ABSTRACT

ZHU, YING. Modeling Dependence in the Design of Crop Insurance Contracts. (Under the direction of Barry K. Goodwin and Sujit K. Ghosh.)

The overall objective of this dissertation is to evaluate and model the underlying risk factors that farmers face and their implications for crop insurance programs. Because these risk factors are not isolated from each other, the focus of this research is on investigating the dependence structure and the interactions among the agricultural production risks. Using both estimation and simulation methods, we analyze the interaction of risk in the presence of time-varying dimension, portfolio dimension and spatial correlation dimension. This study evaluates the risk that arises from changes in prices, yields shortfalls, or both. Several models are used for price and yield risk factors for corn and soybeans. For instance, yield risks can be modeled by a family of Beta distributions, whereas price shocks can be modeled by log-normal distributions. In order to develop a multivariate model that preserves a given set of marginals, a copula approach can be used to characterize the joint yield and price risk of corn and soybeans, which are usually highly correlated. The copula approach has been spurred by the recent developments in whole farm insurance, resulting in an increasing need for the modeling of multivariate risk factors and their interaction. As a part of the dissertation, various copula models are investigated for their suitability in modeling yield and price risks. A time-varying model is proposed in order to account for the yield distribution changes under technology progress. Finally, various dynamic spatial autoregressive models are explored to account for county-level dependence of crop yields.

Modeling Dependence in the Design of Crop Insurance Contracts

by

Ying Zhu

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Economics and Statistics

Raleigh, North Carolina

2010

APPROVED BY:

Dr. Barry K. Goodwin Co-Chair of Advisory Committee	Dr. Sujit K. Ghosh Co-Chair of Advisory Committee
Dr. Peter Bloomfield	Dr. David A. Dickey
Dr. Xiaoyong Zheng	

DEDICATION

To my parents, Lianfu Zhu and Jieming Dai

献给我的父亲朱廉夫，母亲戴洁明

BIOGRAPHY

Ying Zhu was born in China. She received a Bachelor's of Engineering at Beijing Institute of Technology and a Bachelor's of Laws at Renmin University of China. After working in Beijing for a few years, she came to the United States and started her Ph.D. program in Economics in 2004 at North Carolina State University. Ying received her en route Master's degree in Statistics in 2006 and is now completing a Ph.D. co-major in Economics and Statistics. Upon graduation, Ying will work at SAS Institute, Inc in a research position.

ACKNOWLEDGEMENTS

First, I would like to show my sincerest gratitude to my advisers, who have been top notch and a pleasure to work with. Dr. Barry Goodwin has played a major role in developing the ideas for this work and also helped me with professional advice. He has shown enthusiasm and support throughout this process. Dr. Sujit Ghosh has been generous with his good ideas and with his time by meeting with me on a weekly basis. His continuous support and insightful suggestions made it much easier to accomplish this dissertation. A special thanks also to Dr. Peter Bloomfield, for his expertise on this topic and always being willing to answer my statistical questions; and Dr. David Dickey, for his attention to detail with my work and being a great professor of two of my statistical courses and career advice; and Dr. Xiaoyong Zheng, for his insights and expertise on economics questions. Their insightful comments on my research have improved this work significantly.

I would like to thank my parents for their unwavering support and encouragement. I'd like to also thank my brother Xiaobo Zhu for keeping me grounded and always being a role model for me in so many perspectives. Finally, I would like to thank my husband, Hua Li, for his patience and support. Our children, Yuxuan and Jinyuan, have given to me so much to look forward to.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Often in economics, it is important to understand the distribution of several variables interacting simultaneously, not in isolation from one another. The calculation of a premium to be assigned to an insurance risk is a fundamental quantity for risk management. In crop insurance it is necessary to understand how underlying risk variability arises from changes in prices, yields, or both. Ignoring dependencies can lead to possibly biased and inefficient estimates of the risk. Two reasons for the crop insurance market failure have been proposed. First, the asymmetric information that arises because of adverse selection and moral hazard. Second, the systemic risk due to the great variability in yields over time and spaces. To mitigate this negative effect, an accurate measure of the underlying risk is required. Typically, agricultural risks are not isolated from one another. The underlying risks are dependent in different dimensions, such as time dependence, portfolio dependence, and spatial dependence. Thus, it is important to be able to adequately model dependence with multivariate outcomes. Understanding these dependence concepts will help improve the risk management and will help government, insurance/reinsurance companies, and policy makers to evaluate their contract design and policy making.

The key issues in agricultural risk management include technological changes, weather, the portfolio effect, and systemic risk. To understand these, we need to use statistical and econometric methods to evaluate various dependent risks in agriculture, especially the dependence structure of agricultural risk. This study addresses these important concepts and issues, by using statistical modeling to find evidence and patterns of dependence from agricultural production, and to draw implications for crop insurance programs and risk management.

Background on Federally Regulated Crop Insurance Program

Agricultural activity is affected by a variety of risks. From the time of planting in spring until the harvest time in the fall, agricultural producers are exposed to all kinds of risk, associated with negative outcomes stemming from imperfectly predictable natural adversities (for example, pests and diseases), climatic factors, and unanticipated movements in yields, input and output prices and revenues. All this turns the agricultural sector into a sensitive economic sector. It is important to develop various instruments that farmers can access so that the risk can be pooled and rendered diversifiable. Agricultural risk management has been based on government action of developing relevant social schemes such as direct help, disaster relief, subsidized loans, tax exemptions, etc. Among others, crop insurance is a very important risk management tool, developed to insure farmers against agricultural risk. In the United States, Federally regulated crop insurance programs have been an important part of U.S. agricultural policy since 1938. Under this government program, the actual insurance policies are underwritten and serviced by private insurers and administrated by the authorities through the Risk Management Agency (RMA), which is on behalf of the Federal Crop Insurance Corporation (FCIC). Farmers buy insurance directly from these insurers and the federal government subsidizes farmer's premiums that pay for the crop insurance contracts. At the same time, the crop losses of the private insurers are reinsured by the federal government, via the Standard Reinsurance Agreement (SRA).

