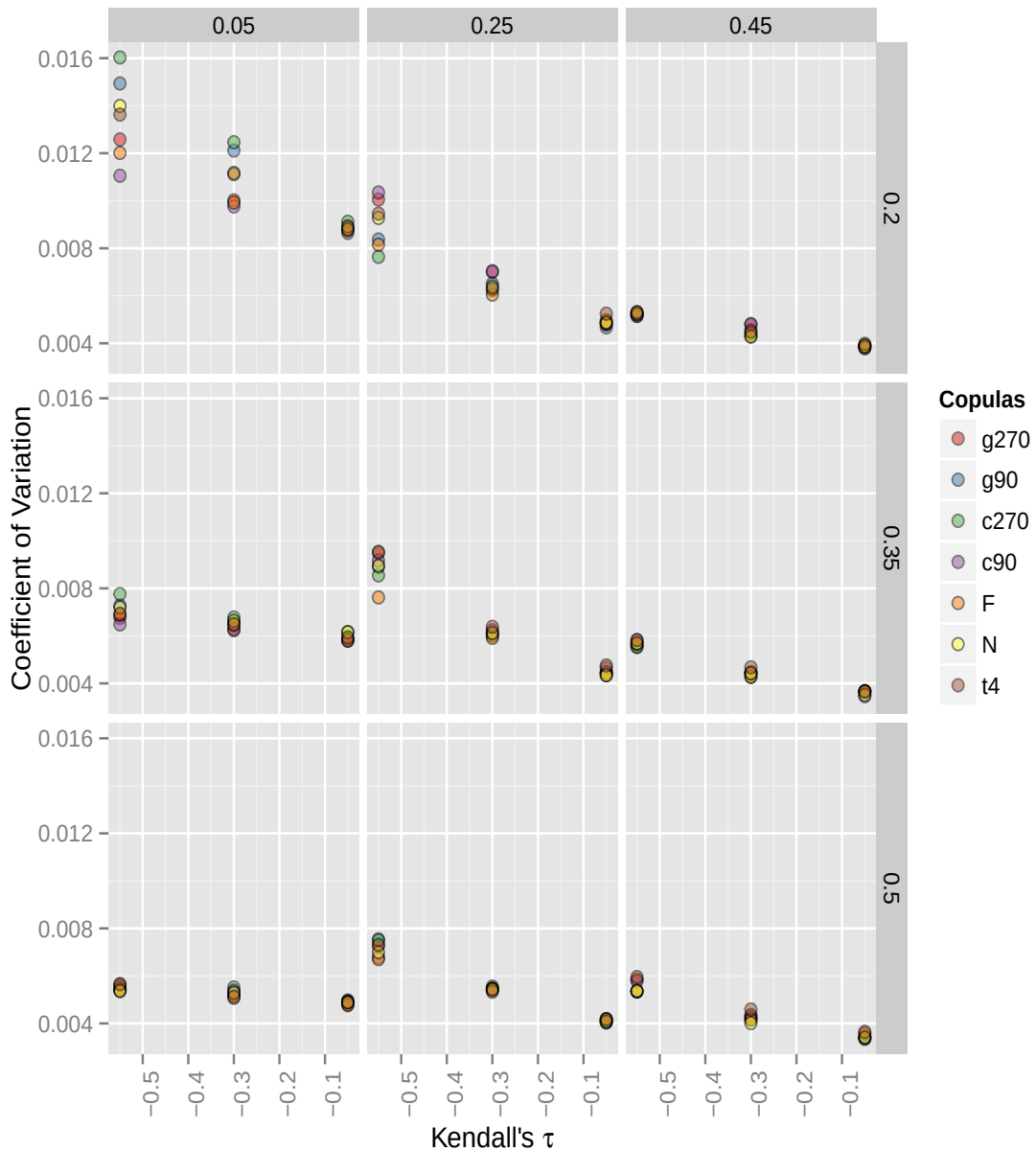


Figure 25: True RPHPE Dispersion Within Copulas 85% Coverage

Notes: This figure maps the coefficient of variation of one thousand simulated “true” premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall’s τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are precisely identified by the all underlying parameters. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .1 are truncated.

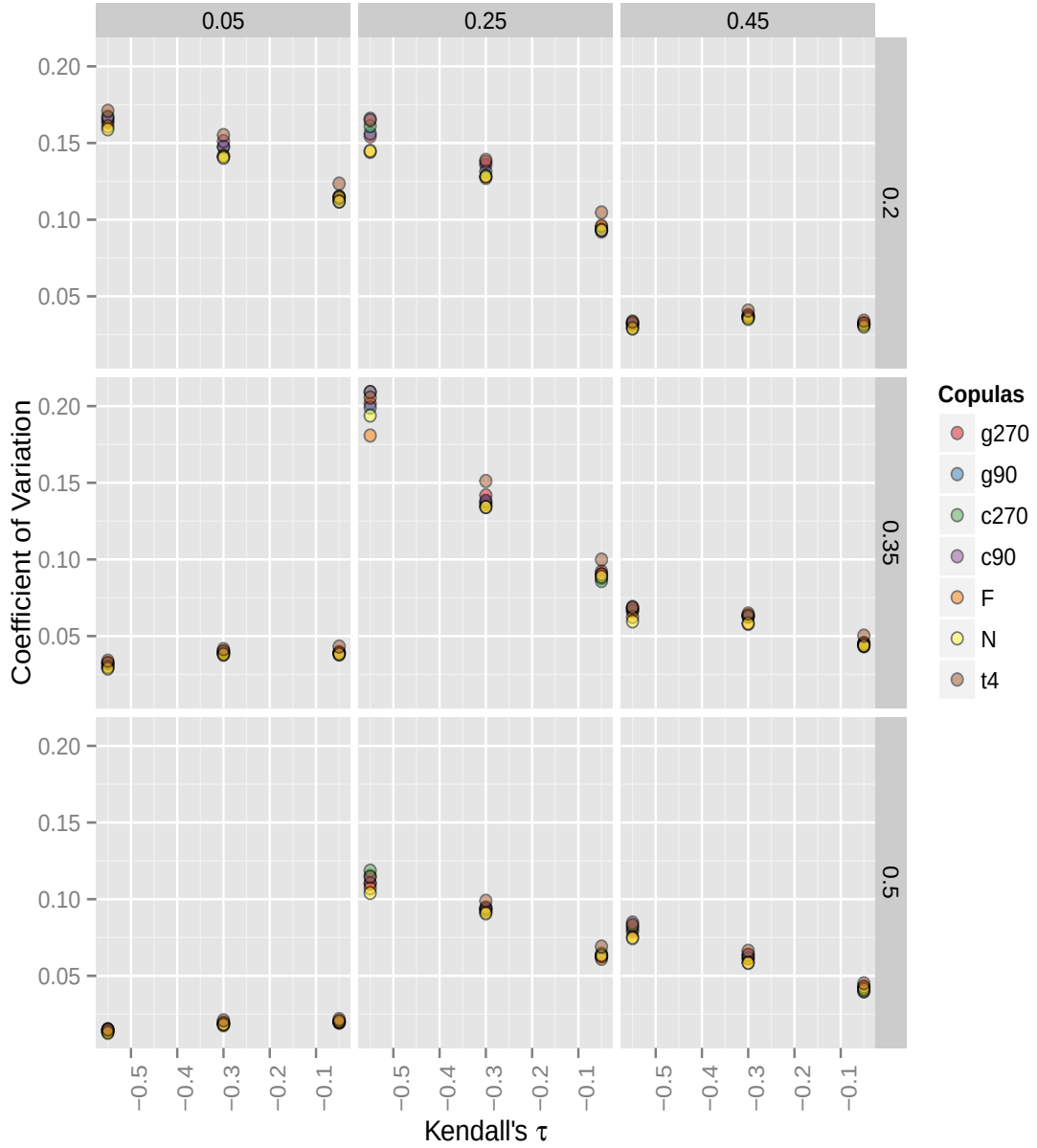


Figure 26: Sample RPHPE Dispersion Within Copulas 85% Coverage

Notes: This figure maps the coefficient of variation of one thousand simulated “true” Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates generated within alternative copulas across alternative specifications for price dispersion and yield variation onto Kendall’s τ . The parametric copulas used to generate premium rates are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270). The dashed 45° line represents no deviation from the average copula generated premium rate. These premium rates are precisely identified by the all underlying parameters. The figure is divided into subplots by grouping the data by the coefficient of variation for yields along the horizontal axis and price dispersion on the vertical axis. Coefficient of variation values greater than .1 are truncated.

For both RP and RPHPE “true” premium rates and RPHPE sample rates, within copula premium rate dispersion decreases as dependence weakens. However, for the RP sample premium rates, within copula premium rate dispersion can increase as a function of weakening dependence as shown in Figures: 16, 18 and 20. Also, *ceterus paribus*, as coverage level increases the within copula premium rate dispersion decreases.

For both RP and RPHPE, the relationship between within copula dispersion for the sample generated premium rates—like the results of the previous simulation examining the potential for premium rate dispersion among alternative copulas—does not have a strict monotonic relationship with price dispersion or yield variation. For example, in Figures 20 and 26, the highest coefficients of variation for the median level of yield variation occur at the median level of price dispersion. Only at the lowest level of price dispersion or yield variation will the coefficients of variation consistently decrease as a function of an increase in the dispersion of the other marginal distribution.

The relationship between the sample premium rates and “true” premium rates is complicated by the fact that the sample coefficients of variation do not necessarily follow the same pattern as the coefficients of variation for the “true” premium rates. For RP premium rates at the highest coverage level, presented in Figures 19 and 20, at the median price dispersion the coefficients of variation for the “true” premium rates decreases as yield variation increases; but, the coefficients of variation for the sample premium rates increase from the lowest level of yield

variation to the median level. A similar pattern is evident for the median yield variation as the level of price dispersion increases from the lowest level to the highest. This difference must be due to the effects of sampling error in estimation of price-yield dependence. However, why sampling error in the estimation of dependence does not simply exacerbate the patterns of the dispersion in the “true” rates is unknown.

The introduction of sampling error through the estimation of dependence on only 50 correlate-pairs exacerbates dispersion. For RPHPE rates the pattern evinced by the true rates is followed by that of the sample rates; but, for RP rates this is not the case. It is unclear, by what path sampling error is traveling, such that sample rates are affected differently than the “true” rates. Like the results examining dispersion between premium rates generated by alternative copulas, sample premium rates generated within copulas become less dispersed as coverage level increases. Stronger negative price-yield dependence tends to increase the dispersion of both “true” and sample premium rates generated within copulas. The effect of dispersion in the marginal distributions is more complicated, but more monotonic results tend to occur along either the lowest level of price dispersion or yield variation. The next section of analysis investigates the causes of premium rate dispersion within copulas.

Parameter Variation Causing Dispersion

The two consistent sources of premium rate dispersion throughout the previous experiments are coverage level and dependence. Sampling error impacts

premium rate dispersion for both true and sampled rates—albeit, differently in some instances. The most direct path by which sampling error can impact premium rates is through the estimate of Kendall’s τ , which induces variation in the calibration of the copula shape parameters. Table 12 summarizes the distributional properties of the parameter estimate by copula, where in each case the hypothesized copula and “true” copula⁶⁸ are the same. The interval between lower 2.5 percentile (2.5%) and upper 97.5% (97.5%) contains 95% of the sampled parameter realizations (97.5% – 2.5%); it is divided by the median parameter estimate to eliminate scale giving a median standardize inter-quantile range of 95%/Median. Because the data are simulated the true parameter estimate is known (Copula True θ).

The true parameters for the Normal and Student-t copulas are identical for the same level of dependence and are simply the linear correlation coefficient Pearson’s ρ . By construction, the cross-rotation parameters between Clayton rotated 90° and Clayton rotated 270° and Gumbel rotated 90° and 270 are identical. In comparing the “true” copula θ parameter to the Median, it is apparent that the central tendency of the distribution for the parameter estimate is a reasonable approximations of the underlying truth. However, the median standardized measure of dispersion (95%/Median) indicates that the interval containing 95% of the parameter estimates surrounding the Median can range from 30% to over 900% of its actual median estimate. Also, as the level of dependence weakens the measure 95%/Median increases; interestingly, this pattern does not hold for either of the

⁶⁸The copula that is generating the multivariate price and yield distribution.

Table 12: Sample Parameter Deviation Compared to True Parameter

Copula	Kendall's τ	2.5%	Median	Copula True θ	97.5%	95%/Median
Normal	-0.55	-0.84	-0.73	-0.76	-0.57	0.36
Normal	-0.30	-0.63	-0.44	-0.45	-0.18	1.03
Normal	-0.05	-0.34	-0.08	-0.08	0.20	7.22
Student-t (4df)	-0.55	-0.85	-0.73	-0.76	-0.52	0.46
Student-t (4df)	-0.30	-0.66	-0.43	-0.45	-0.13	1.24
Student-t (4df)	-0.05	-0.39	-0.08	-0.08	0.26	8.54
Frank	-0.55	-9.92	-6.79	-6.73	-4.40	0.81
Frank	-0.30	-5.10	-2.95	-2.92	-1.12	1.35
Frank	-0.05	-2.26	-0.46	-0.45	1.28	7.62
Gumbel 90°	-0.55	-4.28	-2.50	-2.44	-1.30	1.19
Gumbel 90°	-0.30	-1.82	-0.87	-0.86	-0.24	1.81
Gumbel 90°	-0.05	-0.63	-0.10	-0.11	0.33	9.23
Gumbel 270°	-0.55	-4.35	-2.49	-2.44	-1.29	1.23
Gumbel 270°	-0.30	-1.80	-0.87	-0.86	-0.25	1.79
Gumbel 270°	-0.05	-0.63	-0.10	-0.11	0.34	9.32
Clayton 90°	-0.55	-3.13	-2.24	-2.22	-1.66	0.65
Clayton 90°	-0.30	-1.90	-1.43	-1.43	-1.13	0.54
Clayton 90°	-0.05	-1.32	-1.05	-1.05	1.17	2.37
Clayton 270°	-0.55	-3.11	-2.24	-2.22	-1.66	0.65
Clayton 270°	-0.30	-1.90	-1.43	-1.43	-1.12	0.54
Clayton 270°	-0.05	-1.32	-1.05	-1.05	1.17	2.37

Notes: This table shows the dispersion of the parameter estimate for each copula at each level of Kendall's τ . The 2.5 and 97.5 percentiles are shown. The difference between these percentiles (95.5%-2.5%) represents the distance between upper and lower tail realizations containing 95% of the observations. This distance, is standardized by the median θ parameter estimate for each copula and each level of dependence providing a median standardized inter-quantile range, 95%/Median. The known or "true" calibration of the copula θ parameter is also shown. The Student-t copula is calibrated with 4 degrees of freedom.

rotated Gumbel copulas. A perplexing aspect of these results is that as dependence weakens, the confidence interval surrounding the parameter estimate increases, but the results from the previous simulation experiment indicates that the dispersion of premium rates within copulas tends to decrease.

While premium rate dispersion cannot be a direct function of sampling error in the marginal distributions, it is worth investigation whether variation in price dispersion or yield variation have any effect on the distribution of the parameter

Table 13: Dispersion of Kendall's τ Compared to True Parameter

Kendall's τ	Yield CV	Price σ	2.5%	Median	97.5%	95%/Median
-0.55	0.05	0.20	-0.68	-0.55	-0.40	0.50
-0.55	0.05	0.35	-0.68	-0.55	-0.40	0.51
-0.55	0.05	0.50	-0.68	-0.55	-0.40	0.50
-0.55	0.25	0.20	-0.68	-0.55	-0.40	0.50
-0.55	0.25	0.35	-0.68	-0.55	-0.40	0.51
-0.55	0.25	0.50	-0.68	-0.55	-0.40	0.50
-0.55	0.45	0.20	-0.68	-0.55	-0.40	0.50
-0.55	0.45	0.35	-0.68	-0.55	-0.40	0.51
-0.55	0.45	0.50	-0.68	-0.55	-0.40	0.50
-0.30	0.05	0.20	-0.47	-0.30	-0.11	1.22
-0.30	0.05	0.35	-0.47	-0.30	-0.11	1.20
-0.30	0.05	0.50	-0.47	-0.30	-0.11	1.18
-0.30	0.25	0.20	-0.48	-0.30	-0.11	1.20
-0.30	0.25	0.35	-0.47	-0.30	-0.12	1.17
-0.30	0.25	0.50	-0.47	-0.30	-0.12	1.18
-0.30	0.45	0.20	-0.47	-0.30	-0.11	1.20
-0.30	0.45	0.35	-0.47	-0.30	-0.11	1.19
-0.30	0.45	0.50	-0.47	-0.30	-0.11	1.19
-0.05	0.05	0.20	-0.24	-0.05	0.15	7.59
-0.05	0.05	0.35	-0.24	-0.05	0.14	7.46
-0.05	0.05	0.50	-0.24	-0.05	0.15	7.80
-0.05	0.25	0.20	-0.24	-0.05	0.15	7.87
-0.05	0.25	0.35	-0.24	-0.05	0.14	8.00
-0.05	0.25	0.50	-0.24	-0.05	0.15	7.80
-0.05	0.45	0.20	-0.24	-0.05	0.14	7.74
-0.05	0.45	0.35	-0.25	-0.05	0.15	7.87
-0.05	0.45	0.50	-0.24	-0.05	0.15	8.27

Notes: This table shows the dispersion of the dependence estimate for each combinations of Kendall's τ , price dispersion and yield variation. The 2.5 and 97.5 percentiles are shown. The difference between these percentiles (95.5%-2.5%) represents the distance between upper and lower tail realizations containing 95% of the observations. This distance, is standardized by the median estimate of Kendall's τ at each combination of parameters providing a median standardized inter-quantile range, 95%/Median. Yield (cv) is the coefficient of variation for yields. Price σ is the dispersion in prices.

estimate. Table 13 shows the extent in which sampling error affects the estimation of dependence. The estimation of Kendall's τ does not appear to vary as a function of changes in the dispersion of the marginal distributions. The fact that sampling

error is clearly magnified at low levels of dependence, make the Gumbel results of Table 12 all the more unusual: for the Gumbel copulas the highest level of dispersion between the lower and upper percentiles came at a Kendall's τ of -.35, but the estimation of dependence clearly becomes less precise as the level of Kendall's τ weakens.

Summary

The results of this experiment indicate that regardless of the copula selected to bind simulated price and yield realizations in copula based revenue ratemaking strategy, imprecision of premium rate estimates will be greatest at low coverage levels. That coverage level had this effect on “true” premium rates, where the only potential source of variation is in the resampling of 100,000 price and yield realizations, reinforces the findings of Goodwin (2014) and Hungerford and Goodwin (2014) who observed the acute affect of copulas on determining the relationship between extreme price and yield outcomes that occur in the tail of the multivariate copula distributions.

In general, the sampling error introduced by estimating the dependence used to calibrate the copula shape parameter from 50 price-yield realizations, which come from a known copula and level of dependence, exacerbates the dispersion of premium rates generated within alternative copulas. This result is not surprising; however, that the relationship between the dispersion in the premium rates estimated from the 50 sampled observations differs from the dispersion of “true”

premium rates with respect to the combinations of dependence, price variation and yield variation is unexpected.

The most intuitive path by which sampling error can affect premium rate dispersion is through the calibration of the copula's parameter based on unstable estimates of Kendall's τ . However, the relative magnitude of the sampling error in the estimation of dependence does not appear to directly induce premium rate dispersion, at least as a function of the level of dependence. In fact, as the relative imprecision in dependence estimates increases as dependence weakens, the within copula generated premium rates become less dispersed. Although the relative magnitude of sampling error is greater at low dependence, it can have little effect on rate dispersion because of the countervailing force of the copulas becoming non-unique at weak levels of dependence. This result provides additional evidence that a copula based ratemaking procedure induces less premium rate imprecision on crops and counties with weak price-yield dependence.

The results also indicate that sampling error interacts with the levels of price and yield dispersion, but not as a function of the instability in the estimation of dependence. To account for this result, it is hypothesized that the copula, which is the only remaining source of variation, is interacting with the marginal distributions through the reshaping process. If this hypothesis holds, then it would support the conclusion from the previous experiment and studies that proper identification of the marginal distributions is imperative to a copula based ratemaking procedure for revenue premium rates.

CHAPTER TEN

AKAIKE INFORMATION CRITERION PERFORMANCE

This experiment investigates the finite sample performance of the Akaike Information Criterion (AIC) to select the “true” copula among a set of alternatives based on 50 price and yield observations. The first simulation and empirical experiments showed that premium rates generated by alternative copulas can diverge when price-yield dependence is strong and the coverage level is low. Thus, any method that better models the empirical dependence between prices and yields will reduce the potential for premium rate imprecision as a function of premium rate divergence between copulas. A method of copula selection based upon goodness-of-fit is not practical because it fails to select a copula and exclude false alternatives. The AIC has the distinct practical advantage of at least selecting a copula from a predefined set. It is, however, subject to the same fundamental statistic criticism as the goodness-of-fit tests. Fifty observations reflects the number of correlated price and yield realization upon which a copula could currently be identified to rate revenue product premium rates for multiple peril crop insurance (MPCI). The results indicate the the AIC maybe a viable method for selecting a unique copula for a rating a revenue insurance product.

Design and Data Generation

The AIC is implemented in the **R** package **CD Vine**. The parametric copulas previously examined in this study are considered.⁶⁹ The AIC also performs a test for independence, which if not rejected implies that the Independence copula (I) best fits the data. For three thousand iterations⁷⁰ each copula is calibrated by values of Kendall's τ equal to -0.55, -0.30 or -0.05 and used to bind 100,000 price and yield realizations generated from the assumed marginal price and yield distributions. The price distribution is log-normal and identified by a central tendency of \$5.78. The dispersion of the distribution is either 0.20, 0.35 or 0.50. Yields are assumed to be normally distributed and censored at zero with a mean of 100 bu./ac. and its dispersion varies by a yield standard deviation of 5 bu./ac., 25 bu.ac. or 50 bu./acre. This allows yield variation to be interpreted as a coefficient of variation, which is either 5%, 25%, or 50% of the mean. For each iteration at every unique combination of copula, dependence, and price and yield dispersion, the AIC criterion is used to select a copula that best fits the data. The number of instances in which the AIC correctly identified the “true” data generating copula are summed and divided by three thousand, providing the frequency of accurate copula identification.

⁶⁹The parametric copulas used are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Clayton rotated 90° (c90) and 270° (c270) and Gumbel rotated 90° (g90) and 270° (g270).

⁷⁰Each AIC selection was calculated over 1000 repetitions at 3 coverage levels, which do not affect the distribution of the sample data and can be aggregated, and processed separately for the 7 copulas.

Results

Table 14 shows the ability of the AIC criterion to accurately identify the “true” copula at each unique combination of dependence, yield variation and price dispersion. The AIC results follow a similar pattern as the previous goodness-of-fit results. The ability of the AIC to accurately identify the underlying data generating copula improves markedly as dependence strengthens. The effect of variation in the marginal distributions is slight, if it exists at all. As in the goodness-of-fit findings, the AIC is best able to identify the rotated representation of the Clayton copula. At the lowest level of dependence—virtual independence—the accuracy of the AIC is completely eroded; however, at very low dependence the frequency with which independence cannot be rejected increases; therefore, an Independence copula is most frequently selected. Except for the Student-t copula, all of the copulas in this study converge to the Independence copula at weak levels of dependence; therefore, the inability of the AIC criterion to accurately identify the “true” copula is offset by the convergence of premium rates generated by alternative copulas when calibrated by low levels of dependence.⁷¹

⁷¹A methodological synthesis between the AIC and the goodness-of-fit techniques would not bolster the selection of the correct copula because the inability of the goodness-of-fit tests to reject false alternatives.

Table 14: Percent Identification Akaike Information Criterion

Kendall's τ	Price σ	Yield CV	N	t4	F	c90	c270	g90	g270	I
-0.55	0.20	0.05	41	22	51	74	77	44	42	0
-0.55	0.20	0.25	42	22	51	75	76	43	43	0
-0.55	0.20	0.45	44	22	51	75	78	44	42	0
-0.55	0.35	0.05	42	22	49	75	76	42	42	0
-0.55	0.35	0.25	42	22	51	77	78	42	42	0
-0.55	0.35	0.45	42	22	50	76	75	43	43	0
-0.55	0.50	0.05	41	23	50	76	75	44	42	0
-0.55	0.50	0.25	43	22	51	77	74	44	42	0
-0.55	0.50	0.45	41	22	51	75	77	44	43	0
-0.30	0.20	0.05	24	19	32	57	59	23	21	12
-0.30	0.20	0.25	24	18	33	59	58	23	22	12
-0.30	0.20	0.45	24	20	32	56	57	23	23	12
-0.30	0.35	0.05	23	20	32	58	58	22	22	12
-0.30	0.35	0.25	24	19	33	58	57	21	23	11
-0.30	0.35	0.45	24	19	32	57	58	24	21	12
-0.30	0.50	0.05	24	19	33	57	59	22	22	12
-0.30	0.50	0.25	22	18	31	58	58	24	23	12
-0.30	0.50	0.45	24	19	32	58	58	23	23	12
-0.05	0.20	0.05	1	2	3	2	2	1	1	91
-0.05	0.20	0.25	1	2	2	2	2	1	1	92
-0.05	0.20	0.45	1	2	2	2	2	1	1	92
-0.05	0.35	0.05	2	2	3	2	2	1	1	92
-0.05	0.35	0.25	1	3	2	2	2	1	1	92
-0.05	0.35	0.45	1	3	2	2	2	1	1	91
-0.05	0.50	0.05	1	2	2	2	2	1	1	91
-0.05	0.50	0.25	1	2	2	2	2	1	1	92
-0.05	0.50	0.45	2	2	3	2	2	1	1	92

Notes: This table shows the percent accuracy of the Akaike Information Criterion to identify the "true" copula used to generate 50 price and yield realizations at each combination of price dispersion (Price σ), yield variation as measured by the coefficient of variation (Yield CV) and dependence (Kendall's τ). The frequency values are generated from 3,000 iterations. The set of copulas over which the data was generated and the AIC allowed to select, are the Normal (N), Student-t with 4 degrees of freedom (t4), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected

Summary

Both AIC and goodness-of-fit methodologies are constrained by the copula set over which selection may occur and are subject to the same statistical criticism.

However, the AIC has the distinct practical advantage of selecting a copula. The results suggest that the Akaike Information Criterion is a superior method to identify the appropriate underlying dependence structure. Future research might explore the ability of the AIC criterion to identify the “true” copula at multiple peril crop insurance data characteristics across a larger set of alternative copulas. A ratemaking practice that used the AIC to select a unique copula for a county and crop would need to incorporate the accuracy with which this selection is made, which requires further analysis. Moreover, an analysis of the performance of the AIC over finer increments of dependence would aid any effort to employ this method in practice. The test of independence appears to perform at the minimum level of statistical confidence commonly used.⁷² From a practical perspective the results of the AIC are encouraging and the impact of using an AIC selected copula and test of independence on revenue policy premium rates is explored in the next empirical experiment.

⁷²The performance of the test of independence at the lowest level of dependence never rejects a true null hypothesis more than 10% (i.e. the compliment of the results presented in Table 14), which implies that at the 90% significance level the results of the test are statistically valid.

CHAPTER ELEVEN

COPULA SELECTION STRATEGY APPLIED

In this chapter the potential for premium rate imprecision across all ratemaking methods for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion (RPHPE) are considered for nine counties in Illinois and Iowa. These counties are Hancock, IL.; McDonough, IL.; Logan, IL.; Macon, IL.; McLean, IL.; Tazwell, IL.; Cass, IA.; Delaware, IA.; and Lucas, IA. and represent a range of empirical variation in yields and price-yield dependence for corn. The empirical data are repeatedly resampled to develop distributions for the premium rates generated by the different processes. The most frequently identified copula is selected to generate farm level premium rates by county, based on the Akaike Information Criterion (AIC). The distributional properties of the premium rates generated by the AIC selected copula are compared to those generated from a Normal copula.

Data

Price and yield data are required in order to set premium rates for multiple peril crop insurance (MPCI) revenue policies.⁷³ Producers face the same effective price, which are represented by the futures markets for the commodity. Yields vary across counties, but individual producer yields within a county also vary. Both

⁷³For greater detail regarding the data used in this experiment refer to the section Data for Empirical Experiments of the first empirical experiment.

county level and producer level yield data are used in order estimate producer level premiums by county. Empirical price data are obtained from futures data for corn, which were collected from “Financial Price Data” CDs purchased from R & C Research.⁷⁴ Price data for corn is available from 1969 and were collected until 2013, providing 45 years of price realizations. County level corn yields, as measured by bushels per acre, were collected from National Agricultural Statistic Service (NASS) using the “Quick Stats” interface.⁷⁵ The data must be detrended and adjusted for heteroskedasticity before used in either the empirical rating procedure or to calibrate the copula’s dependency parameter and the censored normal marginal distribution. In addition, a county specific farm level deviation was introduced in order to estimate farm level premium rates.⁷⁶

Design

Each county has 45 years of empirical price-yield correlate pairs. One thousand RP and RPHPE premium rates are estimated by resampling from these data observations, by selecting a sub-sample of 40 price-yield pairs.⁷⁷ Premium rates are generated across the complete set of ratemaking procedures: the parametric and empirical copulas and the pure empirical method.⁷⁸ The counties represent a range in the available empirical variation for farm level yield variation and county

⁷⁴It appears that this company is no longer solvent. The raw data can be provided, upon request.

⁷⁵These data can be found at <http://quickstats.nass.usda.gov/>.

⁷⁶These data were obtained from an RMA survey (USDA, 1999) and can be found at ftp://ftp.rma.usda.gov/pub/Miscellaneous_Files/yield98/.

⁷⁷The combinatorics: $\binom{45}{40}$, yields 1,221,759 unique possible combinations.

⁷⁸For greater detail on the ratemaking procedures refer to the section Ratemaking Procedures in the first empirical experiment.

price-yield dependence, as measured by Kendall's τ . Four of the Illinois counties:

Logan, Macon, McLean and Tazwell were examined by Goodwin (2014).

The empirical data was divided into three equally populated groups of yield variation and price-yield dependence. These groups were intersected to create nine groups with categorically unique intersections of the empirical yield and dependence characteristics. The groups in Table 31 are presented from strong price-yield dependence and low yield variation in the top left corner to weak dependence and low yield variation in the bottom right.⁷⁹ Counties were selected based on their being representative of the median characteristics of the group they occupy. The four counties studied by Goodwin (2014) are fairly representative of the groups they fall within and where in some cases selected over slightly more representative counties for potential comparison.

In the pure empirical method, rates are generated directly from the 40 sampled price-yield observations. For the copula methods, which require simulation from assumed marginal distributions, the copula's shape parameter is calibrated by estimating the dependence between the sub-sample's price and yield pairs.⁸⁰

However, as in the previous experiments, variation from the yield marginal is eliminated because the assumed censored normal distribution is identified by a sample invariant estimate of the county level mean and county yield variation based

⁷⁹When subsequent results are grouped across counties they are presented in the order of Table 31—from top-left to bottom-right.

⁸⁰For greater detail of the ratemaking procedure refer to the section Ratemaking Procedures in the first empirical experiment.

on the complete 45 years of data.⁸¹ Prices are log-normally distributed and generated with a projected price of \$5.78 and price dispersion, which is either 0.20, 0.35 or 0.50. Premium rates are estimated for coverage levels of 50%, 65%, or 85%.

Results

First, the dispersion of premium rates generated by the copulas is compared to the pure empirical method. Next, the frequency results of the AIC criterion for each county in the study are presented. Finally, the selection of an AIC copula for each county is used to generate premium rates and compared to the premium rates generated by the Normal copula, which is currently used by the Risk Management Agency (RMA) to develop revenue policy premium rates.

Empirical Dispersion Within Copulas

The effect of coverage level and price dispersion on RP and RPHPE premium rate dispersion across the study counties is explored graphically in Figures 27 through 32. The scale issue across coverage levels and between copula methods and the pure empirical method necessitated a similar standardization of premium rates, as was used in the previous experiments. In this case, premium rates are standardized by their deviation from the mean premium rate, across both copula and the empirical methods, for each county, at each repetition and level of price dispersion. This allows the mean standardized premium rate deviation to be

⁸¹The implicit assumptions being that the censored normal distribution reflects the shape of producer yields and that 45 years of data identifies the distribution.

compactly represented using subplots. The subplots are divided horizontally between copula and pure empirical rate deviations and vertically by price dispersion. The distribution of the premium rate deviations are presented by county within each sub-plot. The red dashed line marks the location of no deviation from the mean premium rate or a value of one on the deviation axis.

The global difference between premium rates generated by the copula method and pure empirical method make the figures difficult to interpret at the 50% and 65% coverage levels. However, this result reinforces the previous finding that low coverage levels tend to induce premium rate divergence between each copula and between the copula generated premium rates and the pure empirical premium rates. Moreover, it appears that premium rate divergence generated within the copula and pure empirical methods is greatest at the lowest coverage level. Thus, premium rate imprecision caused by sampling error is not only greatest for copulas at low coverage levels, but also for the pure empirical method. This result reinforces the previous claim that the identification of the marginal distributions is critical not only to copula modeling, but to an alternative non-copula based ratemaking procedure. Sampling error is likely exacerbated at low coverage levels because relatively few realizations, in the tail of the bi-variate distributions, are determining premium rates.

As price dispersion increases, across both insurance products and all coverage levels, premium rates tend to converge to the mean. Moreover, from numerical analysis of the distribution properties of these data, within county copula generated

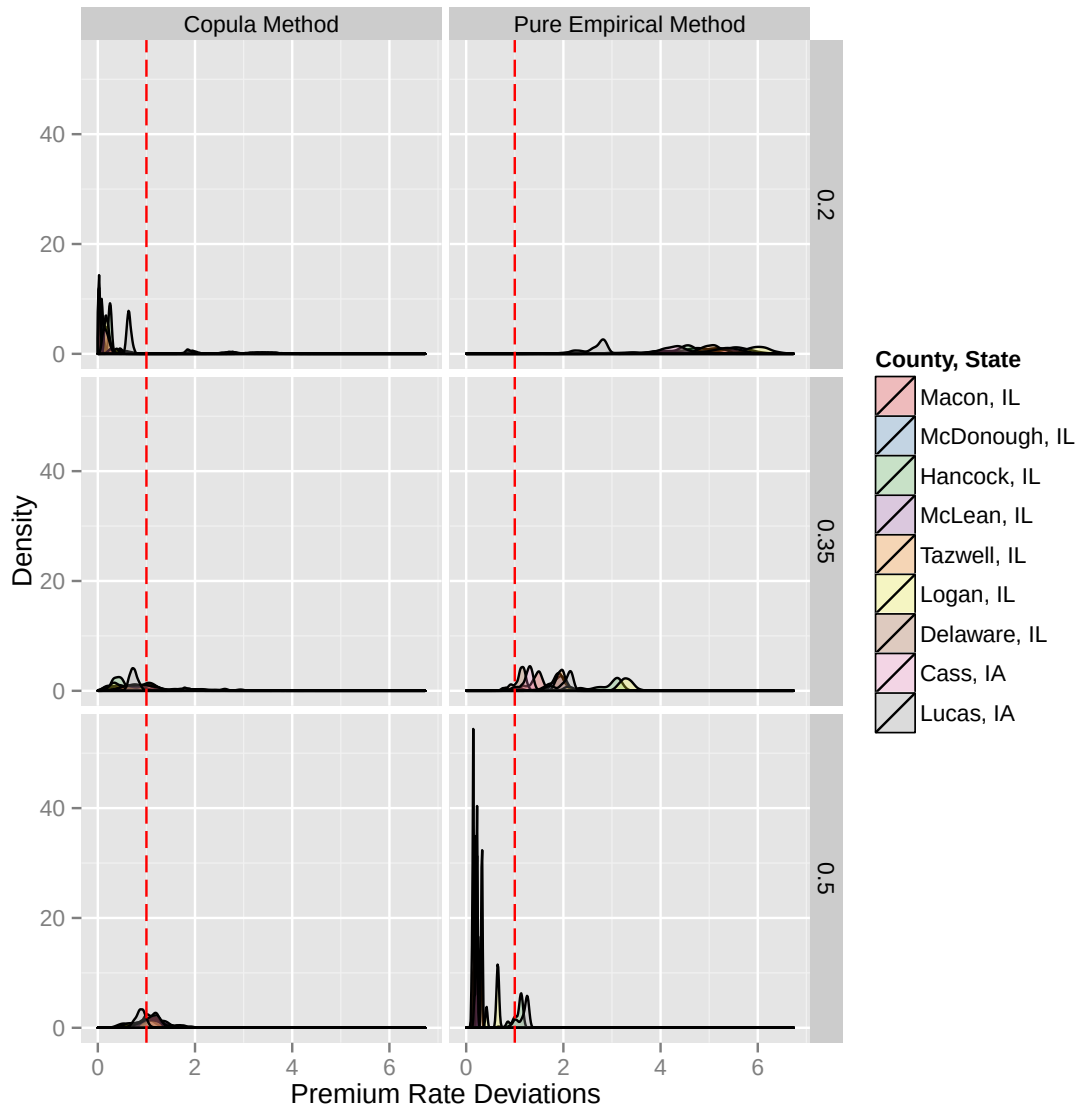


Figure 27: Empirical RP Deviations 50% Coverage

Notes: Revenue Protection (RP) premium rates are standardized by their deviation from the mean premium rate calculated between both copula and the empirical methods, for each county, for each repetition and level of price dispersion. One thousand repetitions populate the distributions at each unique categorical combination. The subplots are divided vertically by price dispersion. The red dashed line marks the location of no deviation from the mean rate or a value of one on the deviation axis.

premium rate deviations become more tightly distributed as price dispersion

increases in all cases. However, the pure empirical method premium rate deviations

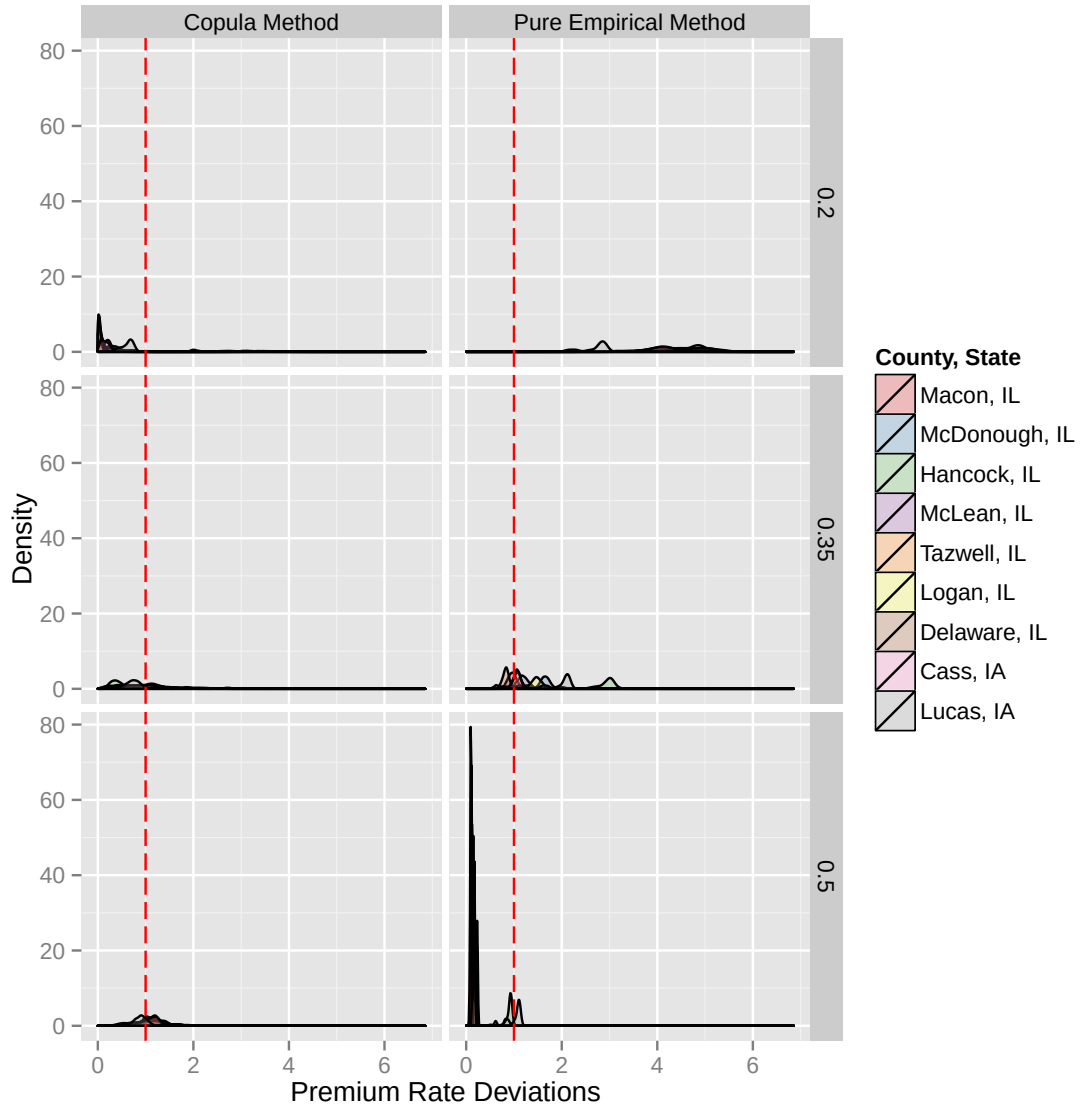


Figure 28: Empirical RPHPE Deviations 50% Coverage

Notes: Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates are standardized by their deviation from the mean premium rate calculated between both copula and the empirical methods, for each county, for each repetition and level of price dispersion. One thousand repetitions populate the distributions at each unique categorical combination. The subplots are divided vertically by price dispersion. The red dashed line marks the location of no deviation from the mean rate or a value of one on the deviation axis.

do not appear to respond to changes in price dispersion.⁸²

From analysis of the data in tabular form, the within county copula

⁸²The numeric tables representing these data are large and effectively analyzing them requires sorting them along different dimensions.