The federal crop insurance program currently insures in excess of $70 billion in crops and livestock commodities. A variety of crop insurance plans are available and a number of new pilot programs are under development[1]. Currently there are 22 existing

[1]The crop insurance programs reviewed in this thesis are the crop policies available on the website of Risk Management Agency (RMA) in United State Department of Agriculture (USDA)

crop insurance products, ranging from yield insurance to revenue insurance. Standard crop yield insurance, termed Multiple Peril Crop Insurance (MPCI), pays an indemnity at a predetermined price to replace yield losses. "Group-risk" yield insurance, termed as the Group Risk Plan (GRP), is based on the county's average yield. Under GRP, insured farmers collect an indemnity when the county average yield falls beneath a yield guarantee, regardless of the farmer's actual yields. There exist two farm-level revenue insurance programs: Crop Revenue Coverage (CRC) and Revenue Assurance (RA). These programs guarantee a minimum level of crop revenue and pay an indemnity if revenues fall beneath the guarantee. The recently developed Group Risk Income Plan (GRIP), which is a variation of the GRP, insures county revenues rather than yields.

The federal crop insurance program aims to be actuarially sound, which requires that the price of the insurance policy equals to the expected loss. To guarantee an actuarially fair federal crop insurance program, the calculation of an actuarially fair premium to be assigned to an insurance risk is crucial. This requires accurate estimates of risk distributions and of the probability of a loss event. Loss probabilities are associated with the probability that the underlying factor (yield, price or revenue) below some threshold will be observed, which is given by the area under the probability density to the left of the guaranteed value. We consider yield insurance as an example. Suppose an insurance contract will insure some proportion (λ) of the mean crop yield (y^e). If the observed yield, $y < \lambda y^e$, the insurer will pay ($\lambda y^e - y$) as an indemnity. The expected loss (in bushels) for this insurance contract that guarantees $\lambda \times 100\%$ of the predicted yield (y^e) takes the form of

$$\text{EL}(y) = E[(\lambda y^e - y) I(y \leq \lambda y^e)]$$

where y denotes the observed annual yield and y^e represents the guaranteed yield.

Likewise, price risk considers some assessment of the likelihood that price below some level will be observed.

Besides the issues underlying the measurement of crop yield and price risk, there are issues underlying the measurement of revenue risk and systemic risk. In this case, one must deal with modeling multiple sources of dependent risk — prices and yields of different crops; yield risk from different counties over multiple time periods.

Proper representation of risk measurement may be complicated by the non-independent nature of the underlying risks. Many topics in risk management in agricultural economics require an accurate representation and interpretation of the dependence structure. For example, the distribution of risk might be time-dependent due to technological progress. New complex products in crop insurance result in portfolios with complex dependence structures, which requires measurement of how yield risk may be correlated with price risk. In addition, yield risk can differ systematically over space due to agronomic condition, complicating efforts to accurately model how spatial systemic yield risk changes across space and over time. In order to accurately measure the underlying risks and appropriately design the crop insurance contracts, it is important to investigate and evaluate the dependence structure between yields, yields and prices, spatial heterogeneity in crop yields, and to develop models that incorporate these dependence structures. Economic models that attempt to capture these dependence effects should appropriately incorporate one of such portfolio, temporal and spatial dependence aspects. All these require multivariate models with flexible and dynamic distributions to capture the dependence. Traditional correlation may not be a satisfactory dependence measure to capture the observed complex dependence structures. Wrong assumptions of the dependence structure may lead to severe underestimation or overestimation of the portfolio (insurance) risk. See Frees and Valdez (1999) for an illustration.

This study evaluates and models various dependence structures in the design and rating of crop insurance contracts in efforts to improve the accuracy of premium rates and to provide important insights into the utilization of National Agricultural Statistics Services (NASS) crop yield and price estimates. Particular attention is given to the GRP and GRIP crop insurance programs, which are all based upon NASS county-level estimates. This thesis focuses on issues relating to the non-independence of county estimates of yields and price across space, over time, and among several U.S. major crops. This study develops time-varying distributional models, a copula approach and spatial analysis to address time-dependence, the portfolio effect, and spatial dependence, and to analyze topics that pertain to modeling production and revenue risk of major agricultural commodity data, estimating the probability distributions across commodities, over time and spatial dimensions, and examining the implications for risk transfer and management at the portfolio, temporal and spatial levels.

Organization of Dissertation

First, economic impacts of accurately modeling production risk in the United States are examined by using a time-varying yield distribution that takes into account technological changes. Conventional methods use a two-stage approach to model yield risk. In Chapter 1, *Modeling Yield Risk Under Technological Change: Dynamic Yield Distributions and the U.S. Crop Insurance Program*, a time-varying yield distribution model is developed to measure conditional yield risk under technological change over time. The production risk is analyzed by using less restrictive distribution models, which allow for the parameter of the distributional family to vary according to technology factors that are featured by time. The conventional method and a proposed time-varying method are compared, and simulation estimates are provided for a yield insurance program. The results indicate

that the accuracy benefits of a time-varying model may be as high as $0.92 per unit of insurance contract with a $10,000 of total liability under the group risk plan. The more flexible time-varying model statistically outperforms the alternative conventional model. This implies the time-varying model may better represent the county-level yield risk, and indicates that the conventional model may underestimate the magnitude of yield risk due to technological progress.

Revenue insurance has become more important in recent years. Thus, Chapter 2, *Modeling Dependence in the Design of Whole Farm Insurance Contract — A Copula-Based Model Approach*, examines the actuarial implications of whole farm insurance. To capture the correlated risk factors in the design and pricing of whole farm insurance contract, we propose a multivariate copula method. The results indicate that the portfolio effect is significant when pricing farm risk as a whole. Premium rate differences range from 75% to 120 % due to the portolio effect.

In Chapter 3, *Dynamic Spatial Models for Estimating Systemic Yield Risk*, a spatial, econometric model of systemic risk is developed to investigate potential risk management improvements in area-based insurance ratings and reinsurance. The results indicate the presence of systemic risk (directional) and support what we observed in agricultural production practice. Ignoring the spatial dependence may have significant impacts on crop insurance demand and participation.