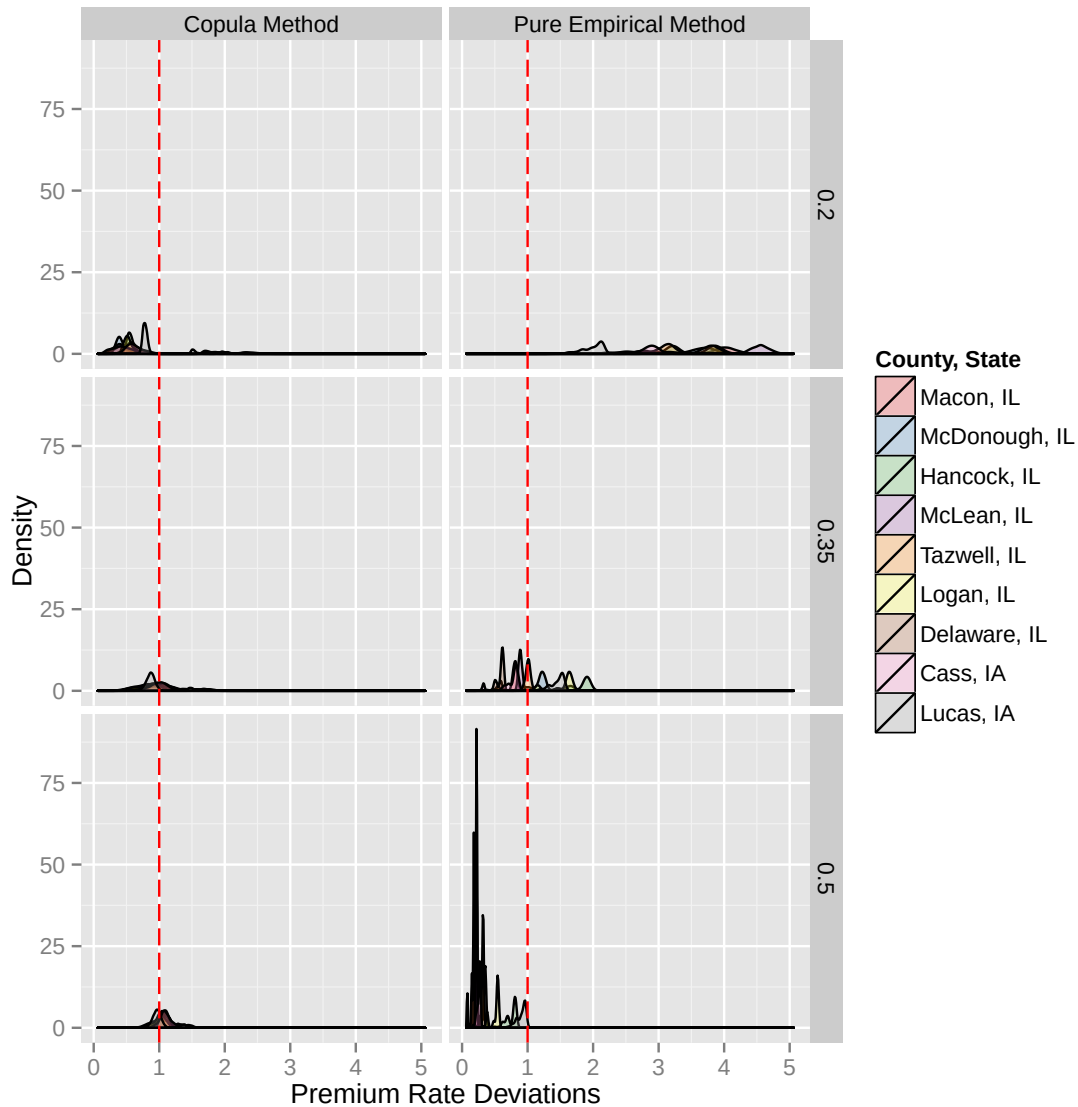


Figure 29: Empirical RP Deviations 65% Coverage Level

Notes: Revenue Protection (RP) premium rates are standardized by their deviation from the mean premium rate calculated between both copula and the empirical methods, for each county, for each repetition and level of price dispersion. One thousand repetitions populate the distributions at each unique categorical combination. The subplots are divided vertically by price dispersion. The red dashed line marks the location of no deviation from the mean rate or a value of one on the deviation axis.

generated premium rate deviations become more tightly distributed as coverage level increases and the pure empirical method deviations tend to become more tightly distributed as well. Also, as price dispersion increases the relationship

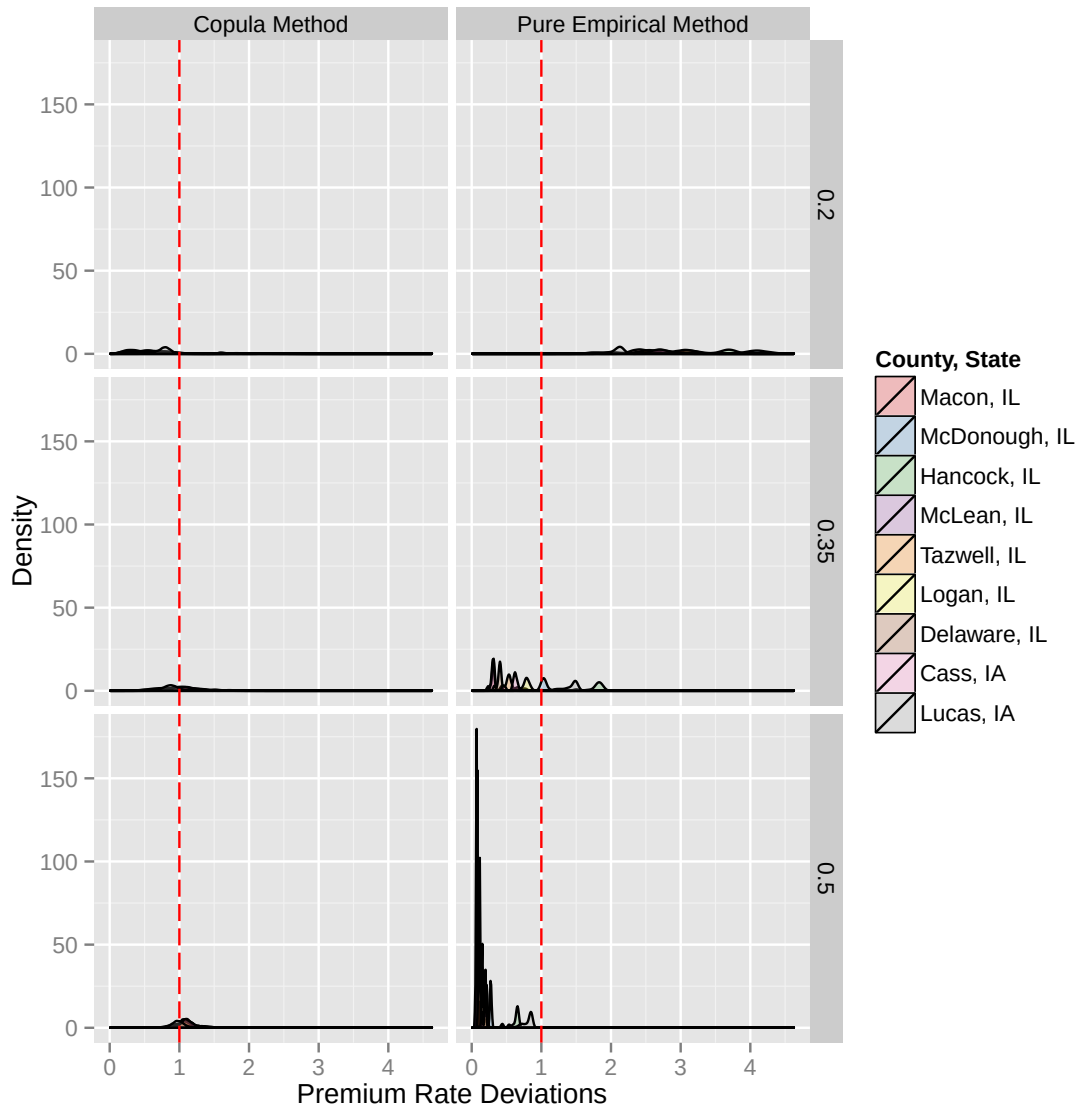


Figure 30: Empirical RPHPE Deviations 65% Coverage Level

Notes: Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates are standardized by their deviation from the mean premium rate calculated between both copula and the empirical methods, for each county, for each repetition and level of price dispersion. One thousand repetitions populate the distributions at each unique categorical combination. The subplots are divided vertically by price dispersion. The red dashed line marks the location of no deviation from the mean rate or a value of one on the deviation axis.

between the copula and pure empirical deviations relative to the mean reverses,

across all coverage levels.⁸³ At the lowest level of price dispersion, pure empirical

⁸³The pattern is at its most extreme at the lowest coverage level.

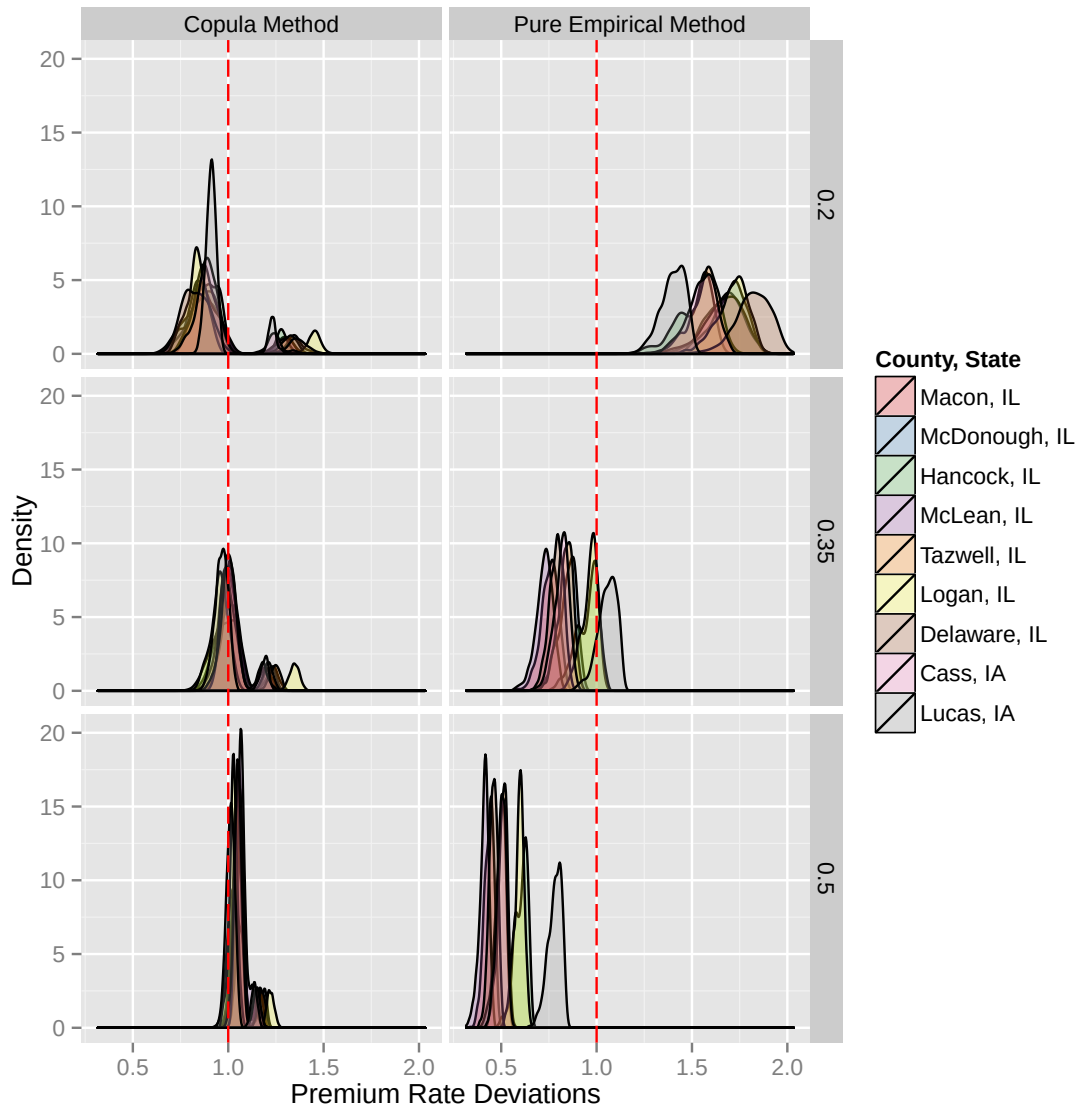


Figure 31: Empirical RP Deviations 85% Coverage

Notes: Revenue Protection (RP) premium rates are standardized by their deviation from the mean premium rate calculated between both copula and the empirical methods, for each county, for each repetition and level of price dispersion. One thousand repetitions populate the distributions at each unique categorical combination. The subplots are divided vertically by price dispersion. The red dashed line marks the location of no deviation from the mean rate or a value of one on the deviation axis.

premium rates are decidedly greater—for all counties—than the mean premium rate

calculated across all ratemaking methods.⁸⁴ The apparent bi-modality in the copula

⁸⁴The pure empirical method tends to generate lower premium rates than the mean premium rate at the highest level of price dispersion, except for Lucas, IA. and Hancock, IA at the lowest coverage

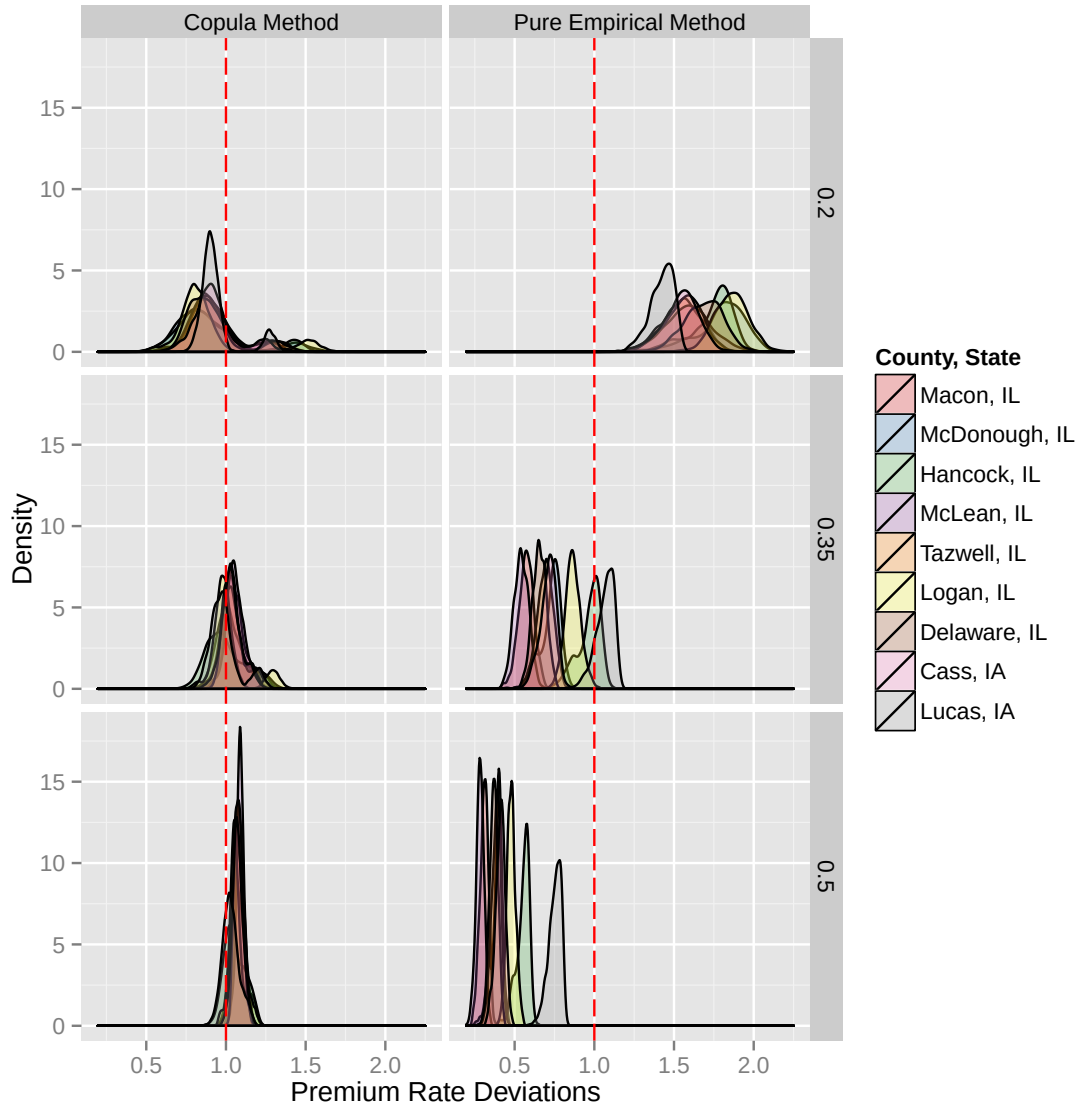


Figure 32: Empirical RPHPE Deviations 85% Coverage

Notes: Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates are standardized by their deviation from the mean premium rate calculated between both copula and the empirical methods, for each county, for each repetition and level of price dispersion. One thousand repetitions populate the distributions at each unique categorical combination. The subplots are divided vertically by price dispersion. The red dashed line marks the location of no deviation from the mean rate or a value of one on the deviation axis.

level. These counties are both characterized by high levels of yield variation and the other high yield variation county—Logan, IL—follows a similar pattern. This result sheds some light on the previously discovered empirical result across all counties presented in Figures 29 and 30, in which even at the highest level of price dispersion some counties generated higher pure empirical premium rates than copula premium rates. The results presented in Figures 27 and 28 indicate that, in general, counties with high yield variation may be those for which the pure empirical premium rates remain higher

method sub-figures, represents the empirical copula's inclusion. In all cases the empirical copula has deviations greater than the parametric copulas. The empirical copula deviations follow the overarching pattern of the data and center to the county mean as a function of increasing coverage level and price dispersion. Unlike the pure empirical method premium rate deviations, they are always greater than the parametric copula deviations; the reason for this result must originate from the parametric marginals, but further research or theoretical insight is required.

Akaike Information Criterion Selection

In practice, the underlying higher order dependence between prices and yields is unknown and must be modeled. It has been shown that available goodness-of-fit statistics cannot confidently discriminate between copulas at the small sample sizes and low levels of dependence that characterize multiple peril crop insurance (MPCI). While the Akaike Information Criterion was at most 60% accurate in identifying the “true” underlying copula in the simulation results, it has the distinct methodological advantage of at least selecting a copula. Therefore, any attempt to generate more accurate premium rates based upon modeling the unique county level price-yield dependence would be better served by using the AIC as the selection criterion.

Table 15 summarizes the AIC selection by county by showing the percent frequency in which the alternative copulas are selected for each county. The table is further separated by the overarching price dispersion environment. The results

than copula premium rates at the lowest coverage level and highest level of price dispersion.

Table 15: AIC Copula Selection for Empirical Study

Copula	Price σ	Macon	McDonough	Hancock	McLean	Tazwell	Logan	Delaware	Cass	Lucas
N	0.20	0.00	0.00	0.00	0.36	0.00	0.00	1.31	0.00	0.00
N	0.20	0.00	0.00	0.00	0.13	0.00	0.00	1.19	0.00	0.00
N	0.20	0.00	0.00	0.00	0.24	0.00	0.00	1.46	0.00	0.00
t	0.20	0.00	30.08	31.27	0.00	0.00	0.00	0.00	4.78	4.13
t	0.20	0.00	30.10	31.23	0.00	0.00	0.00	0.00	5.12	4.21
t	0.20	0.00	29.89	31.14	0.00	0.00	0.00	0.00	5.06	4.09
F	0.20	0.07	0.03	0.04	0.02	0.01	0.00	0.00	0.00	0.00
F	0.20	0.03	0.02	0.07	0.04	0.01	0.00	0.00	0.03	0.00
F	0.35	0.01	0.02	0.01	0.04	0.01	0.00	0.00	0.02	0.00
g90	0.35	1.86	2.93	0.04	1.92	2.01	1.96	1.96	7.16	7.61
g90	0.35	1.94	2.74	0.07	2.32	1.74	2.09	1.93	6.76	7.51
g90	0.35	2.03	3.08	0.10	1.86	1.99	1.90	1.51	7.02	7.56
g270	0.35	0.00	0.11	1.68	0.00	0.00	0.00	0.01	2.72	0.49
g270	0.35	0.00	0.28	1.79	0.00	0.00	0.00	0.00	2.41	0.56
g270	0.35	0.00	0.11	1.81	0.00	0.00	0.00	0.00	2.59	0.56
c90	0.35	0.00	0.00	0.30	0.00	0.00	0.00	0.00	0.00	0.00
c90	0.50	0.00	0.00	0.18	0.00	0.00	0.00	0.00	0.00	0.00
c90	0.50	0.00	0.00	0.27	0.00	0.00	0.00	0.00	0.00	0.00
c270	0.50	30.28	0.18	0.00	31.03	30.80	31.38	1.64	0.68	1.64
c270	0.50	30.14	0.19	0.00	30.83	31.01	31.24	1.58	0.63	1.43
c270	0.50	30.11	0.23	0.00	31.19	30.80	31.43	1.76	0.69	1.56
I	0.50	1.13	0.00	0.00	0.00	0.51	0.00	28.41	18.00	19.46
I	0.50	1.21	0.00	0.00	0.00	0.57	0.00	28.63	18.38	19.62
I	0.50	1.18	0.00	0.00	0.00	0.53	0.00	28.61	17.96	19.58

Notes: This table shows the percent frequency in which the Akaike Information Criterion selected each copula in the available set. The frequency values are generated from 1,000 iterations. The set of copulas are the Normal (N), Student-t with calibrate degrees of freedom (t), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected. Price dispersion is denoted Price σ .

reinforce those presented in Table 14 of the simulation experiment, which show that AIC selection is invariant to changes in the dispersion of the price marginal distribution. Therefore, the results can essentially be aggregated across the same levels of price dispersion. For only one county is the Normal copula selected with any frequency, which is for the weakest dependence county—Delaware, IA, which also has the Independence copula as its most frequently selected copula. Neither the Frank nor the Clayton rotated 90° was selected for any county. The Clayton rotated 90° and Gumbel rotated 270° copulas represent a condition in which price-yield dependence becomes stronger during high-price and low-yield outcomes. Both copulas with this shape were selected infrequently. At least for the counties in this study, modeling price-yield outcomes based upon a theoretical notion that prices and yields become more dependent as price increases, would be unsound. From the simulation results these copulas tend to generate higher premium rates.