The conclusion chapter summarizes the research and discusses the overall implications.

Chapter 1

Modeling Yield Risk Under Technological Change: Dynamic Yield Distributions and the U.S. Crop Insurance Program

Abstract

The objective of this study is to evaluate and model the yield risk associated with major agricultural commodities in the U.S. We are particularly concerned with the nonstationary nature of the yield distribution, which arises primarily because of technological progress in addition to changing environmental conditions. Precise risk assessment depends on the accuracy of modeling the distribution of yields. This problem becomes more challenging as the yield distribution changes over time, a condition that holds for nearly all major crops. A common approach to this problem is based on a two-stage method in which the yield is first detrended and then the estimated residuals

are treated as observed data and modeled using various parametric or nonparametric methods. We propose an alternative parametric model that allows the moments of the yield distributions to change with time and use maximum likelihood method to estimate the parameters. Several model selection techniques (e.g., information criteria, cross-validation measures, non-nested model comparisons) are used to evaluate the suitability of proposed models for yield data obtained from major crop producing counties. In a simulation study, the two alternative yield modeling procedures are used to price an insurance contract that covers the yields at a county level. Our results suggest that the proposed time-varying model outperforms the more conventional model in terms of in-sample goodness-of-fit, out-of-sample predictive power and the prediction accuracy of insurance premium rates.

1.1 Introduction

The Federal Crop Insurance program represents a major component of U.S. agricultural policy and is intended to protect farmers from yield and revenue risk. Accurate modeling of crop yield distributions is essential for the proper design of crop insurance contracts and to the maintenance of an actuarially sound insurance program. Historical agricultural yield data suggest a strong upward trend in annual crop yields (figure 1.1(a)). Advances in technology, germplasm, breeding techniques, and the development of new hybrids may significantly affect the distributions of crop yields, in addition to changes in environmental factors. In particular, various moments of the distributions may evolve over time as technology progresses. These changes can complicate efforts to accurately model yield distributions using data observed over time.

Many studies have attempted to determine the distributional model and estimation methods that best characterize crop yield distributions. These studies have indicated that agricultural yields can be modeled in many different ways. Modeling approaches in the current literature range from non-parametric (Goodwin and Ker, 1998) to parametric methods. The parametric approach of modeling yields usually involves selection and specification of candidate distribution families, parameter estimation and goodness-of-fit assessments. Among others, the Beta distribution is popularly used in practice due to its flexibility and capacity for representing the skewness typically associated with crop yield distributions. The notion of a conditional Beta distribution for yields was introduced by Nelson and Preckel (1989). Some other popular candidates used in the literature include, among others, the lognormal distribution (Day, 1965), the Normal distribution (Just and Weninger, 1999), the Weibull distribution (Chen and Miranda, 2004) and the Logistic distribution (Sherrick et al., 2004). Evidence of non-normal yields has been presented by a number of authors, including Taylor (1990), Ramirez (1997) and Ramirez, Misra, and Field (2003).

In many cases, annually-averaged agricultural yields display a strong upward trend over time and the deviations from trend (residuals) frequently display heteroscedasticity (figure 1.1(a)). A very common approach to modeling yield risk using time-series data has been to first detrend the time series data and then estimate the yield distribution using the detrended yield data, thus treating the estimated, detrended yields as "observed" data. These approaches are often referred to as "two stage" methods; the first stage fits a trend to the data while the second stage uses the detrended data to model the distribution. Examples of such two-stage detrending procedures can be found in, among others, Miranda and Glauber (1997), Swinton and King (1991), Atwood, Shaik, and Watts (2003).

In this two-stage method, it is crucial to determine the correct functional form of the regression representing trend in the first stage and then to establish the correct distributional properties of the detrended data, including such characteristics as skewness, kurtosis, and heteroscedasticity. However, it has been recognized that the resulting estimated residuals, representing the detrended yields, are subject to the estimation uncertainty associated with the detrending method and thus may not necessarily provide an accurate representation of the actual yield distribution. Therefore, the two-stage method of estimating yield distributions may not always adequately capture the estimation uncertainty associated with sampling variability in the first stage estimates of trend. Although any biases induced at the first-stage asymptotically approach zero when the correct functional form is used in the regression and errors are assumed to be homoscedastic, the uncertainty induced at the first stage, if not accounted for in the second stage estimates of the yield distribution, will lead to inaccurate estimation of the variance in the final estimates. The magnitude of this effect can be large especially when the errors are heteroscedastic (Robinson (1987)) and thus can potentially induce significant adverse selection into an insurance program if ignored.

This standard two-stage method has been one of the most popular approaches to removing time trends from crop yields and then modeling the distribution of residuals (usually under the assumption of homoscedasticity) for the purposes of defining and rating crop insurance contracts. A similar two-stage method is used to rate the Group Risk (GRP) and Gross Revenue Insurance (GRIP) programs, though this method does address the potential for heteroscedasticity. However, it is possible to account for the uncertainties associated with the first stage estimates by applying an alternative simultaneous method. At the same time, it is possible to adequately represent characteristics of the yield distribution such as deterministic trends and heteroscedasticity by using alternative estimation techniques. We propose a likelihood based estimation method that simultaneously estimates the trend (conditional mean) and higher moments of the yield density by using a flexible class of parametric distributions. Admittedly, any parametric class of distributions is subject to model misspecification, so we also provide a set of model validation tools that enables a researcher to test the validity of the proposed class of distributions in approximating the true underlying data generation mechanism.

This represents a potential improvement over more common methods typically applied. The method along with its validation measures proposed here allows one to measure conditional yield risk in a dynamic setting and thereby calculate premium rates for crop insurance contracts in a more accurate and systematic way. To analyze the evolution of the yield distribution, we construct a flexible parametric model that attempts to capture the deterministic time (technology) effects associated with the yield distribution parameters. This method essentially models the trend of all moments of the distribution simultaneously by allowing shape, scale, and location parameters of the specific distributional family to evolve over time, while the more common two-stage method usually allows one to model only the mean and possibly the variance to reflect

the change over time. A more complete and coherent picture of technological progress and the consequential changes in yield risk can be provided by simultaneously modeling the time trend and the distributional parameters.

The remainder of the paper is organized as follows. Section 2 outlines a conventional two-stage approach to modeling yield distributions. Section 3 presents the proposed time-varying approach for econometric modeling of yield distributions and illustrates the empirical application of the time-varying model to county-level agricultural yield data for several important crops. Section 4 presents an evaluation of competing models. A simulation study is conducted in Section 5 to compute the insurance premium rate of these models, and its implications are discussed. The final section offers concluding remarks.

1.2 A Conventional Two-Stage Estimation Framework

A common way to capture the effects of technological changes in agricultural production includes modeling the deterministic or stochastic yield effect (trend). In most empirical analyses, a deterministic (rather than stochastic) trend is used to capture the movement in expected yields and thus to represent the variation of yields around this expected level. The main justification for using a deterministic component is that, if crop yield variables evolve slowly through time, then approximation of a deterministic component may be sufficient to model the yield distribution (Just and Weninger, 1999).

In this two-stage estimation model, the trend component is usually controlled for before assessing the distribution of yields–generally based on a homoscedastic model. The usual approach to removing deterministic factors in yields is to use a parametric or a nonparametric regression model. Popular regression models include a log-linear specification based on polynomials, kernel regression, smoothing splines, and partial

linear models (Gyorfi et.al (2002)). We illustrate this idea by using a parametric model with a quadratic trend as well as a nonparametric trend model.[1]

Consider the following trend model:

$$y_t = m(\mathbf{x}_t) + \varepsilon_t \tag{1.1}$$

where y_t is the observed crop yield in year t, $(t = 1, \ldots, T)$, $m(\mathbf{x}_t)$ denotes the regression function $E(Y_t|X_t = x)$, and ε_t represent residuals that are assumed to be identically and independently distributed (i.i.d.) with mean zero. The regression function $m(\cdot)$ can be estimated nonparametrically using kernel methods or smoothing spline methods. Alternatively, if we assume a parametric functional form for $m(\cdot)$, then the regression coefficients can be obtained using ordinary least squares (OLS).[2] In either case, the residuals are obtained as $\hat{\varepsilon}_t = y_t - \hat{m}(x_t)$. We considered both quadratic and nonparametric trend models. The Kolmogorov-Smirnov (K-S) 2-sample goodness of fit (GOF) test suggests that the two residual distributions are not significantly different between the nonparametric and parametric models based on the data in this study. On the basis of this test, the quadratic detrending method is used as a benchmark.

Our empirical analyses presented in this paper are based on applications to National Agricultural Statistics Services (NASS) county-level average yields.[3] Figure 1.1(b) presents the nonparametric residual plot of annual corn yields in Iowa, which shows that the deviations from the trend tend to be proportional to the level of the yields. There is an apparent increase in yield variation over time. To account for this temporal

[1] The selection of these two trend models is intended to provide a benchmark for comparison purposes. There are other detrending methods such as log-linear regression. Since the focus of this study is to compare the two-stage approach and the time-varying method that we propose as an alternative, we use representative methods to illustrate the concepts. A comprehensive survey of all possible detrending methods is beyond the scope of this study.

[2] We assume that $m(x_t) = m_0(x_t, \beta)$, where m_0 is a known functional form up to some finite dimensional regression coefficient vector β.

[3] The data are available at the NASS website at http://www.nass.usda.gov.

heteroscedasticity effect, a rescaled form of the absolute value of deviations from a trend-based, forecasting equation is often suggested. This approach, though ad hoc, is commonly used in practice (see, for example, Miranda and Glauber (1997), Atwood, Shaik, and Watts (2003)). In particular, the detrended normalized values are obtained as:

$$\tilde{y}_t = \hat{y}_T^{(1)}(1 + \frac{\hat{\varepsilon}_t}{\hat{y}_t^{(1)}}) = \frac{\hat{y}_T^{(1)}}{\hat{y}_t^{(1)}} y_t \tag{1.2}$$

where $\hat{y}_t^{(1)} = \hat{m}(x_t)$ and $\hat{y}_T^{(1)} = \hat{m}(x_T)$ represent the predicted values of year t and T, respectively from stage I (equation 1.1) and $\hat{\varepsilon}_t = y_t - \hat{y}_t^{(1)}$. A similar approach has been applied by, among others, Miranda and Glauber (1997). Instead of using standard additive regression residuals, proportional residuals are often used in an effort to achieve homocedasticity in modeling the detrended yield distribution in the second stage. By dividing each error by its associated forecast, the residuals can be scaled to the year T equivalent forecasted yield.

We use a goodness of fit (GOF) specification test to determine the appropriate distribution for the detrended yield $\tilde{y}_t$. A Q-Q plot based on the residuals $\hat{\varepsilon}_t$ (figure 1.1(b)) suggests that the residuals are more negatively skewed than what would be implied by the normal distribution, which suggests that a Beta distribution may be a viable candidate. A GOF test for the Beta distribution (based on a Chi-square statistic) suggests that a Beta distribution provides a reasonable fit for county-level yields. There is considerable empirical evidence from county-level data that the normalized yields are well approximated by a Beta distributed random variate. For example, the GOF test yields a p-value of 0.51 for Kossuth County and 0.62 for Adair County Iowa all-practice corn yields. We use $Beta(\alpha, \beta, \theta, \delta)$ to denote a Beta distribution with

location parameter $\theta \geq 0$, scale parameter $\delta > 0$, and shape parameters $\alpha, \beta > 0$.[4] This implies that the actual yields follow a Beta distribution with constant shape parameters and time-varying location and scale parameters, i.e., $y_t \sim Beta(\alpha, \beta, \tilde{\theta}_t, \tilde{\delta}_t)$, with $\tilde{\theta}_t = \hat{\rho}_t\theta, \tilde{\delta}_t = \hat{\rho}_t\delta$ and $\hat{\rho}_t = \frac{\hat{y}_t^{(1)}}{\hat{y}_T^{(1)}}$. The log-likelihood function of a general Beta distribution based on the detrended data $\tilde{y}_t$ with two shape parameters α, β and location θ and scale δ parameters, is given by,