In contrast to the overarching shape imposed by either the Clayton rotated 90° or Gumbel rotated 270° , for all counties the Gumbel rotated 90° was selected with some frequency. Moreover, for all counties in the middle and low dependence groups and for one of the highest dependence counties, the Clayton rotated 270° copula was selected most frequently. The shape of both of these copulas represent a situation in which the dependence between price and yields becomes stronger at low-price and high-yield outcomes. These results indicate that county yields tend to be better modeled by a higher order dependence that captures a price-yield dependence that increases as prices fall. From the simulation results, these copulas

tended to generate lower premium rates because low price realizations are more likely to be offset by a high yield realization, with no affect on revenue.

For the two remaining high dependence counties, the Student-t copula was selected with the greatest frequency. This copula allows for an increased dependence in either tail. While the simulation results showed that at low dependence the Student-t copula could have a radically different impact on premium rates than all other copulas, these counties' have a relatively high level of price-yield dependence.⁸⁵

For the counties in the weak dependence group, the independence of price and yields could not be rejected and the Independence copula was selected with the greatest frequency. For the most weakly dependent county in the study—Delaware, IA—the frequency of selection is similar to that of the more highly dependent counties. The two remaining low dependence counties have a lower frequency of selection than any of the other counties, but the Independence copula is selected more frequently than the alternatives. For these two counties the Gumbel rotated 90° copula is selected with greater frequency than for any of the other counties. The inability to reject independence is entirely consistent with the weak price-yield dependence that characterizes these counties.⁸⁶

The results of the AIC selection criterion are internally consistent and

⁸⁵The AIC criterion maximizes the fit of the Student-t copula over its dependence and degrees of freedom parameters; the results will be presented subsequently, but 2 degrees of freedom was found to maximize fit of the Student-t copula for the two counties in this study.

⁸⁶The frequency with which the Gumbel rotated 90° is alternatively selected is slightly puzzling: while the Gumbel copula imposes a less structured shape and less tail dependence at higher levels of dependence, at lower levels it imposes a greater structure and tail dependence than its cross rotated Clayton counterpart. Figure 35 illustrates the difference in tail dependence imposed by Clayton and Gumbel copulas. It is not known why, as the data becomes less structured at low dependence, the more structured alternative between the Clayton and Gumbel copulas would be selected.

consistent with previous results from this study. The Clayton and Gumbel copula rotations that represent similar overall price-yield dependence are consistently either selected or not and counties with very low dependence tend to be best modeled by the Independence copula. There is a significant frequency difference between the most frequently selected copula and the next most frequently selected, which indicates that among this limited set of alternative copulas there is a preferred copula. For no county, is the Normal copula the most frequently selected. Future research could expand the set of potential copulas in order allow more flexibility in the unique copula selected for rating revenue insurance products.

The initial analysis of the data was performed on a sub-set of only 200 resampled permutations and the results did not change markedly when all 1,000 repetitions were considered. Therefore, modeling county level dependence for MPCIR revenue products with an AIC method does not require one-thousand repetitions and is a practical alternative to the current practice of simply assuming a Normal copula. Of course, if MPCIR revenue insurance premium rates are tailored by an AIC selected copula, the uncertainty surrounding the copula model selection would need to be incorporated. The potential impact on premium rates of using an AIC selection criterion, when compared to the current assumption of a Normal copula, is examined in the following section.

Copula Selection Premium Rates

Given the AIC copula selection, the properties of the premium rates can be analyzed and compared to the properties of rates generated by an assumed Normal copula. The analysis is based only upon instances in which the most frequently selected AIC identified copula for each county generated the premium rate. Thus, the number of observations upon which subsequent summary statistics and analysis varies by county.⁸⁷

Tables 16 through 24 summarize the difference in distributional properties of the AIC selected copula premium rates and premium rates generated by an assumed Normal copula. Macon county exhibits a strong level of price yield dependence, while Delaware county has the weakest level of dependence in the subset of counties. Values of the Median Ratio measure in Tables 16 through 24 greater than one indicate that the median premium rate generated by the Normal copula is greater than the median premium rate generated by the AIC selected copula.⁸⁸ Values of the Dispersion Ratio measure in the Tables 16 through 24 greater than one indicate that the dispersion of premium rates generated by the Normal copula is greater

⁸⁷Table 32 presents the number of observations upon which summary statistics for AIC selected copula are generated. The complete data set was not used for two reasons. First, the AIC selected the Student-t copula for two counties, but the AIC calibrated the degrees of freedom independently for each repetition, whereas the degrees of freedom parameter for the Student-t in the parametric copula set was always calibrated to four. Second, the Independence copula was selected for the three low dependence counties, but the Independence copula was not included in the original parametric copula set. As can be seen from Table 32 the level difference in observations between counties does not vary radically, the only exceptions being the two of the low dependence counties that selected, but with less relative frequency, the Independence copula.

⁸⁸The Median Ratio measure is calculated by dividing the median premium rate generated by the Normal copula by the median premium rate generated by the AIC selected copula.

Table 16: Macon, IL AIC and Normal Copula Premium Rates

Policy	Coverage	Price σ	AIC Copula	Median Ratio	Dispersion Ratio
RP	50%	0.20	c270	0.36	1.54
RP	50%	0.35	c270	2.07	0.87
RP	50%	0.50	c270	1.39	0.74
RP	50%	0.20	c270	0.74	1.54
RP	50%	0.35	c270	1.32	0.85
RP	50%	0.50	c270	1.07	0.94
RP	65%	0.20	c270	1.11	1.03
RP	65%	0.35	c270	1.04	1.08
RP	65%	0.50	c270	1.00	1.22
RPHPE	65%	0.20	c270	0.08	3.40
RPHPE	65%	0.35	c270	2.07	0.87
RPHPE	65%	0.50	c270	1.39	0.74
RPHPE	85%	0.20	c270	0.53	2.28
RPHPE	85%	0.35	c270	1.31	0.85
RPHPE	85%	0.50	c270	1.07	0.95
RPHPE	85%	0.20	c270	1.02	1.27
RPHPE	85%	0.35	c270	1.01	1.20
RPHPE	85%	0.50	c270	0.98	1.35

Notes: This table shows the different distributional properties of premium rate estimates for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion policies, generated by a copula selected with the Akaike Information Criterion (AIC) versus a Normal copula. The measure Median Ratio is the median premium rate estimate generated by the Normal copula divided by the median premium rate estimate generated by the AIC selected copula. The Dispersion Ratio measures the relative dispersion between premium rates estimated from a Normal copula versus an AIC selected Copula. It is calculated by taking the difference between the lower 2.5 percentile and upper 97.5 percentile of premium rate estimates generated by the normal copula, dividing them by the median estimate and dividing this value by the same formula applied to the premium rates generated by the AIC copula. The set of copulas over which the AIC could select are the Normal (N), Student-t with degrees of freedom calibrated by the AIC (t), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected. The dispersion in prices is denoted Price σ .

than the dispersion of premium rates generated by the AIC selected copula.⁸⁹

In order to limit the analysis, the results for RP premium rates generated at the 85% coverage level are discussed with greater detail. The Median Ratio measure

⁸⁹The Dispersion Ratio measure was calculated by dividing the difference between the difference between the upper 97.5 percentile and lower 2.5 percentile premium rates generated by the Normal copula by the median Normal copula premium rate. This value is then divided by the difference between the upper 97.5 percentile and lower 2.5 percentile of premium rates generated by the AIC selected copula, which are divided median AIC copula premium rate.

Table 17: McDonough, IL AIC and Normal Copula Premium Rates

Policy	Coverage	Price σ	AIC Copula	Median Ratio	Dispersion Ratio
RP	50%	0.20	t	0.11	1.37
RP	50%	0.35	t	0.40	1.58
RP	50%	0.50	t	1.10	1.16
RP	50%	0.20	t	0.53	1.31
RP	50%	0.35	t	1.15	1.03
RP	50%	0.50	t	1.15	1.10
RP	65%	0.20	t	1.16	1.12
RP	65%	0.35	t	1.13	1.01
RP	65%	0.50	t	1.06	1.09
RPHPE	65%	0.20	t	0.03	3.70
RPHPE	65%	0.35	t	0.39	1.63
RPHPE	65%	0.50	t	1.10	1.16
RPHPE	85%	0.20	t	0.39	1.98
RPHPE	85%	0.35	t	1.16	1.05
RPHPE	85%	0.50	t	1.15	1.11
RPHPE	85%	0.20	t	1.47	0.89
RPHPE	85%	0.35	t	1.20	1.04
RPHPE	85%	0.50	t	1.08	1.14

Notes: This table shows the different distributional properties of premium rate estimates for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion policies, generated by a copula selected with the Akaike Information Criterion (AIC) versus a Normal copula. The measure Median Ratio is the median premium rate estimate generated by the Normal copula divided by the median premium rate estimate generated by the AIC selected copula. The Dispersion Ratio measures the relative dispersion between premium rates estimated from a Normal copula versus an AIC selected Copula. It is calculated by taking the difference between the lower 2.5 percentile and upper 97.5 percentile of premium rate estimates generated by the normal copula, dividing them by the median estimate and dividing this value by the same formula applied to the premium rates generated by the AIC copula. The set of copulas over which the AIC could select are the Normal (N), Student-t with degrees of freedom calibrated by the AIC (t), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected. The dispersion in prices is denoted Price σ .

indicates that the the Normal copula generates a higher median premium rate than the AIC selection for all middle and high dependence county groups, regardless of underlying yield variation and across the range of price dispersion.⁹⁰ Thus, for the

⁹⁰As price dispersion increases the magnitude of the Median Ratio decreases, indicating that the median rates from the Normal and AIC selection copulas tend to converge as price dispersion increases. Consistent exceptions from this trend arise at the lowest price dispersion at the two lower coverage levels for RPHPE premium rates. Thus, during low price dispersion periods, farmers within higher dependence counties are being charged radically reduced RPHPE rates with a Normal copula

Table 18: Hancock, IL AIC and Normal Copula Premium Rates

Policy	Coverage	Price σ	AIC Copula	Median Ratio	Dispersion Ratio
RP	50%	0.20	t	0.56	0.94
RP	50%	0.35	t	0.65	1.51
RP	50%	0.50	t	1.32	1.07
RP	50%	0.20	t	0.90	1.35
RP	50%	0.35	t	1.30	1.21
RP	50%	0.50	t	1.32	1.07
RP	65%	0.20	t	1.16	1.25
RP	65%	0.35	t	1.20	0.94
RP	65%	0.50	t	1.10	1.08
RPHPE	65%	0.20	t	0.42	1.57
RPHPE	65%	0.35	t	0.61	1.69
RPHPE	65%	0.50	t	1.34	1.07
RPHPE	85%	0.20	t	1.09	1.31
RPHPE	85%	0.35	t	1.51	1.08
RPHPE	85%	0.50	t	1.38	1.07
RPHPE	85%	0.20	t	1.63	0.84
RPHPE	85%	0.35	t	1.43	0.83
RPHPE	85%	0.50	t	1.18	1.07

Notes: This table shows the different distributional properties of premium rate estimates for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion policies, generated by a copula selected with the Akaike Information Criterion (AIC) versus a Normal copula. The measure Median Ratio is the median premium rate estimate generated by the Normal copula divided by the median premium rate estimate generated by the AIC selected copula. The Dispersion Ratio measures the relative dispersion between premium rates estimated from a Normal copula versus an AIC selected Copula. It is calculated by taking the difference between the lower 2.5 percentile and upper 97.5 percentile of premium rate estimates generated by the normal copula, dividing them by the median estimate and dividing this value by the same formula applied to the premium rates generated by the AIC copula. The set of copulas over which the AIC could select are the Normal (N), Student-t with degrees of freedom calibrated by the AIC (t), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected. The dispersion in prices is denoted Price σ .

most popular product at the highest available coverage level, it is likely that farmers within many Illinois corn producing counties are being charged a higher premium rate than their county level price-yield dependence warrants. This results holds for the higher dependence counties regardless of whether the AIC selected the Clayton

dependent rating process, than their AIC selected dependence would suppose—if they elected the lowest coverage level.

Table 19: McLean, IL AIC and Normal Copula Premium Rates

Policy	Coverage	Price σ	AIC Copula	Median Ratio	Dispersion Ratio
RP	50%	0.20	c270	0.39	1.48
RP	50%	0.35	c270	2.43	0.81
RP	50%	0.50	c270	1.48	0.87
RP	50%	0.20	c270	0.71	1.38
RP	50%	0.35	c270	1.44	0.88
RP	50%	0.50	c270	1.09	1.04
RP	65%	0.20	c270	1.13	1.11
RP	65%	0.35	c270	1.05	1.16
RP	65%	0.50	c270	1.00	1.21
RPHPE	65%	0.20	c270	0.02	11.66
RPHPE	65%	0.35	c270	2.44	0.81
RPHPE	65%	0.50	c270	1.48	0.87
RPHPE	85%	0.20	c270	0.42	2.31
RPHPE	85%	0.35	c270	1.43	0.87
RPHPE	85%	0.50	c270	1.09	1.04
RPHPE	85%	0.20	c270	1.03	1.32
RPHPE	85%	0.35	c270	1.01	1.28
RPHPE	85%	0.50	c270	0.98	1.27

Notes: This table shows the different distributional properties of premium rate estimates for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion policies, generated by a copula selected with the Akaike Information Criterion (AIC) versus a Normal copula. The measure Median Ratio is the median premium rate estimate generated by the Normal copula divided by the median premium rate estimate generated by the AIC selected copula. The Dispersion Ratio measures the relative dispersion between premium rates estimated from a Normal copula versus an AIC selected Copula. It is calculated by taking the difference between the lower 2.5 percentile and upper 97.5 percentile of premium rate estimates generated by the normal copula, dividing them by the median estimate and dividing this value by the same formula applied to the premium rates generated by the AIC copula. The set of copulas over which the AIC could select are the Normal (N), Student-t with degrees of freedom calibrated by the AIC (t), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected. The dispersion in prices is denoted Price σ .

rotated 90° or Student-t copula.⁹¹ The Dispersion Ratio indicates that the dispersion of premium rates generated by the Normal tend to be greater than that the AIC selected copula. It appears that the magnitude of this metric is larger for

⁹¹A similar pattern is evident at the highest coverage level for the RPHPE product; however, the smaller magnitudes of the Median Ratio, as well as the decreasing trajectory of the measure as a function of price dispersion, tends to produce lower median AIC selected copula RPHPE premium rates at the highest level of price dispersion.

Table 20: Tazwell, IL AIC and Normal Copula Premium Rates

Policy	Coverage	Price σ	AIC Copula	Median Ratio	Dispersion Ratio
RP	50%	0.20	c270	0.66	1.22
RP	50%	0.35	c270	1.49	1.23
RP	50%	0.50	c270	1.56	0.81
RP	50%	0.20	c270	0.87	1.57
RP	50%	0.35	c270	1.40	0.95
RP	50%	0.50	c270	1.13	1.02
RP	65%	0.20	c270	1.12	1.16
RP	65%	0.35	c270	1.08	1.04
RP	65%	0.50	c270	1.02	1.23
RPHPE	65%	0.20	c270	0.21	2.86
RPHPE	65%	0.35	c270	1.47	1.23
RPHPE	65%	0.50	c270	1.56	0.81
RPHPE	85%	0.20	c270	0.57	2.57
RPHPE	85%	0.35	c270	1.38	0.98
RPHPE	85%	0.50	c270	1.13	1.04
RPHPE	85%	0.20	c270	1.04	1.45
RPHPE	85%	0.35	c270	1.04	1.20
RPHPE	85%	0.50	c270	1.00	1.41

Notes: This table shows the different distributional properties of premium rate estimates for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion policies, generated by a copula selected with the Akaike Information Criterion (AIC) versus a Normal copula. The measure Median Ratio is the median premium rate estimate generated by the Normal copula divided by the median premium rate estimate generated by the AIC selected copula. The Dispersion Ratio measures the relative dispersion between premium rates estimated from a Normal copula versus an AIC selected Copula. It is calculated by taking the difference between the lower 2.5 percentile and upper 97.5 percentile of premium rate estimates generated by the normal copula, dividing them by the median estimate and dividing this value by the same formula applied to the premium rates generated by the AIC copula. The set of copulas over which the AIC could select are the Normal (N), Student-t with degrees of freedom calibrated by the AIC (t), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected. The dispersion in prices is denoted Price σ .

the RPHPE policy than RP policy. This result may provide some ancillary support for the AIC selected copulas as they tend generate a more precise premium rate estimate than their Normal counterpart.

In contrast to the previous experiments the most stark results come from the counties representing the lowest dependence. Across the three counties, variation in

Table 21: Logan, IL AIC and Normal Copula Premium Rates

Policy	Coverage	Price σ	AIC Copula	Median Ratio	Dispersion Ratio
RP	50%	0.20	c270	0.72	1.15
RP	50%	0.35	c270	1.13	1.42
RP	50%	0.50	c270	1.61	0.90
RP	50%	0.20	c270	0.86	1.34
RP	50%	0.35	c270	1.35	1.08
RP	50%	0.50	c270	1.14	1.10
RP	65%	0.20	c270	1.10	1.28
RP	65%	0.35	c270	1.08	1.14
RP	65%	0.50	c270	1.02	1.29
RPHPE	65%	0.20	c270	0.18	3.18
RPHPE	65%	0.35	c270	1.03	1.57
RPHPE	65%	0.50	c270	1.61	0.89
RPHPE	85%	0.20	c270	0.47	2.94
RPHPE	85%	0.35	c270	1.32	1.13
RPHPE	85%	0.50	c270	1.13	1.11
RPHPE	85%	0.20	c270	0.99	1.65
RPHPE	85%	0.35	c270	1.03	1.34
RPHPE	85%	0.50	c270	0.99	1.46

Notes: This table shows the different distributional properties of premium rate estimates for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion policies, generated by a copula selected with the Akaike Information Criterion (AIC) versus a Normal copula. The measure Median Ratio is the median premium rate estimate generated by the Normal copula divided by the median premium rate estimate generated by the AIC selected copula. The Dispersion Ratio measures the relative dispersion between premium rates estimated from a Normal copula versus an AIC selected Copula. It is calculated by taking the difference between the lower 2.5 percentile and upper 97.5 percentile of premium rate estimates generated by the normal copula, dividing them by the median estimate and dividing this value by the same formula applied to the premium rates generated by the AIC copula. The set of copulas over which the AIC could select are the Normal (N), Student-t with degrees of freedom calibrated by the AIC (t), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected. The dispersion in prices is denoted Price σ .

coverage levels and levels of price dispersion, the Normal copula generates radically lower premium rates, than the AIC selected Independence copula. This difference is ameliorated as coverage level increases. However, at the highest coverage level the median Normal copula premium rates tends to be no more than 95% of the median AIC selected premium rates and at the lowest level of price dispersion it can still be

Table 22: Delaware, IA AIC and Normal Copula Premium Rates

Policy	Coverage	Price σ	AIC Copula	Median Ratio	Dispersion Ratio
RP	50%	0.20	I	0.04	6.64
RP	50%	0.35	I	0.36	6.30
RP	50%	0.50	I	0.66	4.63
RP	50%	0.20	I	0.23	10.53
RP	50%	0.35	I	0.60	7.16
RP	50%	0.50	I	0.80	4.68
RP	65%	0.20	I	0.71	8.53
RP	65%	0.35	I	0.85	5.11
RP	65%	0.50	I	0.92	3.00
RPHPE	65%	0.20	I	0.03	7.23
RPHPE	65%	0.35	I	0.36	6.30
RPHPE	65%	0.50	I	0.66	4.63
RPHPE	85%	0.20	I	0.21	11.72
RPHPE	85%	0.35	I	0.60	7.25
RPHPE	85%	0.50	I	0.80	4.72
RPHPE	85%	0.20	I	0.64	11.46
RPHPE	85%	0.35	I	0.82	5.92
RPHPE	85%	0.50	I	0.91	3.53

Notes: This table shows the different distributional properties of premium rate estimates for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion policies, generated by a copula selected with the Akaike Information Criterion (AIC) versus a Normal copula. The measure Median Ratio is the median premium rate estimate generated by the Normal copula divided by the median premium rate estimate generated by the AIC selected copula. The Dispersion Ratio measures the relative dispersion between premium rates estimated from a Normal copula versus an AIC selected Copula. It is calculated by taking the difference between the lower 2.5 percentile and upper 97.5 percentile of premium rate estimates generated by the normal copula, dividing them by the median estimate and dividing this value by the same formula applied to the premium rates generated by the AIC copula. The set of copulas over which the AIC could select are the Normal (N), Student-t with degrees of freedom calibrated by the AIC (t), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected. The dispersion in prices is denoted Price σ .

almost 30% less for RP premium rates and almost 40% for RPHPE premium rates.

From these results, low dependence counties are receiving a significant premium rate discount as a function of the implicit structure the Normal copula imposes.⁹²

⁹²While this result appears in conflict with previous simulation results, which showed that as dependence weakened premium rates generated by alternative copulas converged, it must be noted that the lowest level of dependence examined in this experiment is a Kendall's τ of -.17. The level of dependence is more than three times larger than the -0.05 Kendall's τ used throughout the simulations to represent low dependence. Therefore, the premium rate convergence found in the

Table 23: Cass, IA AIC and Normal Copula Premium Rates

Policy	Coverage	Price σ	AIC Copula	Median Ratio	Dispersion Ratio
RP	50%	0.20	I	0.20	3.61
RP	50%	0.35	I	0.31	10.08
RP	50%	0.50	I	0.59	8.84
RP	50%	0.20	I	0.41	9.58
RP	50%	0.35	I	0.57	12.22
RP	50%	0.50	I	0.76	8.33
RP	65%	0.20	I	0.78	11.15
RP	65%	0.35	I	0.86	7.73
RP	65%	0.50	I	0.92	4.28
RPHPE	65%	0.20	I	0.12	6.26
RPHPE	65%	0.35	I	0.30	10.40
RPHPE	65%	0.50	I	0.59	8.82
RPHPE	85%	0.20	I	0.30	15.01
RPHPE	85%	0.35	I	0.55	12.99
RPHPE	85%	0.50	I	0.75	8.58
RPHPE	85%	0.20	I	0.66	17.61
RPHPE	85%	0.35	I	0.80	10.38
RPHPE	85%	0.50	I	0.88	5.80

Notes: This table shows the different distributional properties of premium rate estimates for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion policies, generated by a copula selected with the Akaike Information Criterion (AIC) versus a Normal copula. The measure Median Ratio is the median premium rate estimate generated by the Normal copula divided by the median premium rate estimate generated by the AIC selected copula. The Dispersion Ratio measures the relative dispersion between premium rates estimated from a Normal copula versus an AIC selected Copula. It is calculated by taking the difference between the lower 2.5 percentile and upper 97.5 percentile of premium rate estimates generated by the normal copula, dividing them by the median estimate and dividing this value by the same formula applied to the premium rates generated by the AIC copula. The set of copulas over which the AIC could select are the Normal (N), Student-t with degrees of freedom calibrated by the AIC (t), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected. The dispersion in prices is denoted Price σ .

Across the low dependence counties represented by Tables 22 through 24 the dispersion of the Normal copula premium rates versus the premium rates generated by the Independence copula is markedly greater. However, there is no instability surrounding the parameter estimate for the Independence copula: it is invariant.

simulation results may only begin to take effect for counties with essentially independent price-yield dependence.

Table 24: Lucas, IA AIC and Normal Copula Premium Rates

Policy	Coverage	Price σ	AIC Copula	Median Ratio	Dispersion Ratio
RP	50%	0.20	I	0.76	2.32
RP	50%	0.35	I	0.60	7.08
RP	50%	0.50	I	0.62	11.12
RP	50%	0.20	I	0.80	4.13
RP	50%	0.35	I	0.74	9.65
RP	50%	0.50	I	0.79	9.62
RP	65%	0.20	I	0.90	5.80
RP	65%	0.35	I	0.90	6.24
RP	65%	0.50	I	0.93	4.76
RPHPE	65%	0.20	I	0.60	5.45
RPHPE	65%	0.35	I	0.48	11.62
RPHPE	65%	0.50	I	0.56	14.45
RPHPE	85%	0.20	I	0.67	8.40
RPHPE	85%	0.35	I	0.63	14.90
RPHPE	85%	0.50	I	0.73	13.12
RPHPE	85%	0.20	I	0.81	10.77
RPHPE	85%	0.35	I	0.81	10.97
RPHPE	85%	0.50	I	0.87	8.42

Notes: This table shows the different distributional properties of premium rate estimates for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion policies, generated by a copula selected with the Akaike Information Criterion (AIC) versus a Normal copula. The measure Median Ratio is the median premium rate estimate generated by the Normal copula divided by the median premium rate estimate generated by the AIC selected copula. The Dispersion Ratio measures the relative dispersion between premium rates estimated from a Normal copula versus an AIC selected Copula. It is calculated by taking the difference between the lower 2.5 percentile and upper 97.5 percentile of premium rate estimates generated by the normal copula, dividing them by the median estimate and dividing this value by the same formula applied to the premium rates generated by the AIC copula. The set of copulas over which the AIC could select are the Normal (N), Student-t with degrees of freedom calibrated by the AIC (t), Frank (F), Gumbel 90° (g90), Gumbel 270° (g270), Clayton 90° (c90) and Clayton 270° (c270). A test for independence was also performed and if not rejected, then the Independence (I) copula is selected. The dispersion in prices is denoted Price σ .

While the premium rate generating process for both the AIC and Normal copulas is identified by the same estimate of Kendall's τ , it seems logical that Normal copula rate imprecision is magnified by additional instability in its parameter estimation.

Summary

The first result of this experiment showed that pure empirical premium rates will tend to converge, within counties, as coverage levels increase. This finding further reinforces the importance of the marginal distributions, not only for a copula based rating of revenue products, but for rating revenue products in general. It was again shown that higher price dispersion environments, will tend to produce higher copula premium rates than pure empirical premium rates. These results are important because the choice to rate multiple peril crop insurance (MPCI) revenue products with copulas, introduces a certain amount of variability in both premium rate dispersion and direction as a function of the overarching price dispersion environment

The second result showed that the AIC criterion clearly selects a best-fit copula for each copula and the selections are at least internally consistent. While this selection is limited by a predefined set of copulas, this set could be expanded and from a practical perspective it would not be difficult to perform a resampling procedure to establish the best fit copula for a county-crop. Moreover, the number of resampled iterations could be as low as 200, possibly lower, and a reasonable assessment of the best fitting copula could be attained using an AIC criterion. For that matter generating a distribution of premium rates conditional upon the AIC selected copula is not be computationally inconceivable. Generating a distribution of premium rates would allow the instability of the premium rate estimates to be

incorporated into premium rates, although the additional uncertainty regarding the accuracy of the AIC selection would also need to be incorporated.

At very least, the decision to impose a Normal copula on counties with weakly dependent price-yield characteristics, imposes more structure on the data than is statistically valid. If the results of the counties in this study can be extrapolated, then the additional structure of the Normal copula, allows weakly dependent counties to receive a large premium rate discount, relative to the essential independence between their county yields and national prices. On the other hand, if the AIC selection for more strongly dependent counties is true or at least a better model than the Normal copula, these counties tend to have higher premium rates except at low levels of price dispersion. The crux of incorporating an AIC copula selection method into current ratemaking practices is how much confidence can be placed in the AIC's identification. Incorporating the uncertainty surrounding the AIC selected copula—as a function of dependence—into the ratemaking process would be a necessary component of using the AIC to rate MPCIR revenue products.

CHAPTER TWELVE

CONCLUSION

Publicly provided and subsidized multiple peril crop insurance (MPCI) is an important risk management tool for the United States farm sector. Policies that protect a producer's revenue were introduced in 1996. With these policies came the need to not only model the distribution of producer yields, but the distribution of prices and the relationship between these distributions in order to develop actuarially fair premium rates. To this end, the USDA Risk Management Agency's current ratemaking strategies impose the equivalent of a Normal copula on independently simulated price and yield realizations to model county and crop specific price and yield relationships.

This study investigates the viability of using a copula based ratemaking strategy by evaluating the implications of copula modeling through six empirical and simulation experiments. The first concern of copula modeling is the selection of an appropriate copula and the statistical basis upon which a Normal copula could be selected was examined in a Monte-Carlo simulation that investigated the performance of available goodness-of-fit statistics at the number of observations and dependence that characterize data sets typically used to develop premium rates for multiple peril crop insurance (MPCI) revenue policies. An empirical experiment examined the potential for Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates generated by alternative, equally

viable, copulas tend to diverge for most counties in the major corn producing states of Illinois, Indiana and Iowa. Limits with respect to the amount of empirical data available for rate setting warranted a simulation experiment to investigate the same potential for premium rate divergence, but across a set of yield and dependence characteristic not empirically evident in corn. Next, the potential for premium rates to diverge within copulas was examined with a simulation experiment by generating distributions of “true” premium rates, obtained from precisely identified marginal distributions and copulas, as well as sample premium rates generated by alternative copulas calibrated with measures of price-yield dependence based on 50 paired observations. The design of this experiment allowed an assessment of the performance of an alternative copula selection strategy, the Akaike Information Criterion (AIC), to be assessed. Finally, the AIC was used to select a county specific copula for a subset of the counties in the first empirical experiment that represented a range in empirical yield and dependence characteristics. Premium rate distributions were generated from the AIC selected copula and compared to the distribution of premium rates generated from the Normal copula, which is currently used in the ratemaking procedure for revenue policies.

Copulas are used in the existing MPCIR ratemaking process to attempt to overcome data limitations and establish more actuarially accurate revenue product premium rates. However, the same limits on the number of observations available to rate setting that leads to their use render statistical identification of the appropriate copula infeasible. This inevitably results in imprecision in revenue premium rate

estimates and raises the issue of the viability of alternative copulas. Moreover, factors that affect the level of premium rates not only affect the dispersion between premium rates generated by different copulas, but the dispersion of premium rates generated within copulas. Premium rates would more accurately reflect the uncertainty of copula modeling, if they were adjusted by the potential for implicit premium rate imprecision; such an adjustment would increase premium rates for all revenue policies. In addition, the sensitivity of premium rate dispersion to changes in the *assumed* price and yield marginal distributions implies that copula based ratemaking depends critically on the accurate identification of the price and yield marginals.

The use of an AIC selection criterion could improve the ratemaking process by allowing the copula to vary by county and crop. This method of copula selection has distinct practical advantages over a goodness-of-fit based approach; however, it is subject to the same fundamental criticisms. Its implementation would not only require additional quantitative work, but qualitative assessments, as well as the compatibility of its results with theoretical insight. In fact, in counties where dependence between yields and prices is weak, premium rates are likely to be more accurately rated, if their price and yield realizations were treated as independent. The current use of a Normal copula cannot be supported and, according to AIC criterion, is sub-optimal. In light of the attractive features of an alternative non-copula based pure empirical ratemaking strategy and the implicit premium rate imprecision not currently incorporated into revenue policy premium rates, whether a

copula based ratemaking meets the overt policy objectives for MPCl warrants reconsideration.

REFERENCES CITED

- Ahsan, S. M., Ali, A. A., and Kurian, N. J. (1982). Toward a theory of agricultural insurance. *American Journal of Agricultural Economics*, 64(3), 510–529.
- Arrow, K. (1971). *Essays in the theory of risk bearing*. Markham Publishing Co.
- Atwood, J., Baquet, A., and Watts, M. (1996). *Income protection: technical report* (Tech. Rep.). Agricultural Division and Montana State University.
- Atwood, J., Shaik, S., and Watts, M. (2002). Can normality of yields be assumed for crop insurance? *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 50(2), 171–184.
- Atwood, J., Shaik, S., and Watts, M. (2003). Are crop yields normally distributed? a reexamination. *American Journal of Agricultural Economics*, 85(4), 888–901.
- Babcock, B. A., Hart, C. E., and Hayes, D. J. (2004). Actuarial fairness of crop insurance rates with constant rate relativities. *American Journal of Agricultural Economics*, 86(3), 563–575.
- Babcock, B. A., and Hennessy, D. A. (1996). Input demand under yield and revenue insurance. *American Journal of Agricultural Economics*, 78(2), 416–427.
- Beaudoin, D. (2007). *Estimation de la dépendance et choix de modèles pour des données bivariées sujettes à censure et troncature* (Unpublished doctoral dissertation). PhD thesis, Université Laval, Québec, Canada.
- Borch, K. (1960). The safety loading of reinsurance premiums. *Scandinavian Actuarial Journal*, 1960(3-4), 163–184.
- Botts, R. R., and Boles, J. N. (1958). Use of normal-curve theory in crop insurance ratemaking. *Journal of Farm Economics*, 40(3), 733–740.
- Brechmann, E. C., and Schepsmeier, U. (2013). Modeling dependence with c- and d-vine copulas: The R package CDVine. *Journal of Statistical Software*, 52(3), 1–27.
- Chambers, R. G. (1989). Insurability and moral hazard in agricultural insurance markets. *American Journal of Agricultural Economics*, 71(3), 604–616.
- Coble, K. H., and Barnett, B. J. (2013). Why do we subsidize crop insurance? *American Journal of Agricultural Economics*, 95(2), 498–504.
- Coble, K. H., Knight, T. O., Goodwin, B. K., Miller, M. F., Rejesus, R. M., and Duffield, G. (2010). A comprehensive review of the rma aph and combo rating

methodology, final report. *prepared by sumaria systems for the Risk Management Agency.*

- Coble, K. H., Knight, T. O., Pope, R. D., and Williams, J. R. (1996). Modeling farm-level crop insurance demand with panel data. *American Journal of Agricultural Economics*, 78(2), 439–447.
- Coble, K. H., Knight, T. O., Pope, R. D., and Williams, J. R. (1997). An expected-indemnity approach to the measurement of moral hazard in crop insurance. *American Journal of Agricultural Economics*, 79(1), 216–226.
- Edwards, W., and Plastina, A. (2014). *Trend-adjusted actual production history (aph)* (Tech. Rep.). Iowa State University Extension and Outreach.
- Egli, D. (2008). Comparison of corn and soybean yields in the united states: Historical trends and future prospects. *Agronomy journal*, 100(Supplement.3), S–79.
- Genest, C., Rémillard, B., and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and economics*, 44(2), 199–213.
- Glauber, J. W. (2004). Crop insurance reconsidered. *American Journal of Agricultural Economics*, 86(5), 1179–1195.
- Glauber, J. W. (2012). The growth of the federal crop insurance program, 1990–2011. *American Journal of Agricultural Economics*, 95(2), 482–488.
- Goodwin, B. K. (1993). An empirical analysis of the demand for multiple peril crop insurance. *American Journal of Agricultural Economics*, 75(2), 425–434.
- Goodwin, B. K. (1994). Premium rate determination in the federal crop insurance program: what do averages have to say about risk? *Journal of Agricultural and Resource Economics*, 382–395.
- Goodwin, B. K. (2001a). Problems with market insurance in agriculture. *American Journal of Agricultural Economics*, 83(3), 643–649.
- Goodwin, B. K. (2001b). U.s. farm safety nets and the 2000 agricultural risk protection act. *Canadian Journal of Agricultural Economics/Revue canadienne d’agroeconomie*, 49(4), 543–555.
- Goodwin, B. K. (2014). Agricultural policy analysis: The good, the bad, and the ugly. *American Journal of Agricultural Economics*.

- Goodwin, B. K., Harri, R., Rejesys, R. M., Coble, K. H., and Knight, T. O. (2014). *Actuarial review for price volatility factor methodology* (Tech. Rep.). Sumaria Systems, Inc.
- Goodwin, B. K., and Hungerford, A. (2014). Copula-based models of systemic risk in us agriculture: implications for crop insurance and reinsurance contracts. *American Journal of Agricultural Economics*.
- Goodwin, B. K., Roberts, M. C., and Coble, K. H. (2000). Measurement of price risk in revenue insurance: implications of distributional assumptions. *Journal of Agricultural and Resource Economics*, 195–214.
- Goodwin, B. K., and Smith, V. H. (1995). The economics of crop insurance and disaster aid. *AEI studies in agricultural policy (USA)*.
- Goodwin, B. K., and Smith, V. H. (2003). An ex post evaluation of the conservation reserve, federal crop insurance, and other government programs: program participation and soil erosion. *Journal of Agricultural and Resource Economics*, 201–216.
- Goodwin, B. K., and Smith, V. H. (2013). What harm is done by subsidizing crop insurance? *American Journal of Agricultural Economics*, 95(2), 489–497.
- Goodwin, B. K., Vandever, M. L., and Deal, J. L. (2004). An empirical analysis of acreage effects of participation in the federal crop insurance program. *American Journal of Agricultural Economics*, 86(4), 1058–1077.
- Halcrow, H. G. (1949). Actuarial structures for crop insurance. *Journal of Farm Economics*, 31(3), 418–443.
- Harri, A., Erdem, C., Coble, K. H., and Knight, T. O. (2009). Crop yield distributions: A reconciliation of previous research and statistical tests for normality. *Applied Economic Perspectives and Policy*, 31(1), 163–182.
- Hofert, M., Kojadinovic, I., Maechler, M., and Yan, J. (2014). copula: Multivariate dependence with copulas [Computer software manual]. (R package version 0.999-12.)
- Hojjati, B., and Bockstael, N. E. (1988). *Modeling the demand for crop insurance*. State University.
- Horowitz, J. K., and Lichtenberg, E. (1993). Insurance, moral hazard, and chemical use in agriculture. *American Journal of Agricultural Economics*, 75(4), 926–935.

- Hull, J. (2012). *Options, futures, and other derivatives* (8th ed.). Prentice Hall.
- Hungerford, A. E., and Goodwin, B. (2014). Big assumptions for small samples in crop insurance. *Agricultural Finance Review*, 74(4).
- Josephson, G. R., Lord, R. B., and Mitchell, C. W. (2000). Actuarial documentation of multiple peril crop insurance ratemaking procedures. *Report of Milliman and Robertson, Inc. prepared for the Risk Management Agency*.
- Just, R. E., and Weninger, Q. (1999). Are crop yields normally distributed? *American Journal of Agricultural Economics*, 81(2), 287–304.
- Knight, T. O., and Coble, K. H. (1997). Survey of us multiple peril crop insurance literature since 1980. *Review of Agricultural Economics*, 19(1), 128–156.
- Knight, T. O., and Coble, K. H. (1999). Actuarial effects of unit structure in the us actual production history crop insurance program. *Journal of Agricultural and Applied Economics*, 31(3), 519–536.
- Kojadinovic, I., Yan, J., Holmes, M., et al. (2011). Fast large-sample goodness-of-fit tests for copulas. *Statistica Sinica*, 21(2), 841.
- Kramer, R. A. (1983). Federal crop insurance 1938-1982. *Agricultural History*, 181–200.
- Lence, S. H., and Hayes, D. J. (2002). Us farm policy and the volatility of commodity prices and farm revenues. *American Journal of Agricultural Economics*, 84(2), 335–351.
- Miranda, M. J. (1991). Area-yield crop insurance reconsidered. *American Journal of Agricultural Economics*, 73(2), 233–242.
- Miranda, M. J., and Glauber, J. W. (1997). Systemic risk, reinsurance, and the failure of crop insurance markets. *American Journal of Agricultural Economics*, 79(1), 206–215.
- Mishra, A. K., Wesley Nimon, R., and El-Osta, H. S. (2005). Is moral hazard good for the environment? revenue insurance and chemical input use. *Journal of environmental management*, 74(1), 11–20.
- Nelson, R. B. (1999). *An introduction to copulas*. Springer.
- Nieuwoudt, W., and Bullock, J. (1985). The demand for crop insurance. In *1985 conference, august 26-september 4, 1985, Málaga, Spain*.

- Paulson, N. D., and Babcock, B. A. (2008). Get a grip: should area revenue coverage be offered through the farm bill or as a crop insurance program? *Journal of Agricultural and Resource Economics*, 137–153.
- Plastina, A., and Edwards, W. (2014). *Proven yields and insurance units for crop insurance*. Retrieved from <http://www.extension.iastate.edu/agdm/crops/pdf/a1-55.pdf> ([Online; accessed 24-September-2014])
- Quiggin, J. C., Karagiannis, G., and Stanton, J. (1993). Crop insurance and crop production: an empirical study of moral hazard and adverse selection. *Australian Journal of Agricultural and Resource Economics*, 37(2), 95–113.
- Ramaswami, B. (1993). Supply response to agricultural insurance: Risk reduction and moral hazard effects. *American Journal of Agricultural Economics*, 75(4), 914–925.
- Raviv, A. (1979). The design of an optimal insurance policy. *The American Economic Review*, 84–96.
- RMA. (2009). *Rma revenue rating an analysis of the combo rating method* (Tech. Rep.). Risk Management Agency.
- Roberts, M. J., Key, N., and O'Donoghue, E. (2006). Estimating the extent of moral hazard in crop insurance using administrative data. *Applied Economic Perspectives and Policy*, 28(3), 381–390.
- Rothschild, M., and Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4), pp. 629-649.
- Schmidt, T. (2007). Coping with copulas. *Copulas-From theory to application in finance*, 3–34.
- Schurle, B. (1996). The impact of size on yield variability and crop insurance premiums. *Review of Agricultural Economics*, 415–422.
- Serra, T., Goodwin, B. K., and Featherstone, A. M. (2003). Modeling changes in the us demand for crop insurance during the 1990s. *Agricultural Finance Review*, 63(2), 109–125.
- Shaik, S., and Atwood, J. A. (2003). Demand for optional units in crop insurance. In *2003 annual meeting, july 27-30, montreal, canada*.
- Sherrick, B. J., Zanini, F. C., Schnitkey, G. D., and Irwin, S. H. (2004). Crop

- insurance valuation under alternative yield distributions. *American Journal of Agricultural Economics*, 86(2), 406–419.
- Shields, D. A. (2012). Federal crop insurance: Background. In *Congressional research service. crs report for congress, december* (Vol. 6).
- Skees, J. R., and Reed, M. R. (1986). Rate making for farm-level crop insurance: Implications for adverse selection. *American Journal of Agricultural Economics*, 68(3), 653–659.
- Sklar, M. (1959). *Fonctions de répartition à n dimensions et leurs marges*. Publications de l'Institut de Statistique de Université de Paris 8.
- Smith, V. H., and Baquet, A. E. (1996). The demand for multiple peril crop insurance: evidence from montana wheat farms. *American journal of agricultural economics*, 78(1), 189–201.
- Smith, V. H., and Glauber, J. W. (2012). Agricultural insurance in developed countries: Where have we been and where are we going? *Applied Economic Perspectives and Policy*, pps029.
- Smith, V. H., and Goodwin, B. K. (1996). Crop insurance, moral hazard, and agricultural chemical use. *American Journal of Agricultural Economics*, 78(2), 428–438.
- Society, C. A. (1988). *Statement of principles regarding property and casualty insurance ratemaking* (Tech. Rep.). Author.
- Stiglitz, J. E. (1975). Incentives, risk , and information: Notes towards a theory of hierarchy. *The Bell Journal of Economics*, 6(2), 552–579.
- Swanson, E. R., and Nyankori, J. C. (1979). Influence of weather and technology on corn and soybean yield trends. *Agricultural Meteorology*, 20(4), 327–342.
- Tew, B. V., and Reid, D. W. (1988). Probability distributions of crop prices, yields, and gross revenue. *Northeastern Journal of Agricultural and Resource Economics*, 17(2).
- USDA. (1999). *yield98*. Retrieved from ftp://ftp.rma.usda.gov/pub/Miscellaneous_Files/yield98/
- USDA. (2014). *History of the crop insurance program*. Retrieved from <http://www.rma.usda.gov/aboutrma/what/history.html> ([Online; accessed 2-September-2014])

- Walters, C. G., Shumway, R. C., Chouinard, H. H., and Wandschneider, P. R. (2012). Crop insurance, land allocation, and the environment. *Journal of Agricultural and Resource Economics*, 37(2), 301.
- Wang, H. H., and Zhang, H. (2003). On the possibility of a private crop insurance market: A spatial statistics approach. *Journal of Risk and Insurance*, 70(1), 111–124.
- Woodard, J. D., Sherrick, B. J., and Schnitkey, G. D. (2011). Actuarial impacts of loss cost ratio ratemaking in us crop insurance programs. *Journal of Agricultural and Resource Economics*, 36(1), 211.
- Wright, B. D., and Hewitt, J. A. (1994). All-risk crop insurance: lessons from theory and experience. In *Economics of agricultural crop insurance: theory and evidence* (pp. 73–112). Kluwer Academic Publishers.