$$\begin{aligned}\text{LLF}(\alpha, \beta, \theta, \delta | \tilde{y}_t, t = 1, \ldots, T) =& (\alpha - 1) \sum_{t=1}^{T} \log(\tilde{y}_t - \theta)^+ + (\beta - 1) \sum_{t=1}^{T} \log(\delta + \theta - \tilde{y}_t)^+ \\ & - T \log(B(\alpha, \beta)) - T(\alpha + \beta - 1) \log(\delta) \end{aligned} \tag{1.3}$$

where $\log(B(\alpha, \beta)) = \log(\Gamma(\alpha)) + \log(\Gamma(\beta)) - \log(\Gamma(\alpha + \beta))$ and $\log a^+ = \log a$ if $a > 0$ and $\log a^+ = 0$ otherwise, which ensures that $\theta \leq \tilde{y}_t \leq \theta + \delta \ \forall t$, for any $\theta, \delta > 0$. Notice that this implies that a pseudo log-likelihood function of parameters based on the original data y_t is

$$\begin{aligned}\text{LLF}(\alpha, \beta, \theta, \delta | y_t, t = 1, \ldots, T) =& (\alpha - 1) \sum_{t=1}^{T} \log(y_t - \tilde{\theta})^+ + (\beta - 1) \sum_{t=1}^{T} \log(\tilde{\delta} + \tilde{\theta} - y_t)^+ \\ & - T \log(B(\alpha, \beta)) - T(\alpha + \beta - 1) \log(\tilde{\delta}) \end{aligned} \tag{1.4}$$

where $\tilde{\theta}_t = \hat{\rho}_t\theta, \tilde{\delta}_t = \hat{\rho}_t\delta$ and $\hat{\rho}_t = \frac{\hat{y}_t^{(1)}}{\hat{y}_T^{(1)}}$.

We obtain the parameter estimates $(\hat{\alpha}, \hat{\beta}, \hat{\theta}, \hat{\delta})$ by maximizing the $\text{LLF}(\alpha, \beta, \theta, \delta)$ based on the normalized values of $\tilde{y}_t$. The results are presented in table 1.1. The predicted detrended value is calculated as: $\hat{\tilde{y}} = (\hat{\theta} + \hat{\delta}\frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}})$. The predicted actual mean

[4] In other words, $\frac{\tilde{y}_t - \theta}{\delta} \sim Beta(\alpha, \beta)$, where $Beta(\alpha, \beta)$ represents a standard Beta distribution defined on $(0, 1)$ with shape parameters $\alpha, \beta > 0$.

yield can be calculated from the detrended model as:

$$\hat{\hat{y}}_t = \frac{\hat{y}_T^{(1)}}{\hat{y}_t^{(1)}}\hat{\hat{y}}_t = \frac{\hat{m}(x_T)}{\hat{m}(x_t)}(\hat{\theta} + \hat{\delta}\frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}}) \tag{1.5}$$

We use these second stage predicted values $\hat{\hat{y}}_t$ to compute expected yield loss and compare with the results obtained by the proposed time-varying Beta model.

Using a first stage estimation to detrend yield data and then treating the resulting detrended yields as if they were observed without error may not be appropriate in that the first stage estimation error is ignored (e.g., $\hat{\varepsilon}_t$'s are assumed known for the LLF in equation 1.3.) Robinson (1987) pointed out that there is a loss of efficiency of the OLS estimator when the data generating mechanism is in fact heteroscedastic. A more systematic inferential method might be needed to accurately capture trend effects and model conditional yield risk.

1.3 A Time-Varying Yield Distribution Model

In this section, a flexible class of parametric models is proposed which allows us to simultaneously and coherently specify the first four moments using suitable polynomials of time, and the coefficients of the polynomials are estimated simultaneously by maximizing the resulting likelihood function. Several alternative models are examined to measure conditional yield risk. For instance, instead of using polynomials to model the first four moments of the proposed distribution, one may use knot-based splines. As we have noted, this approach differs from most approaches in the existing literature, which typically estimate trends and distributional parameters in two consecutive but independent steps. The time-varying model accounts for parameter uncertainty by simultaneously estimating the entire likelihood function, which includes time-trend parameters and the distributional parameters in one step. The results of this proposed

model are compared to those based on the conventional two-stage approach described in the previous section for several important crops and counties drawn from U.S. county-level data.

Our proposed modeling strategy, which attempts to estimate the entire structure of the yield distribution model jointly, accounts for the complexities of accurately capturing yield distributions and characterizes probabilistic models of the temporal process of yields involved with crop production. The basic assumption of the time varying model is that the parameters of the distribution follow a specific temporal pattern, such that the whole temporal changes of the yield distribution can possibly be captured by the time-varying shape and scale parameters. The model is estimated by maximizing a Beta log-likelihood function which includes time-varying shape, location and scale parameters.

These time varying parameters evolve over time according to an exponential form. As we discuss below, this particular functional form ensures that the Beta shape, scale, and location parameters are positive at every observation.[5] The log-likelihood of the time-varying Beta distribution is as follows.

$$\begin{aligned}\text{LLF}(\mathbf{b}|y_t, t = 1, 2, ..., T) =&(\alpha_t - 1)\sum_{t=1}^{T}\log(y_t - \theta_t)^+ + (\beta_t - 1)\sum_{t=1}^{T}\log(\delta_t + \theta_t - y_t)^+ \\ &- T\log(B(\alpha_t, \beta_t)) - T(\alpha_t + \beta_t - 1)\log(\delta_t) \qquad (1.6)\end{aligned}$$

where $\log(B(\alpha_t, \beta_t)) = \log(\Gamma(\alpha_t)) + \log(\Gamma(\beta_t)) - \log(\Gamma(\alpha_t + \beta_t))$ with $\alpha_t, \beta_t, \theta_t, \delta_t$ being functions of time and depending on a finite-dimensional vector $\mathbf{b}$ of parameters, i.e., $\alpha_t = \alpha(t, \mathbf{b})$, $\beta_t = \beta(t, \mathbf{b})$, $\theta_t = \theta(t, \mathbf{b})$ and $\delta_t = \delta(t, \mathbf{b})$. We evaluated two different time trend structures for the parameters of the yield distributions—a standard linear trend and a quadratic trend model. However our method is not restricted to these