APPENDICES

APPENDIX A

TABLES

Table 25: Abbreviations

Abbreviation	Name
RMA	Risk Management Agency
MPCI	Multiple Peril Crop Insurance
APH	Actual Production History
LCR	Lost Cost Ratio
RP	Revenue Protection
RPHPE	Revenue Protection with Harvest Price Exclusion
<i>PP</i>	Projected Price
<i>HP</i>	Harvest Price
N	Normal or Gaussian copula
t4	Student-t copula with 4 degrees of freedom
F	Frank copula
c90	Clayton copula rotated 90°
c270	Clayton copula rotated 270°
g90	Gumbel copula rotated 90°
g270	Gumbel copula rotated 270°
<i>Pb</i>	Parametric Bootstrap Method
<i>Mult</i>	Fast Multiplier Method
i- τ	Inversion of Kendall's τ
i- ρ	Inversion of Spearman's ρ
<i>mpl</i>	Maximum Pseudo-Likelihood
empP	Pure Empirical Method
empC	Empirical copula
AIC	Akaike Information Criterion
LRM	Linear Regression Model

Table 26: Experimental Design for Goodness-of-Fit Studies

		Study	
	g2009	k2011	s2015
Copula	C,G,F,P,N, t_4,p_4	C,G,F,P,N, t_4	C,G,F,N, t_4
Sim.	Pb	Mult	Mult
Est.	i- τ	i- τ ,i- ρ , mpl	i- τ , mpl
Meth.	$S_n, T_n, S_n^K, S_n^B, S_n^C, A_n$	S_n	S_n
Rep.	10,000	10,000	10,000
n	{50, 150}	{75, 150, 300}	{50, 75, 150}
N	1,000	1,000	1,000
Dep.	{.25, .5, .75}	{.25, .5, .75}	{.01, .25, .50}

	Variable		Copula
Copula	True and H_0 copulas	C**†	Clayton
Sim.	Generate gof stat.	G**†	Gumbel
Est.	Estimate σ	F**†	Frank
Meth.	Test statistic	P**†	Plackett
Rep.	Simulation size	N**†	Normal
n	Sample size	p4	Pearson 4df
N	Primary bootstraps	t4**†	Student-t 4df
τ	Dependence		

	Method		Simulation
S_n^{**}	Cramer-von Mises	Pb*	Parametric Bootstrap
T_n	Kolmogorv-Smirnov	Mult*	Fast Multiplier
S_n^K	Kendall's S_n		
T_n^K	Kendall's T_n		
S_n^{B*}	S_n Variate B		
S_n^{C*}	S_n Variate C		
A_n^*	χ^2 Anderson-Darling		

	Estimation Method		Study
i- τ^*	Inv. of Kendall's τ	b2007	Beaudoin (2007)
i- ρ^*	Inv. of Spearman's ρ	g2009	Genest et al. (2009)
mpl^*	Max-Pseudo Likelihood	k2011	Kojadinovic et al. (2011)
		s2015	Present Study

*Implemented in R			
**Implemented for multiplier simulation method			
†Rotated representation in R			

Table 27: Percent Rejection of Null Hypothesis with $n = 75$

Study	genCop	Sim	Est	$\tau = .01$					$\tau = .25$					$\tau = .50$				
				C	F	G	N	t4	C	F	G	N	t4	C	F	G	N	t4
s2015	C	Mult	<i>mpl</i>	13	16	16	15	15	9	60	75	50	47	9	94	99	87	85
			i- τ	8	8	9	8	13	8	41	68	39	44	8	78	97	78	82
s2015	F	Mult	<i>mpl</i>	9	9	10	9	15	23	8	28	11	17	60	9	53	23	30
			i- τ	0	0	0	0	0	3	1	4	1	1	46	4	37	10	17
s2015	G	Mult	<i>mpl</i>	9	10	9	10	11	52	23	5	14	11	93	47	5	20	15
			i- τ	7	7	7	7	11	41	10	8	8	11	85	14	7	10	13
s2015	N	Mult	<i>mpl</i>	9	10	10	10	17	13	11	16	6	11	45	22	27	5	8
			i- τ	8	8	8	8	13	19	9	24	8	13	48	11	34	7	12
s2015	t4	Mult	<i>mpl</i>	8	8	8	8	9	17	16	19	12	7	46	41	31	13	7
			i- τ	7	8	9	8	9	18	9	21	8	9	46	12	31	7	9

These results present the percent rejection of the H_0 hypothesis that the data come from the hypothesized copula for goodness-of-fit tests at the 5% significance level from the simulations conducted at 75 observations ($n = 75$) for the present study (s2015). The “true” copulas (*genCop*) and hypothesized copulas are the Clayton (C), Frank (F), Gumbel (G), Normal (N) and Student-t with 4 degrees of freedom (t4) and values are in bold font when the copula generating the data is the same as the hypothesized copula. The simulation methods (Sim), for generating the test statistics and p-values are the fast multiplier (*Mult*) and parametric bootstrap (*Pb*) methods. The estimation methods (*Est*) generating the hypothesized copula θ parameter are maximum-pseudo likelihood (*mpl*) and the inversion of Kendall’s τ (i- τ).

Table 28: Percent Rejection of Null Hypothesis with $n = 150$

Study	genCop	Sim	Est	$\tau = .01$					$\tau = .25$					$\tau = .50$				
				C	F	G	N	t4	C	F	G	N	t4	C	F	G	N	t4
k2011	C	Mult	<i>mpl</i>						7	80	91	66	58	7	100	100	98	97
s2015				9	11	11	10	9	7	79	91	65	58	7	100	100	98	97
g2009		Pb	<i>i-τ</i>						5	56	86	50	57	5	96	100	94	95
k2011		Mult	<i>i-τ</i>						6	57	85	52	55	5	95	100	94	95
s2015				5	6	6	6	10	6	56	86	52	54	5	95	100	94	95
k2011	F	Mult	<i>mpl</i>						45	5	34	9	17	91	5	70	26	41
s2015				5	6	6	6	11	44	5	33	9	17	92	5	70	26	40
g2009		Pb	<i>i-τ</i>						40	5	33	8	19	89	5	63	20	35
k2011		Mult	<i>i-τ</i>						46	5	34	9	17	92	5	69	24	42
s2015				0	0	0	0	0	13	0	5	0	1	89	2	53	14	27
k2011	G	Mult	<i>mpl</i>						83	40	4	21	15	100	80	4	33	24
s2015				7	7	6	6	8	82	40	4	21	14	100	79	4	32	23
g2009		Pb	<i>i-τ</i>						72	15	5	10	14	100	40	5	18	22
k2011		Mult	<i>i-τ</i>						76	16	5	12	14	100	38	5	19	24
s2015				7	7	7	7	12	76	16	5	12	14	100	37	4	19	22
k2011	N	Mult	<i>mpl</i>						31	14	19	5	9	83	44	37	4	6
s2015				6	6	6	6	12	30	15	19	5	9	82	43	36	4	7
g2009		Pb	<i>i-τ</i>						32	8	24	5	11	80	20	38	5	8
k2011		Mult	<i>i-τ</i>						38	7	26	5	9	85	16	42	5	8
s2015				5	6	5	5	11	38	7	25	5	10	85	15	42	4	8
k2011	t4	Mult	<i>mpl</i>						39	28	23	15	5	84	72	42	17	4
s2015				6	7	7	7	5	38	27	24	14	4	83	71	41	17	4
g2009		Pb	<i>i-τ</i>						32	8	24	5	5	77	27	34	5	5
k2011		Mult	<i>i-τ</i>						37	10	23	6	5	84	24	38	5	5
s2015				6	7	7	7	6	37	10	23	6	6	84	23	38	5	5

These results present the percent rejection of the H_0 hypothesis that the data come from the hypothesized copula for goodness-of-fit tests at the 5% significance level conducted at 150 observations ($n = 150$) for the study performed by Kojadinovic et al. (2011) (k2011), Genest et al. (2009) (g2009) and the present study (s2015). The “true” copulas (*genCop*) and hypothesized copulas are the Clayton (C), Frank (F), Gumbel (G), Normal (N) and Student-t with 4 degrees of freedom (t4) and results are in bold font when the copula generating the data is the same as the hypothesized copula. The simulation methods (Sim), for generating test statistics and p-values are the fast multiplier (*Mult*) and parametric bootstrap (*Pb*) methods. The estimation methods (*Est*) generating the hypothesized copula θ parameter are maximum-pseudo likelihood (*mpl*) and the inversion of Kendall’s τ (*i- τ*).

Table 29: Copula Versus Pure Empirical RP Premium Rates

Price Dispersion	Coverage Level	Copula Methods	Pure Empirical Method	Method Difference
0.20	%50	0.37	0.61	-0.24
0.20	%55	0.52	0.82	-0.29
0.20	%60	0.74	1.10	-0.36
0.20	%65	1.06	1.49	-0.43
0.20	%70	1.54	2.04	-0.51
0.20	%75	2.24	2.84	-0.61
0.20	%80	3.24	4.00	-0.76
0.20	%85	4.61	5.59	-0.98
0.35	%50	0.56	0.61	-0.05
0.35	%55	0.86	0.82	0.04
0.35	%60	1.32	1.10	0.22
0.35	%65	1.99	1.49	0.50
0.35	%70	2.93	2.04	0.89
0.35	%75	4.18	2.84	1.34
0.35	%80	5.78	4.00	1.78
0.35	%85	7.72	5.59	2.13
0.50	%50	1.39	0.61	0.78
0.50	%55	2.14	0.82	1.32
0.50	%60	3.13	1.10	2.03
0.50	%65	4.37	1.49	2.88
0.50	%70	5.88	2.04	3.83
0.50	%75	7.64	2.84	4.80
0.50	%80	9.67	4.00	5.67
0.50	%85	11.93	5.59	6.35

Notes: All values are scaled by 100 and can be interpreted in terms of percentages. The table presents the average copula generated Revenue Protection (RP) premium rate for most Illinois, Indiana and Iowa corn producing counties, as well as the premium rate generated by the pure empirical method at each unique combination of price dispersion and coverage level. The difference in these values is presented in the final column, which was calculated by subtracting the pure empirical premium rate from the average copula premium rate.

Table 30: Copula Versus Pure Empirical RPHPE Premium Rates

Price Dispersion	Coverage Level	Copula Methods	Pure Empirical Method	Method Difference
0.20	%50	0.30	0.43	-0.13
0.20	%55	0.41	0.57	-0.17
0.20	%60	0.57	0.77	-0.20
0.20	%65	0.82	1.06	-0.24
0.20	%70	1.19	1.47	-0.28
0.20	%75	1.73	2.09	-0.36
0.20	%80	2.52	3.03	-0.51
0.20	%85	3.60	4.35	-0.75
0.35	%50	0.46	0.43	0.03
0.35	%55	0.71	0.58	0.14
0.35	%60	1.10	0.77	0.33
0.35	%65	1.68	1.06	0.63
0.35	%70	2.49	1.47	1.02
0.35	%75	3.55	2.09	1.46
0.35	%80	4.88	3.03	1.85
0.35	%85	6.46	4.35	2.11
0.50	%50	1.28	0.43	0.84
0.50	%55	1.97	0.58	1.40
0.50	%60	2.89	0.77	2.12
0.50	%65	4.03	1.06	2.98
0.50	%70	5.39	1.47	3.92
0.50	%75	6.95	2.09	4.86
0.50	%80	8.69	3.03	5.66
0.50	%85	10.57	4.35	6.23

Notes: All values are scaled by 100 and can be interpreted in terms of percentages. The table presents the average copula generated Revenue Protection with Harvest Price Exclusion (RPHPE) premium rate for most Illinois, Indiana and Iowa corn producing counties, as well as the premium rate generated by the pure empirical method at each unique combination of price dispersion and coverage level. The difference in these values is presented in the final column, which was calculated by subtracting the pure empirical premium rate from the average copula premium rate.

Table 31: Characteristics of Selected Counties

			Yield Variation (CV)		
			[0.13,0.17)	[0.17,0.21)	[0.21,0.36]
			(Yield CV, τ)	(Yield CV, τ)	(Yield CV, τ)
Kendall's τ	[-0.48,-0.31)	Counties in Group	11	35	36
		Group Attributes	(.15,-.34)	(.18,-.35)	(.23,-.37)
		County, State	Macon, IL*	McDonough,IL	Hancock, IL
		County Identifier	17113	17109	17067
		County Attributes	(.15,-.34)	(.18,-.37)	(.23,-.37)
	[-0.31,-0.23)	Counties in Group	20	31	32
		Group Attributes	(.15,-.28)	(.19,-.28)	(.23,-.28)
		County, State	McLean, IL*	Tazwell,IL *	Logan, IL *
		County Identifier	17115	17179	17107
		County Attributes	(.16,-.28)	(.19,-.30)	(.22,-.30)
	[-0.23,-0.04]	Counties in Group	51	16	13
		Group Attributes	(.15,-.17)	(.18,-.22)	(.31,-.21)
		County, State	Delaware, IA	Cass, IA	Lucas, IA
		County Identifier	19055	19029	19177
		County Attributes	(.15,-.17)	(.18,-.24)	(.31,-.21)

Notes: This table presents the attributes of the nine groups, which represent the range in empirically observed variation in corn yield variation and price-yield dependence and the counties selected to represent these groups. Yield variation is measured by the coefficient of variation (CV) *Indicates counties studied by Goodwin (2014).

Table 32: Number of Observations for Empirical Statistics

Coverage	Price σ	Macon	McDonough	Hancock	Mclean	Tazwell	Logan	Delaware	Cass	Lucas
50%	0.20	912	902	938	937	923	940	855	529	591
50%	0.20	891	903	943	917	931	935	868	556	590
50%	0.20	900	908	943	941	930	946	853	537	584
65%	0.35	914	902	930	929	924	949	848	541	593
65%	0.35	911	919	938	921	925	937	862	541	581
65%	0.35	898	903	934	932	922	935	857	507	582
85%	0.50	899	903	946	927	925	935	854	550	567
85%	0.50	911	887	930	937	935	940	847	557	595
85%	0.50	912	879	926	934	920	948	865	572	596

Notes: This table presents the total number of observations underlying analysis of the distributional properties of the revenue insurance policy premium rates. The dispersion in prices is denoted Price σ .

APPENDIX B

FIGURES

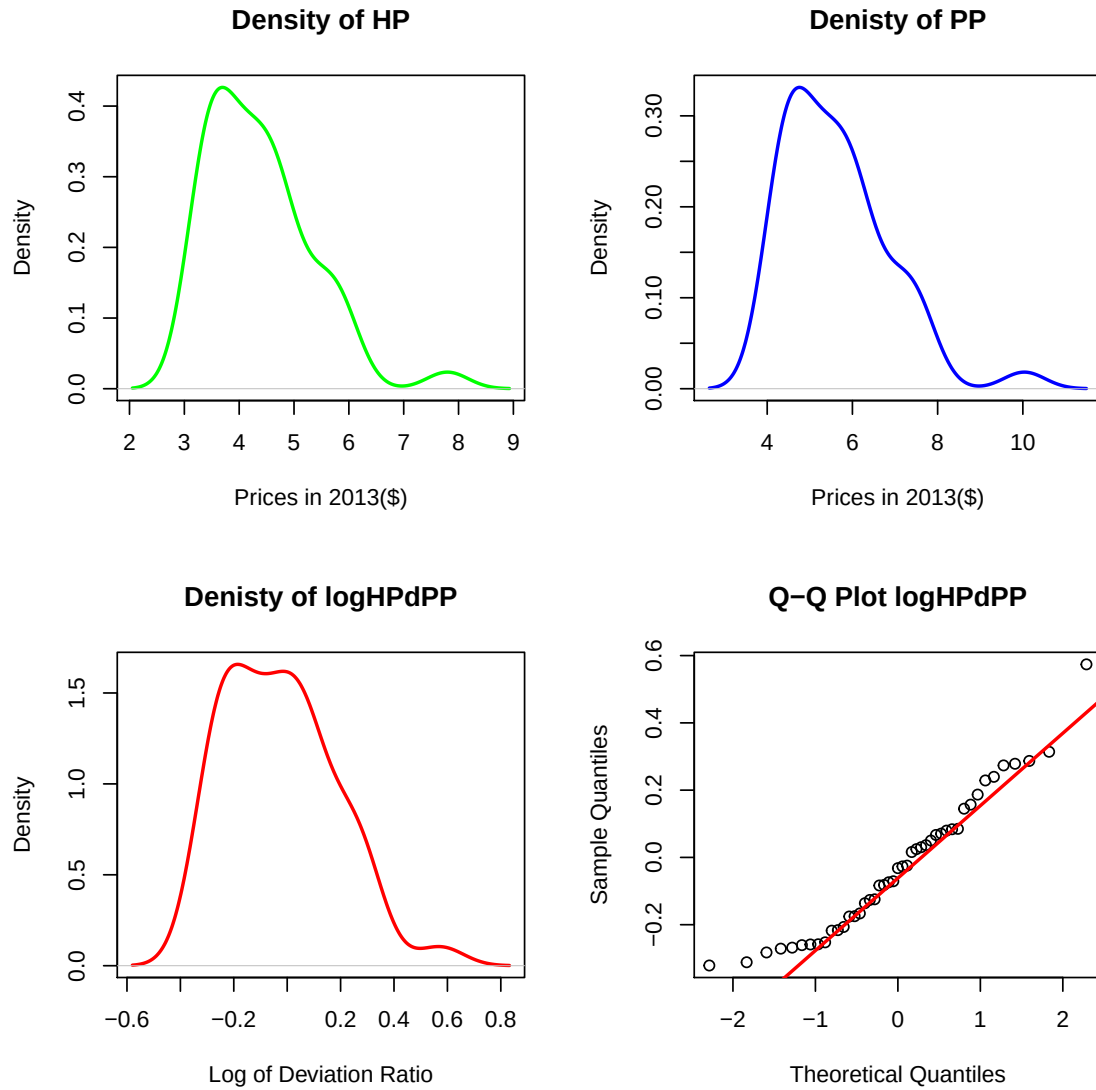


Figure 33: Density and Q-Q plot for Corn Futures Price Data

Notes: This figure shows the distributions of harvest prices (HP), projected prices (PP) and the transformed deviations between HP and PP (logHPdPP). A plot of the logHPdPP quantiles are also presented in a Q-Q plot.

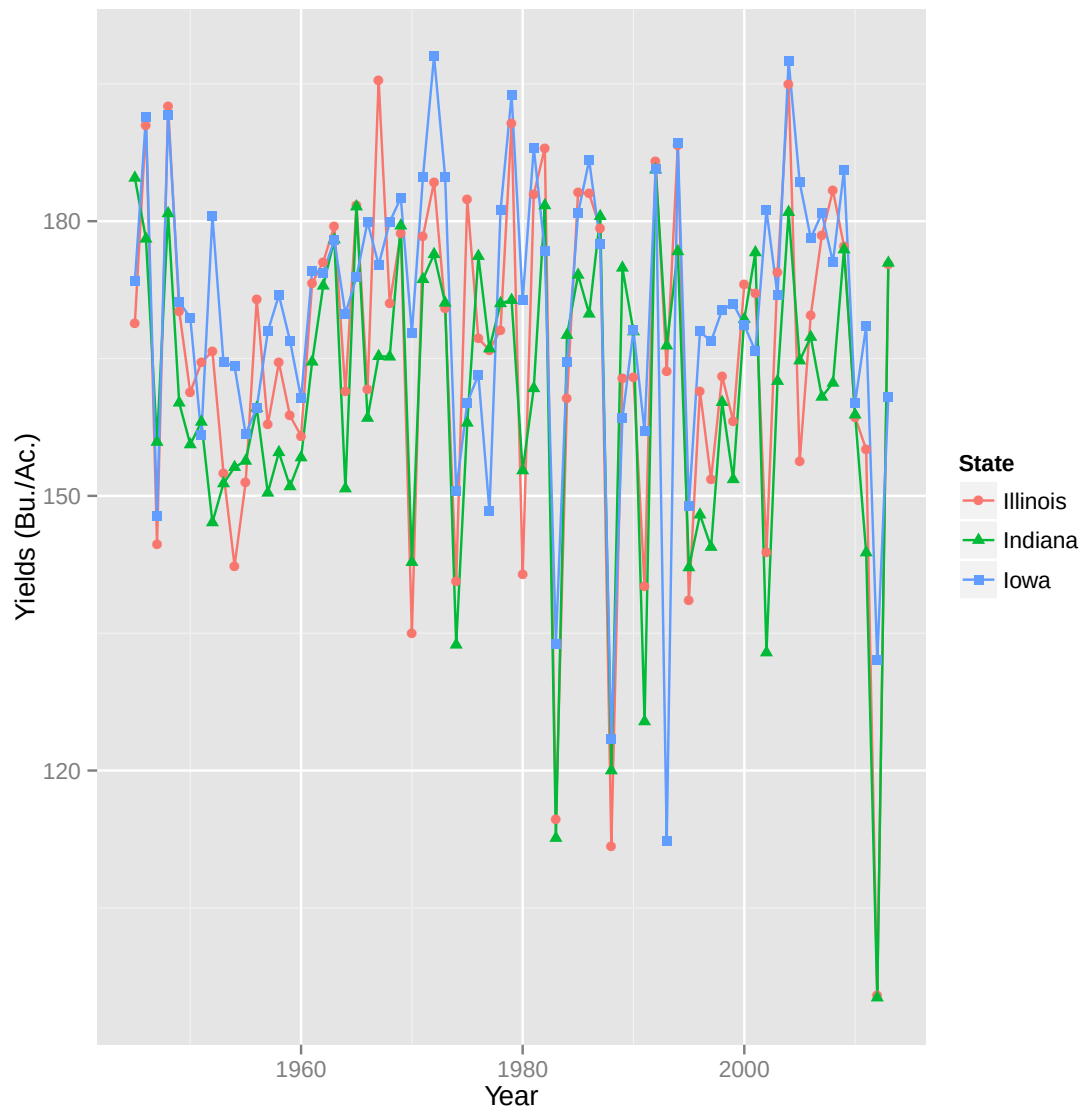


Figure 34: State Level Detrended Corn Yields from 1945-2013

APPENDIX C

TAXONOMY OF CROP INSURANCE POLICIES

The taxonomic crop insurance family tree, bifurcates first between “named-peril,” also called “specific-peril” insurance, and “multiple-peril” or “all-risk” insurance—MPCI. Private companies have successfully offered named-peril insurance for over a hundred years (Smith and Glauber, 2012). As a result there has been no incentive for governments to offer a similar product. However, MPCI has never been successfully offered and is instead currently funded by the federal government (Smith and Glauber, 2012).

Multiple-peril products are divided into two broad categories: individual protection, which is primarily administered under the Common Crop Insurance Policy (CCIP) and while area-based products, which reside under the umbrella of Area Risk Protection Insurance (ARPI) (Shields, 2012). Individual based coverage policies base premium rates and indemnification on a farm’s actual production history (*APH*). Area based policies set premium rates and pay indemnities based on an index related to for the individual producer’s yield or revenue outcome, usually county level yields or revenue per acre.

Both area and *APH* policies include yield and revenue products. Yield products pay an indemnity based solely on yield short falls, while revenue products protect the insured against revenue declines caused by either a yield or price short fall or both. CCIP yield and revenue products are both based on *APH*. There is both a specific yield product, called “APH,” and a standardized protocol and database for establishing a producer’s actual production history using the farm’s historical yield records and, where necessary, supplementing them with county yield

information. Revenue and yield products indemnify a producer based on a value of their crop, but there are two distinct possibilities for how this value might be established depending on whether or not there is a viable futures market for the crop. If an active futures market exists for the crop being insured, then the RMA has developed a protocol called the Commodity Exchange Price Provisions (CEPP) that establishes both the projected price (*PP*) and harvest price (*HP*); however, if there is no futures market for the crop, then only a yield insurance product is offered and the RMA establishes the price at which yields are valued (Coble et al., 2010). CEPP revenue products have the option of the Harvest Price Exclusion (HPE), which allows the producer to elect to receive indemnification based solely on the project price in the form of Revenue Protection with Harvest Price Exclusion (RPHPE). Without this exclusion the producer elects Revenue Protection (RP) and may be indemnified based upon the higher of the *PP* or *HP*.

All ARPI policies are ones for which prices are set by CEPP, this allows for consistency between ARPI and most CCIP products (Shields, 2012). Some CCIP products allow a producer to insure the revenue of their whole farm and not just the revenue from a specific crop. The most basic level of coverage is a yield product that offers the minimum catastrophic coverage (CAT). Any level of either yield or revenue coverage that exceeds CAT, is called “buy-up” coverage (Shields, 2012).

APPENDIX D

CURRENT RATEMAKING PRACTICES

The consolidation of many different policies into the “COMBO” plan greatly simplified the ratemaking process for multiple peril crop insurance (MPCI). However, COMBO itself is rather complicated. The most recent comprehensive review of rate setting in the COMBO plan was completed by Coble et al. (2010).⁹³ COMBO provides a consistent framework for rating all individual Actual Production History (APH) yield or revenue products and only pertains to crops for which there is a viable futures market. The Risk Management Agency’s (RMA) ratemaking process consists of three basic steps: 1) adjusting previous loss and exposure experience, 2) the derivation of unloaded county base premium rates, and 3) the reloading of these premium rates. As this study deals with revenue rates, which are added onto base APH rates through the reloading process, more attention will be given to this aspect of the process and the specifics of how the revenue loads are devised.

The popularity of revenue programs means that much of the RMA’s historical loss experience is for these products. However, the RMA elects to convert this information into equivalent APH values. The process involves the recalculation of revenue based liability to an equivalent APH liability. In addition to converting historical liability and indemnity information to APH, the RMA also chooses to convert historical information at all coverage levels to the 65% level.

Converting all historical information to the 65% coverage level increases the available data upon which premium rates are based. The adjustment to liability is a

⁹³Much of this section is based on Coble et al. (2010) or an earlier review by Josephson et al. (2000).

straight forward multiplication of the historical cumulative liability at a given coverage level by the ratio of 65% coverage level divided by the given coverage level. The adjustment to indemnity is not so straightforward because for coverage levels below 65% an indemnity was not paid. An interpolation procedure is implemented to generate this information; the details of this procedure as well as the adjustment methods in general are provided by Coble et al. (2010). The adjustment to a common coverage level and policy is an attempt to increase credibility by increasing sample size.

Even when aggregating at the county level, MPCCI ratemaking is based on limited historical loss experience for some crops and regions.⁹⁴ The RMA uses a step-wise processes to evaluate the credibility of the target county's experience and supplements it with loss experiences from surrounding counties using the Bühlmann method. This "credibility weighting" process augments the target county's experience, if the target county has either limited or highly variable historical experience. The resulting lost cost ratio (*LCR*) is unloaded because it is calculated with various caps and without any experience from the lowest coverage level. Such loads, along with other miscellaneous loads, are then reintroduced. At this point the county unloaded APH base premium rate is individualized to reflect heterogeneity amongst individual producers': 1) crops, 2) practices,⁹⁵ 3) insurable units⁹⁶ and

⁹⁴Illinois, Indiana and Iowa corn has a relatively robust loss experience, but when compared to more common forms of insurance it is not a large amount of data.

⁹⁵Practices can vary by such difference as whether a farmer irrigates or their use of chemicals.

⁹⁶A farmer can elect to insure his business by electing a policy that insures revenue from an aggregation of fields producing the same crop, or the farmer can purchase separate insurance policies for individual fields.

4) coverage levels.

A separate procedure is used to derive both Revenue Protection with Harvest Price Exclusion (RPHPE) and Revenue Protection (RP) premium rates. There are four basic steps in rating revenue products: 1) calculating the price-yield correlations, 2) calibrating the price and yield distributions, 3) generating the correlated draws from the calibrated distributions and 4) simulating losses and calculating revenue premium rates.⁹⁷

Price and yield correlations are determined from historical futures price data⁹⁸ and National Agricultural Statistics Service (NASS) county yield data. Prices are transformed into deviation by calculating the percent deviation of each year's harvest price from the projected price. Yield data are linearly detrended and transformed into deviations from the detrended mean. Correlations between yields and prices are calculated over data since 1990 because of a 1985 policy change in federal surplus storage policy that made prices more sensitive to yields (Coble et al., 2010). Ultimately, this correlation measure of dependence is used to impose county level dependence on prices and yields through the Iman-Conover method.

Price and yield realizations come from assumed distributions. Therefore, the parameters identifying these distributions must be estimated. Prices are assumed to be log-normally distributed, which is supported by the models used actively by many traders (Coble et al., 2010). The price distribution is identified by the

⁹⁷The following exposition relies extensively on the RMA (2009) and Goodwin, Harri, Rejesys, Coble, and Knight (2014) reviews of the revenue rating process.

⁹⁸Corn uses the Chicago Board of Trade (CBOT) exchange.

projected price (PP) and a measure of dispersion, the implied volatility ($ivol$). The properties of the log-normal price distribution are:

$$\tilde{p} = e^{\sigma_p \cdot p_i + \mu_p} \quad (23)$$

where:

$$\begin{aligned} \sigma_p &= \ln((ivol)^2) \\ \mu_p &= \ln(PP) - \frac{1}{2} \cdot \sigma_p \end{aligned} \quad (24)$$

Yields are assumed to be distributed normally and censored at zero. The justification for this assumption is that it allows for the yield μ_y and σ_y to be more easily calibrated than alternative distributions (Coble et al., 2010). The general procedure to calibrate the yield parameters is to: 1) normalize the expected yield to 100, 2) identify the target APH rate, 3) find μ and σ such that the following equations hold:

$$\begin{aligned} 100 &= \frac{1}{5000} \sum_{i=1}^{5000} \max(y_i, 0) \\ Target\ Rate &= \frac{\frac{1}{5000} \sum_{i=1}^{5000} \max(0, 65 - \max(y_i, 0))}{65} \\ y_i &= z_i \cdot \sigma + \mu; z_i \sim \phi(0, 1) \end{aligned} \quad (25)$$

This procedure is carried out for *Target Rate* ranging from .001 to .50 by intervals of .0001, but always at the 65% coverage level. Finally, these parameters are adjusted

for actual production history yields not equal to 100, where:

$$\begin{aligned} \text{adjusted } \mu &= \tilde{\mu}_y = \frac{APH \cdot \mu}{100} \\ \text{adjusted } \sigma &= \tilde{\sigma}_y = \frac{\sigma}{\mu} \cdot \tilde{\mu}_y \end{aligned} \tag{26}$$

Once marginal price and yield parameters are identified, the price and yield realization can be simulated from the parametric distributions. However, these simulated marginal realizations are independent and the appropriate county level dependence must be imposed upon them.

The log-normality assumption for prices allows standard normal deviates to be log-normally transformed in order to generate simulate price realizations.⁹⁹ Standard normal deviates are generated out of a normal distribution: $z_i \sim \phi(0, 1)$. These draws are then transformed to a log-normal distribution where:¹⁰⁰

$$\ln(p_i) = \mu_p + z_i \cdot \sigma_p \tag{27}$$

Where:

$$\begin{aligned} \mu_p &= \ln(PP) - \frac{1}{2}\sigma_p^2 \\ \sigma_p &= ivol \end{aligned} \tag{28}$$

⁹⁹The following exposition the RMA's process for generating simulated price realization relied on Goodwin and Hungerford (2014).

¹⁰⁰As Goodwin and Hungerford (2014) states, the formulas provided by RMA (2009) are incorrect: the RMA document states that $\sigma_p = \ln(ivol + 1)$ and $\mu_p = \ln(PP) + \frac{1}{2}\sigma_p^2$. Only the correct formulations are detailed to follow.

The simulated harvest price (HP_i) realizations are the transformation of the log-prices:

$$HP_i = e^{\ln(p_i)} \quad (29)$$

The distribution of HP is log-normal.

The simulated yield realizations (y_i) are then generated out of standard normal distribution with the previously calibrated parameters and censored at zero:

$$y_i = \text{Max}(0, \phi(\tilde{\mu}_y, \tilde{\sigma}_y)) \quad (30)$$

Five hundred draws are calculated from the inverse normal of Babcocks Nearly Uniform sequence (Coble et al., 2010). The Iman-Conover procedure is used to impose the appropriate county level correlation of the independent simulated marginal realizations. This procedure generates a equivalently dimensioned reference matrix of uniformly distributed marginals with the desired correlation structure. The simulated price and yield realizations are then rank order sorted against the reference matrix to achieve the desired dependence while retaining the unique properties of the underlying marginal distributions. As Hungerford and Goodwin (2014) note, when a normal score function is used, the Iman-Conover method is analogous to the use of a Gaussian or Normal copula.

Given the calibrated marginal parameters and price-yield draws: *Yield*

premium rates, *RPHPE* premium rates, and *RP* premium rates are calculated:¹⁰¹

$$\begin{aligned}
 APH &= \sum_{i=1}^{500} \frac{Max(0, cvg*Y - Max(0, y_i))}{500*Y*cvg} \\
 RPHPE &= \sum_{i=1}^{500} \frac{Max(0, cvg*Y*P - Max(0, y_i))*Min(2*P, H_i)}{500*Y*cvg*P} \\
 RP &= \sum_{i=1}^{500} \frac{Max(0, cvg*Y*Min(2*P, Max(P, H_i)) - Max(0, y_i*Min(2*P, H_i)))}{500*Y*cvg*P}
 \end{aligned} \tag{31}$$

where Y is the producer's $\mathbb{E}(y_i)$ or actual production history, P is the projected price and H_i is a simulated harvest price realizations. The minimum functions serve to protect against price movements that are two times greater than the projected price, placing an upper bound on indemnity.

The revenue premium rate loads are then calculated based upon the difference in revenue premium rate from the *APH* premium rate:

$$\begin{aligned}
 RP_{load} &= RP - APH \\
 PP_{load} &= PP - APH
 \end{aligned} \tag{32}$$

Therefore, the final revenue rates are the individuals base APH rate plus the revenue load:

$$\begin{aligned}
 RP_{final} &= APH_{rate} + RP_{load} \\
 PP_{final} &= APH_{rate} + PP_{load}
 \end{aligned} \tag{33}$$

¹⁰¹The information regarding the RMA's revenue rating practice comes from RMA (2009) , Goodwin et al. (2014) and Coble et al. (2010).

APPENDIX E

GUMBEL AND CLAYTON COPULA TAIL DEPENDENCE

The tail dependence for the Clayton and Gumbel copulas, do not have a strict relationship across their range of potential $\theta(\tau)$ calibration parameters. Figure 35 shows that for dependence lower than $\tau \approx .2$, the Gumbel copula imposes stricter tail dependence than the Clayton. 75% of the counties in Illinois, Indiana and Iowa have a dependence stronger in absolute value than $\tau \approx .22$. Thus, for the majority of the counties the Clayton copula will induce higher tail dependence, but for 25% of the counties (i.e. the weak dependence counties), the Gumbel copula will induce higher tail dependence relative to the Clayton. Thus, it is hypothesized that rates for weaker dependent counties will tend to be lower out of Gumbel rotated 90° than Clayton rotated 270° copulas; conversely, rates for stronger dependence counties will be lower out of the Clayton rotated 270° than Gumbel rotated 90° copulas.

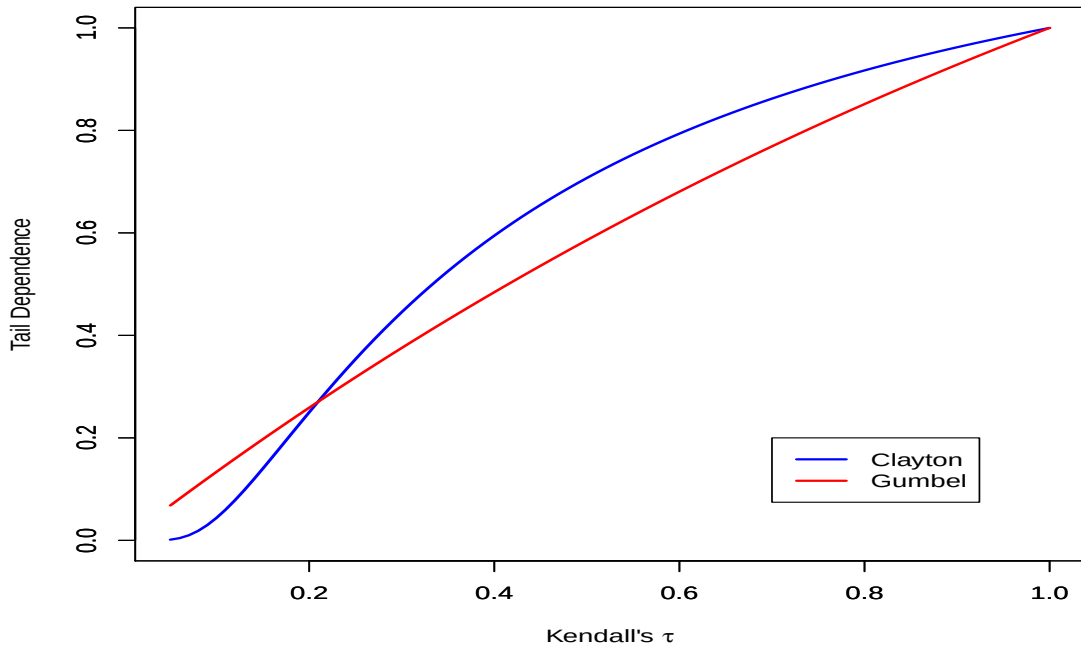


Figure 35: Tail Dependence of Clayton and Gumbel Copulas

APPENDIX F

ITÔ'S PROCESS FOR PRICES

A Markov process is one for which the distribution of a random variable in the future is not contingent on the path of the variable in the past.¹⁰² This aligns with the weak form of market efficiency, wherein the current price embodies all relevant past information. It can be show that the probability of change for a price for a short period of time is:

$$\Delta P = \phi(0, \Delta t) \quad (34)$$

If the mean change is equal to zero, then the process is known as a Wiener process and describes Brownian motion. The two properties of a general Wiener process is that the change in the underlying variable—price—must follow: $\Delta z = \epsilon \sqrt{\Delta t}$, where $\epsilon \sim \phi(0, 1)$ and the values of any Δz are independent between Δt . The generalized Wiener process incorporates mean drift (a) and variance rates (b):

$$dx = a \cdot dt + b \cdot dz \quad (35)$$

As Hull (2012) notes it would not be appropriate to model a price trajectory as a generalized Wiener process because the standard deviation in price should be proportional to the level value of that price. Thus, the following equivalent continuous models are thought to accurately reflect price movements, conditional on

¹⁰²This section calls extensively from Hull (2012).

the assumptions:

$$\begin{aligned} dSP &= \mu P dt + \sigma P dz \\ \frac{dP}{P} &= \mu dt + \sigma dz \end{aligned} \tag{36}$$

where μ is the expected rate of return, σ is the volatility in the stock (*ivol*) and P is the price at time t . This model is called geometric Brownian motion and the following discrete representation can be used to simulate any number of potential price trajectories across a specified time duration at discrete time intervals:

$$\Delta P = \mu P \Delta t + \sigma P \epsilon \sqrt{\Delta t} \tag{37}$$

where ΔP is the change in the price P over a small time interval Δt and $\epsilon \sim \phi(0, 1)$.

With this process, n number of unique price trajectories can be simulated for a specified length of time T at $T\Delta t$ number of discrete time intervals through Monte Carlo simulation.

A more elegant simulation can be achieved through Itô's calculus and the lemma, which allows unique realizations to be “jumped” to and is also one avenue for showing the log-normality of prices.

Itô's Lemma says that if a random variable follows an Itô process:

$$dx = a(x, t)dt + b(x, t)dz \tag{38}$$

where dz is a Wiener process, then Itô's Lemma shows that there is a function

$G(x, t)$ that follows the process:

$$dG = \left(\frac{\partial G}{\partial x} a + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} b^2 \right) dt + \frac{\partial G}{\partial x} b dz \quad (39)$$

If equation 36, is indeed a valid model for the motion of prices, then Itô's Lemma suggest that prices can be modeled:

$$dG = \left(\frac{\partial G}{\partial P} \mu P + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial P^2} \sigma^2 P^2 \right) dt + \frac{\partial G}{\partial P} \sigma P dz \quad (40)$$

From this the log-normality of prices can be show by stepping through process followed by $\ln(P)$ if P follows the process in equation 36. If,

$$G = \ln(P) \quad (41)$$

then

$$\frac{\partial G}{\partial P} = \frac{1}{P}, \quad \frac{\partial^2 G}{\partial P^2} = -\frac{1}{P^2}, \quad \frac{\partial G}{\partial t} = 0 \quad (42)$$

therefore, G follows:

$$dG = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dz \quad (43)$$

Prices are simulated effectively out of equation 43. Moreover, Hull (2012) goes on to assert that the constant drift rate and variance rate of the process for prices implies that prices are distributed log-normally:

$$P_T \sim \phi \left[\ln S_0 + \left(\mu - \frac{\sigma^2}{2} T, \sigma^2 T \right) \right] \quad (44)$$

where P_T is the price at a future time and P_0 is the price at time zero.