[5] Of course, other functional forms including quadratic specifications could be used to ensure positive parameters. Alternatively, one may also specify functional form using the first four moments of the Beta distribution, which may require a constrained optimization of the likelihood function.

chosen functional forms. For instance, quite generally we can model any of these Beta parameters as $\exp\{\sum_{j=1}^{J} \psi_j(t) b_j\}$, where $\psi_j(\cdot)$'s may represent members of a collection of J basis functions (e.g., choosing $\psi_j(t) = t^{j-1}$ we obtain polynomials while choosing $\psi_j(t) = (t - t_j)_+^3$ we obtain cubic splines with knots t_j's). The time-varying parameter estimates of these structural models are obtained by a standard MLE method. Because the quadratic specification nests the linear trend, a standard likelihood ratio test can be used to evaluate the statistical significance of the quadratic terms and thus to select the optimal trend specification. Note that the Beta distribution is characterized by four parameters (α, β, θ, and δ). For simplicity and numerical stability of the maximum likelihood approach, we fix the minimum possible yield to be equal to zero in each case (i.e., by setting $\theta_t = 0$ for all t). We allow each parameter of the Beta distribution to vary over time through a functional relationship of the form $\alpha = exp(f(\mathbf{b}, t))$ where $f(\cdot)$ is a linear or quadratic function of time (though other functional forms such as knot-based splines mentioned earlier could also be used). Such a specification allows us to use an unconstrained maximization of the likelihood function. As our results demonstrate below, the quadratic terms were not found to be statistically significant for the data sets that we have analyzed and thus our final representation of the conditional mean uses a standard linear trend.

The predicted value $\hat{y}_t$ from the time-varying model is calculated by simultaneously estimating the time trend and the parameters of the distribution in a full MLE approach. Let independently distributed yields data follow a time-varying Beta, i.e., $y_t \sim Beta(\alpha_t, \beta_t, \delta_t), t = 1, \ldots, T$, where $\alpha_t = \alpha(t, \mathbf{b}), \beta_t = \beta(t, \mathbf{b}), \delta_t = \delta(t, \mathbf{b})$, $\mathbf{b} = (b_1, \ldots, b_K)^T$. $\hat{\mathbf{b}}$ denotes the MLE obtained by maximizing the log-likelihood of equation (1.6). The predicted yield of the time-varying model is given by

$$\hat{y}_t = \hat{\delta}_t \frac{\hat{\alpha}_t}{\hat{\alpha}_t + \hat{\beta}_t} \tag{1.7}$$

where $\alpha_t = \alpha(t, \hat{\mathbf{b}}), \beta_t = \beta(t, \hat{\mathbf{b}})$, and $\delta_t = \delta(t, \hat{\mathbf{b}})$.

1.3.1 *Empirical Application*

The time-varying model not only addresses the dynamic characteristics of yield distributions, but also provides a more flexible specification of heteroscedasticity and higher order moments (e.g., skewness and kurtosis). We implement the time-varying model by applying the methods to county-level yield data in major crop producing areas obtained from the National Agricultural Statistics Service (NASS). In particular, the top 10 producing counties are selected in the major producing states for corn, soybeans, cottons. These county/crop combinations include the following: Iowa all-practice corn from Kossuth, Sioux, Pottawattamie, Plymouth, Webster, Pocohontas, Hardin, Franklin, Clinton and Woodbury counties; Iowa soybeans from Kossuth, Sioux, Pattawattamie, Plymouth, Webster, Woodbury, Benton, Grundy, Crawford and Tama counties; Texas upland cotton from Gaines, Lubbock, Hockley, Lynn, Dawson, Hale, Terry, Crosby, Floyd and Martin counties.

It is widely recognized that the rate of technological progress varies considerably across different crops. Our results are presented in figure 1.2 and demonstrate that Iowa corn and soybean yields are skewed, kurtotic and exhibit strong time trend effects and varying degrees of heteroscedasticity through time. In contrast, Texas cotton yields appear to have a more modest time trend, though strong evidence of temporal heteroscedasticity is exhibited.

By maximizing the LLF($\mathbf{b}|y_t, t = 1, \ldots, T$) as defined in equation (1.6) using county-level yield data (Adair County Iowa all-practice corn is presented as an example), the maximum likelihood estimates of this time-varying Beta distribution with a linear time trend in the exponent and a quadratic time trend structure are shown in table 1.1. A likelihood ratio test statistic of the two alternative models has a value of 4.12, which

does not reject the null hypothesis that the quadratic trend parameters are equal to zero and thus supports the adequacy of the linear specification.

Substitution of the MLE estimates can be used to evaluate the time-varying Beta density for any given year. Figures 1.2(d), 1.2(e) and 1.2(f) illustrate the dynamic evolution of the densities that are estimated by each time-varying model for corn, soybeans and cotton yields. Various moments of the distributions appear to evolve over time. The density plots of these estimated time-varying distributions suggest different means, skewness coefficients, and maximum values of corn yields for each year. In figures 1.2(a), 1.2(b) and 1.2(c), we present estimated densities for both the time-varying model and the more conventional detrended model. In every case, the time-varying densities show a smaller degree of leptokurtosis than is the case for standard, two-stage detrended yield data.

Table 1.2 presents log-likelihood values for the two alternative models for a number of counties. In almost every case, the LLF value of the time-varying model is greater than the conventional model, even after adjustments (within the context of alternative information criteria) for the number of parameters. This suggests a better goodness-of-fit for the time-varying model. This is also illustrated in figure 1.3, which contains a side-by-side bar plot of the LLF values for all major county/crop combinations considered in our analysis.[6]

1.4 Model Performance and Specification Tests

We considered a number of specification tests and evaluations of forecasting performance of the alternative models. Vuong's nonnested specification test (Vuong (1989)) is a

[6]MLEs for these other counties are available upon request from the authors.

likelihood-based test for model selection. Vuong's test statistic is given by:

$$v = \frac{n^{\frac{1}{2}} LR_n(\hat{\theta}_n, \tilde{\theta}_n)}{\hat{\omega}_n} \tag{1.8}$$

where $LR_n(\hat{\theta}_n, \tilde{\theta}_n) = L_n^f(\hat{\theta}_n) - L_n^g(\tilde{\theta}_n)$, $L_n^f(\hat{\theta}_n)$ is the maximum likelihood function of the time-varying model and $L_n^g(\tilde{\theta}_n)$ is the maximum likelihood function of the detrended model. $\hat{\omega}_n$ is defined as:

$$\hat{\omega}_n^2 = \frac{1}{n}\sum_{t=1}^{n}\left(\log\frac{f(Y_t|X_t;\hat{\theta}_n)}{g(Y_t|X_t;\tilde{\theta}_n)}\right)^2 - \left(\frac{1}{n}\sum_{t=1}^{n}\log\frac{f(Y_t|X_t;\hat{\theta}_n)}{g(Y_t|X_t;\tilde{\theta}_n)}\right)^2$$

The test statistic v is approximately distributed as a standard normal random variable. As specified, if $v > c$, where c is the critical value[7], we reject null that the models are the same in favor of the alternative time-varying model F_θ. Alternatively, if $v \leq -c$, we would reject the null in favor of the detrended model G_θ. Vuong's test statistics v are presented in table 1.2 and in a majority of cases (87%) support the time-varying specification.

Table 1.2 also presents goodness-of-fit comparisons for conventional models (model I) and time-varying models (model II) based on the Akaike Information Criterion (AIC) (Akaike (1974)) and Schwarz's Bayesian Information Criterion (BIC) (Schwarz (1978)). Smaller values of the AIC or BIC indicate a better fit. Both figure 1.3 and table 1.2 show that the time-varying Beta has the lowest AIC and BIC for most if not all counties, which indicates that it is the most parsimonious and optimal model among the set of models that we have considered in this article. Moreover, ΔAIC ($\Delta AIC = AIC - \min(AIC)$) and ΔBIC ($\Delta BIC = BIC - \min(BIC)$) in table 1.2 are significantly large for the

[7]We can choose a critical value c from the standard normal distribution that corresponds to the desired level of significance (e.g. for $c = 1.96$; $Pr(z \geq |\pm c|) = 0.05$).

conventional detrended Beta model,[8] which also offers evidence in support of the time-varying model (Burnham and Anderson, 2003, section 2.6).

Table 1.3 presents the comparison results of ten-year, out-of-sample forecast, two-step-ahead forecast and a cross validation (leave-one-out) test. The out-of-sample forecast method essentially evaluates which method is better in forecasting the first moment of yields. This, of course, has direct relevance for the estimation of crop yield distributions and the subsequent rating of crop insurance contracts. Note, however, that these tests only compare models in one aspect of the yield distribution—the first moment (the mean). Thus, likelihood based specification tests may provide more information about goodness of fit for the entire distribution assuming that one model is correct.

The cross-validation method ranks competing models based on their out-of-sample forecasting performance with some observations being randomly left out. For example, the leave-one-out cross-validation test is conducted for all counties considered for Iowa all-practice corn for the 82 years of county-level annual yields from 1926 to 2007. We drop each observation from the sample, fit the model, and use the estimates to forecast the omitted observation. The predicted and actual yields are compared to get the cross-validation Root Mean Squared Errors (RMSE) in each period.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_{(i)})^2}$$

where $\hat{Y}_{(i)}$ is the prediction for Y_i obtained by fitting the model with observation i omitted.

We sum the cross-validation errors and obtain the RMSE for the two competing models. Results (table 1.3) indicate that the time-varying Beta distribution model out-performs the constant Beta model in most of the major agricultural production

[8] As an example, $\Delta AIC = 88.16, \Delta BIC = 88.25$ for detrended model for Webster county soybean yields in Iowa.

counties. Specifically, eight of the ten top Iowa corn production counties, nine of ten Iowa top soybean production counties, and six of seven Texas top cotton production counties exhibit a better cross-validation performance in the time-varying model. The resulting RMSEs of the time-varying model for these yield data are smaller than that of the conventional model. The differences of the RMSE between the two competing models are bigger for corn and cotton than soybeans. This is consistent with what we have observed in the practice of genetic improvement and biotechnological progress in agriculture. There have been less biotech innovations for soybeans than for corn and cotton and therefore the yield distribution of soybeans is less affected. As a result, the two competing methods do not make a big difference in out-of-sample predictive power for soybeans yields. In addition to computing RMSEs, one may also compute the Spearman's correlation between the Y_i's and $\hat{Y}_{(i)}$'s or generate a Q-Q plot to check other distributional characteristics between the observed and (leave-one-out) predicted values.

In the current group risk crop insurance programs in the U.S., yields are forecast two years into the future. These forecasts are then used to establish insurance guarantees. In light of this, we considered an additional out-of-sample forecast evaluation intended to provide an analog to the forecasts used in these area-wide programs. In this approach, models are ranked based on their out-of-sample forecasting performance for a series of two-year ahead and ten-year ahead forward forecasts. The out-of-sample predictive accuracy of the alternative models is again compared to see which model performs better. For example, to predict 1993's yield, the estimates are based on the sample from 1926 to 1991; to predict 1994's yield, the estimates are based on the sample from 1927 to 1992, etc. The two-year-ahead prediction errors are recorded each year by moving forward to predict yields from 1993 up to 2007 and the average of this two-year ahead prediction

error ($RMSPE_2$) is given by:

$$\text{RMSPE}_2 = \sqrt{\frac{1}{15}\sum_{t=67}^{80}(\hat{y}_{t,t+2} - y_{t+2})^2} \tag{1.9}$$

where $\hat{y}_{t,t+2} = E[y_{t+2}|y_1, \ldots, y_t]$ is the predicted value of y_{t+2} based on observing $y_1, \ldots, y_t$. Another out-of-sample test is conducted by partitioning the entire sample into two parts and estimating parameters based on the first part of the data for the period 1926 to 1997 (the first 72 observations), then the estimated parameters are used to compute the expected (mean) yields for the out-of-sample period spanning 1998 to 2007 (the second part of the data). Prediction errors are computed by comparing the predicted yields and the actual yields out-of-sample. The mean of the squared difference between the predicted value and the actual yield value is calculated as a "leave-ten-out" forecast error $RMSPE_{10}$.