APPENDIX G

YIELD TRENDS AND HETEROSKEDASTICITY

The county level yield data are first subset to include only years after 1945; therefore, yields are detrended based upon trend estimates from 1945 to 2013 data. The data are further subset to include only counties with the maximal 69 years of data. A first stage robust linear model is run on yields against time:

$$y_{i,t} = \beta_0 + \beta_1(t) \quad (45)$$

In the absence of heteroskedasticity, detrended yields ($\bar{y}_{i,t}$) are the sum of the largest fitted value ($\hat{y}_{i,t}$) from the first stage regression and the residual ($\hat{u}_{i,t}$):

$$\bar{y}_{i,t} = \hat{y}_{i,t} + \hat{u}_{i,t} \quad (46)$$

Heteroskedasticity is determined at the .015 level and determined with an F-statistic constructed:

$$F = \frac{Var(|\hat{u}_{i,t}| - Var(\hat{u}_{i,t}))}{Var(\hat{u}_{i,t})} * 67 \quad (47)$$

If heteroskedasticity is detected, then heteroskedastically adjusted errors are constructed from a second robust linear estimated:

$$\begin{aligned} |\hat{u}_{i,t}| &= \beta_0 + \beta_1(t) \\ \bar{\hat{u}}_{i,t} &= \hat{u}_{i,t} * \frac{Max(|\hat{u}_{i,t}|)}{|\hat{u}_{i,t}|} \end{aligned} \quad (48)$$

The trend and heteroskedasticity adjusted yields are:

$$\bar{y}_{i,t} = \hat{y}_{i,t} + \bar{\hat{u}}_{i,t} \tag{49}$$

APPENDIX H

MOTIVATION FOR DEVIATION MEASURE

Motivating the development of this deviation measure was a preliminary investigation into the evolution of premium rate differences between as few as two copulas across coverage levels. Initially it was hypothesized that premium rates from one copula would be consistently biased relative to another regardless of coverage level. However, further investigation showed that the direction of premium rate bias could be a function of coverage level and could switch. However, the scale of premium rates varies by orders of magnitude between the lowest and highest coverage levels making it impossible to depict this evolution across all counties and copulas.

Figure 36 shows the county level rates generated by a Frank copula versus a Normal copula for counties that have the most potential for dispersion. Both copulas are symmetric, but when their relative rates are examined under the alternative price distributions, premium rates from a Normal copula will tend to be higher at low coverage levels, but lower at high coverage levels. However, the magnitude of this deviation is almost eliminated at low price dispersion, if not switched. Rates between copulas diverge most at high price dispersion.

Examining the evolution of rate bias of this limited grouping requires a high level of graphical complexity because the premium rates at the highest coverage level are an order of magnitude greater than that of the lowest. Considering all the possible combinations of copulas and characteristic sources of variation, warranted the development of the deviation measure that eliminated the scale difference as a function of coverage level.

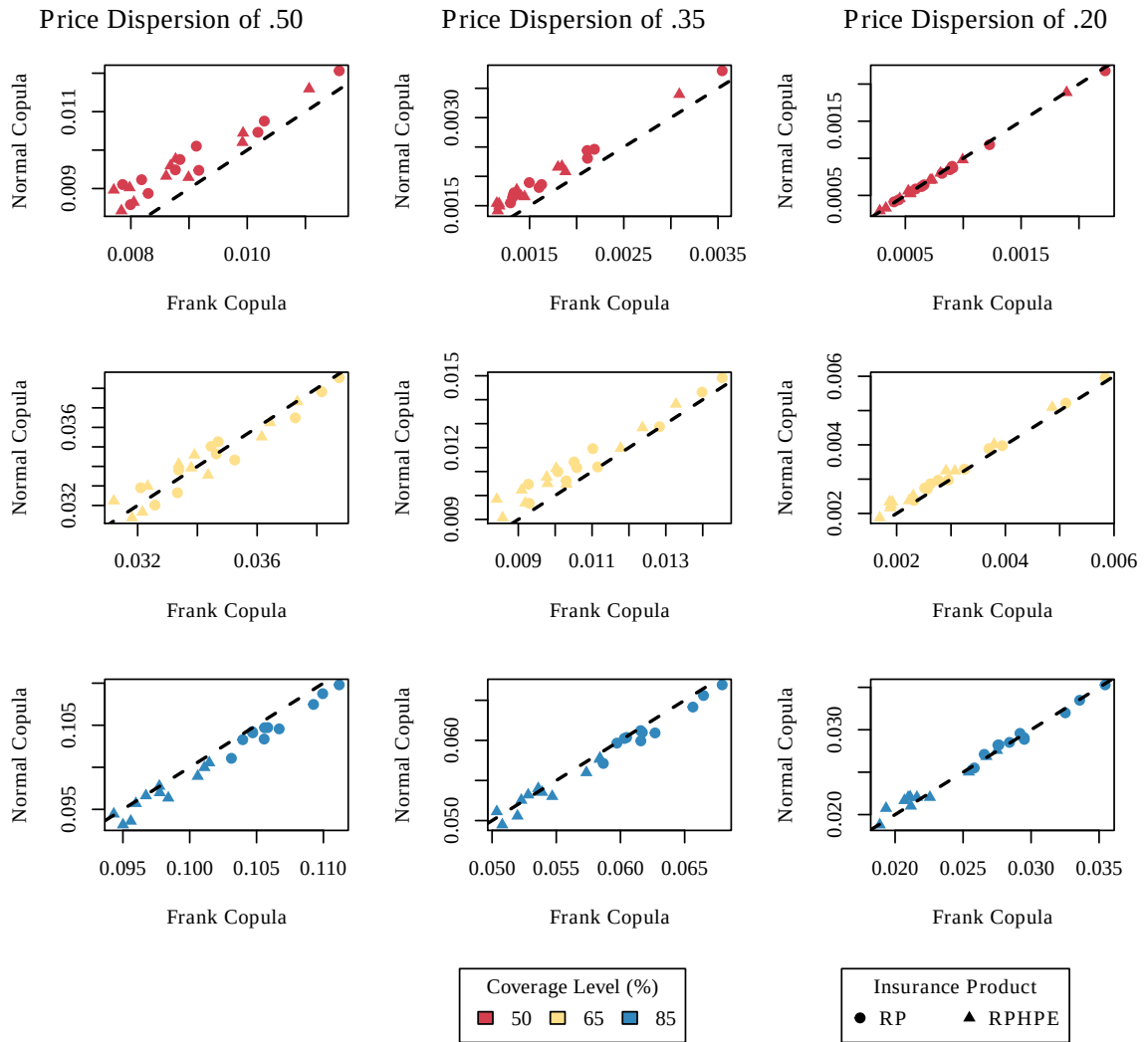


Figure 36: Motivation for Deviation Metric

Notes: These data the premium rates for the 11 Illinois, Indiana or Iowa corn producing counties with low levels of yield deviation ($.125 \leq \text{yield variation as measure by the coefficient of variation} \leq .167$) and strong price-yield dependence ($-.481 \leq \text{Kendall's } \tau \leq -.305$). Revenue Protection (RP) and Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates are plotted. The dashed 45° line represents equality between premium rate pairs.

APPENDIX I

GOODNESS-OF-FIT STUDIES

In order to test the goodness-of-fit of any copula (*gofCop*), the copula’s shape parameter θ must first be estimated. This issue has been extensively addressed in the literature Genest et al. (2009). However, the issue of addressing the null hypothesis that the data being investigated are in fact validly represented by a specific copula, has received less attention (Genest et al., 2009). Recently, the power and nominal performance of proposed goodness-of-fit statistics and the techniques used to implement them have been studied by Beaudoin (2007), Genest et al. (2009)¹⁰³ and Kojadinovic et al. (2011) in large-scale Monte Carlo simulations. Kojadinovic et al. (2011) uses the best performing statistic as identified by Genest et al. (2009), conduct an analogous simulation and explores the performance of a much faster method (*Mult*) for generating realizations of the test statistic and associated p-values. They find that their *Mult* method approaches the performance of the parametric bootstrap (*Pb*) method of Genest et al. (2009), as sample size increases.

The general experimental design shared by the previous studies is to assess the finite sample performance of goodness-of-fit test statistics over a limited set of copulas by generating bi-variate pairs out of a known copula, the “true” copula (*genCop*) used to generate the data. The power¹⁰⁴ and nominal performance¹⁰⁵ of the test statistics are assessed by numerically counting the number of rejections, dividing them by the total number of simulations and comparing the ratio to the 5%

¹⁰³The complete findings from this work are presented in the dissertation of Beaudoin (2007).

¹⁰⁴The power of the test is its power of discernment or ability to correctly reject the H_0 when it is false (i.e. there is a low probability of committing a type-II error).

¹⁰⁵The nominal performance of the test statistic is assessed by whether the percent rejection of an incorrect H_0 hypothesis asymptotically approaches the significance level used as the rejection criterion (i.e. if a 5% significance level is being used, then the false rejection of a true H_0 —or probability of committing a type-I error—should approach 5%).

nominal rejection level.¹⁰⁶

When Genest et al. (2009) performed their experiment, many proposed test-statistics were considered for estimating goodness-of-fit. The design of the study was to identify which of these proposed test statistics proved the most powerful in practice. The authors limited their investigation to a class of tests called “Blanket Tests,” so called because these test provide a goodness-of-fit estimate without user imposed specifications. In all cases the authors estimated the copula’s shape parameter θ by inversion of Kendall’s τ (i- τ). However, several other methods by which the copula parameter can be estimated are available: maximum pseudo-likelihood (*mpl*), the inversion of Spearman’s ρ (i- ρ) and maximum likelihood. The authors found that while most test statistics approached their nominal levels,¹⁰⁷ the Cramér-von Mises S_n statistic had the most power to discriminant between false alternative.

The *Pb* method of generating realizations of the test statistic and p-values, is computationally expensive: Genest et al. (2009) note that their experiment required the almost exclusive use of 140 CPUs over a one month period. When the sample size becomes large or the goodness-of-fit must be assessed many times, this method becomes untenable. In response to this limitation, Kojadinovic et al. (2011) developed a much faster ”multiplier” method (*Mult*) of generating realizations of test statistics and p-values. The authors reduce computation by exploiting the

¹⁰⁶The 5% level is commonly used when reporting the significance of statistics.

¹⁰⁷This means that the Parametric Bootstrap (*Pb*) method the authors developed for generating realizations of the test statistics and p-values, is a consistent estimation process.

properties of multiplier central limit theorems, testing the finite sample performance of their method with an experimental design very similar to Genest et al. (2009).

Ultimately, Kojadinovic et al. (2011) find that the *Mult* method provides reasonable nominal performance at sample size as low as $n = 150$; however, their empirically counted proportions exceeded the nominal significance level at sample size $n = 75$.¹⁰⁸

At smaller sample sizes, the parametric bootstrap method in Genest et al. (2009) holds its nominal level, but has little power to discriminate between false alternatives.¹⁰⁹ While large sample sizes are not realized in crop insurance, the need to repeatedly estimate unique dependencies across many counties makes the fast multiplier approach of Kojadinovic et al. (2011) more attractive. Although the trend of their results can be extrapolated to multiple peril crop insurance (MPCI) data characteristics, neither study explicitly examined the performance of these tests and methods at low data availability and low dependence.

¹⁰⁸The authors do not report these results because of the over-rejection of the a true null-hypothesis. It was hoped that these results could be obtained, for comparison.

¹⁰⁹Results for $n = 50$ were reported in the doctoral dissertation of Beaudoin (2007).

APPENDIX J

CORN AS SPECIAL CASE OF SIMULATIONS

There is an apparent incompatibility between the empirical corn results and the simulated results: the price dispersion with the greatest level of dispersion for the empirical data was 0.35. Yet, the simulated results show that as price dispersion decreases, dispersion tends increases. Perhaps this incompatibility arises from the fact that yield variation of 0.15 is quite distinct from even the next lowest level in the simulated data, but the lowest empirically observed category of yield variation spans this range, possibly introducing some additional noise. Figures 37 and 38 show the simulated dispersion results subset over the empirical corn variation along with the empirical corn premium rate dispersion results.

For price dispersion levels of 0.35 and 0.50, the empirical results and simulated results have a high degree of correspondence. However, at the lowest level of price dispersion the simulated and empirical results diverge. Perhaps most perplexing, is the fact that the empirical results show higher levels of premium rate dispersion for higher coverage levels, especially for the low yield variation groups. There is no case where this is true for the simulated results. Also, the magnitude of variation is not similar between the simulated and empirical results at the lowest level of price dispersion. However, dispersion is similar for the higher coverage levels despite the different scales. For both insurance policies, the empirical results match the simulated results as yield variation increases, and it appears that the empirical RP results even capture the simulated outcome, in which dispersion increases as a function of dependence for the 0.20 yield variation group.¹¹⁰

¹¹⁰The lack of correspondence in the lowest yield variation group may be the results of too much noise in conjunction with a high degree of sensitivity to possible empirical combinations of depen-

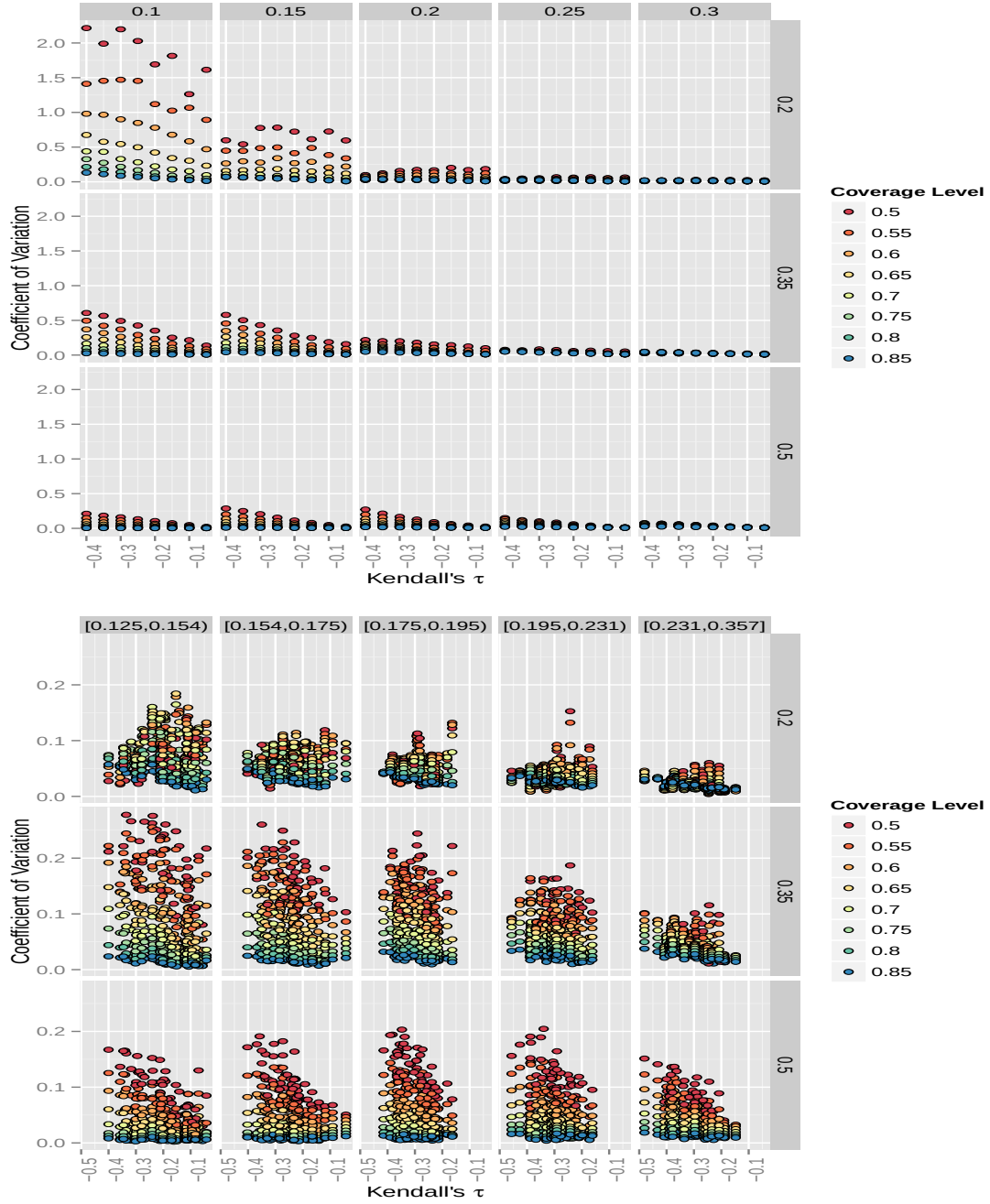


Figure 37: Simulated and Empirical Corn RP Dispersion

Notes: This figure compares the coefficient of variation between Revenue Protection (RP) premium rates generated by alternative copulas from the empirical corn results to the simulated results across there comparable ranges of price dispersion, yield variation and dependence. Empirical premium rate dispersion was calculated for most Illinois, Indiana and Iowa corn producing counties. The separate plots are divided into subplots by grouping the data by farm level yield variation on the horizontal axis and the three levels of price dispersion, used in the empirical experiment, on the vertical axis.

dence and yield variability within the group, which is shown in the simulation results.

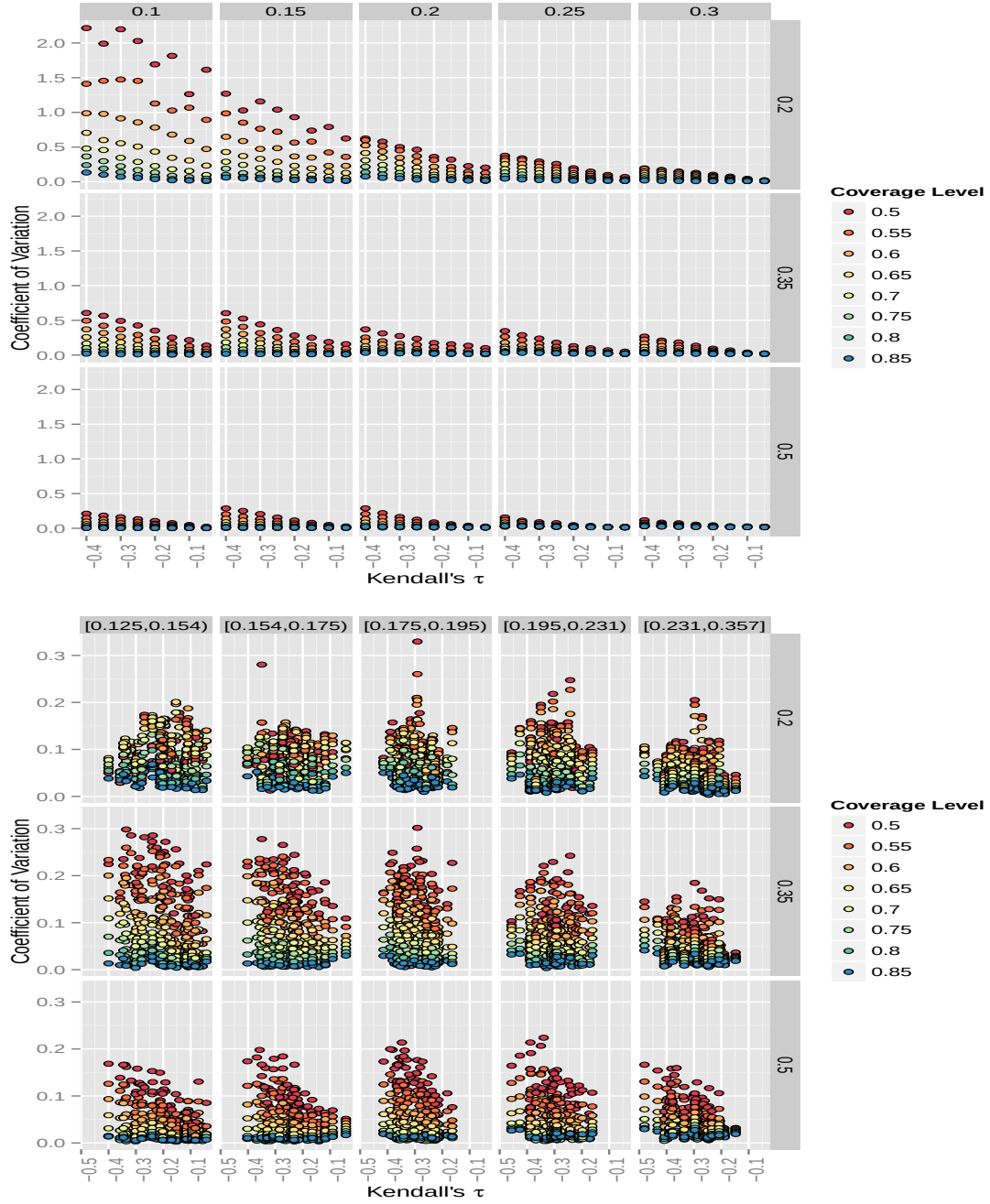


Figure 38: Simulated and Empirical Corn RPHPE Dispersion

Notes: This figure compares the coefficient of variation between Revenue Protection with Harvest Price Exclusion (RPHPE) premium rates generated by alternative copulas from the empirical corn results to a the simulated results across there comparable ranges of price dispersion, yield variation and dependence. Empirical premium rate dispersion was calculated for most Illinois, Indiana and Iowa corn producing counties. The separate plots are divided into subplots by grouping the data by farm level yield variation the horizontal axis and the three levels of price dispersion, used in the empirical experiment, on the vertical axis.