The out-of-sample measures are computed for selected major crop/county combinations in the U.S. and such predictive measures again provide comprehensive evidence that the time-varying approach represents an improvement across all criteria considered. Table 1.3 shows that time-varying model has smaller values of both $RMSPE_2$ and $RMSPE_{10}$ in most cases. However, there are some cases in table 1.3 showing that the conventional model has smaller values of $RMSPE_2$ and $RMSPE_{10}$ as well. For example, Iowa corn production in Plymouth and Franklin counties, Iowa soybeans in Webster county and Texas cotton in Gaines and Crosby counties show smaller $RMSPE_2$ or $RMSPE_{10}$ when fitted with the conventional model. Having noted this, we must point out that the out-of-sample comparison test is only based on the accuracy of first moment mean prediction, which is not an overall evaluation of the entire yield distribution. Since the time-varying model is an alternative to the conventional two-stage model to estimate the yield distribution and to forecast the mean, these two models

may display different out-of-sample performance based on different yield data in terms of mean prediction. Strong evidence is still found, as presented in table 1.2, that supports the time-varying model's performance in estimating the entire yield distribution in terms of likelihood based tests and nonnested model distribution tests.

Table 1.4 presents alternative methods to compare the two competing models, based on the regression method and the rate cross validation method. By using the regression method, we can consider which model's predicted values better explain the variation of the actual yields. The regression model is shown as:

$$y_t = \gamma_0 + \gamma_1 \hat{\hat{y}}_t + \gamma_2 \hat{y}_t + e_t \tag{1.10}$$

where $\hat{\hat{y}}_t$ is obtained from the detrended model (equation 1.5) and $\hat{y}_t$ is taken from the time-varying model (equation 1.7).

The first panel in table 1.4 presents the results of regressing the actual yield on the predicted yields from the two competing models. The results indicate that only the coefficient γ_2 of predicted yields from the time-varying model is significantly different from zero in equation (1.10), which suggests the time-varying model yields a better prediction of the actual yield. Further, the intercept term is also not significantly different from zero, indicating that the chosen model has no systematic bias. The null hypothesis of $H_0 : \gamma_2 = 1$ suggests a large p-value equal to 0.465 which again indicates that the chosen model has no scale bias.

1.5 Simulation of a Group Risk Insurance Program

Yield–based insurance policies in the federal crop insurance program include the individual, farm-level multiple peril crop insurance (MPCI) and the county-level Group Risk Plan (GRP), which is based upon county-average yields from NASS. An important

policy parameter in the GRP program is the premium rate, which is based on the county-average yield distribution. In this section, we evaluate the economic impacts of adopting rates based on the time-varying distribution methods. If the yield distributions are changing over time, the premium rates should be adjusted accordingly. The premium rates from the proposed time-varying approach are illustrated with simulated data and a rate cross-validation test is conducted to compare the predictive accuracy of the premium rates from the time-varying approach with those of the conventional two-stage approach. Standard crop yield insurance pays an indemnity at a predetermined price to replace yield losses. Under the GRP, insured farmers collect an indemnity when the county average yield falls beneath a guarantee, regardless of the farmers' actual yields. Loss probabilities correspond to the likelihood that yields y below some threshold will be observed, which is given by the area under the density curve to the left of the guaranteed yield. Suppose an insurance contract will insure some proportion ($\lambda \in (0,1)$) of the expected crop yield (y^e). If $y < \lambda y^e$, the insurer will pay $(\lambda y^e - y)p$ as an indemnity, where p is a predetermined price. An actuarially fair premium is defined by the expected loss of this contract, which takes the form of

$$\text{E(Loss)} = E[(\lambda y^e - y)I(y \leq \lambda y^e)]p = E[(\lambda y^e - y)^+]p \tag{1.11}$$

where $a^+ = \max(0, a)$ for a number $a \in \mathbb{R}$. In the preceding discussion, y denotes the observed annual county level yield and y^e represents the predicted (guaranteed) yield. To calculate the expected loss requires estimation of the distribution of yields. The conventional two-stage estimation method will be compared to the proposed time-varying distribution in terms of calculating expected loss and premium rates.

In our simulation, one million random variables are generated from these time-varying Beta distributions. For example, under the time-varying Beta distribution

with an exponential time trend structure (model II), simulated yields are generated from $Beta(\alpha_t, \beta_t, 0, \delta_t)$ for each $t = 1927, \ldots, 2006$. The probability of yield loss, the expected yield loss, and the actuarially fair premium rate associated with a contract that guarantees 75 percent of the expected yield are calculated for each year. As shown in figure 1.4, the premium rates range from 0.83 percent in 1985 to 0.36 percent in 2006 for the case in which the yields are coming from model II. The rates change with time because all moments of the time-varying distribution are evolving. In contrast, the premium rates calculated from a conventional two-stage Beta distribution model (model I) indicate a constant premium rate around 1.88 percent from 1927 to 2006 (figure 1.4). For crop insurance in 2006, the premium rate from the detrended Beta model is 1.52 percentage points higher than the premium from the time-varying Beta model (0.36 percent versus 1.88 percent). Thus, the conventional model tends to significantly over-price for the same level of coverage.

Rate cross-validation is proposed to measure the predictive accuracy of premium rates of one model when the alternative model is true. Rate cross-validation can be tested as follows:

> Step 1: Assume one of the alternative yield distribution models, denoted by j, is true and simulate a set of actuarially fair premium rates (denoted as $r_{true_j,t}$).
>
> Step 2: Simulate 1000 sets of 80 pseudo-observations of corn yields from the corresponding true yield distribution.
>
> Step 3: Obtain 1000 sets of MLEs based on these pseudo-observations; then calculate the pseudo actuarially fair premium rates (denoted as $r_{j',t}$) based on the MLEs.

Then we can compare the pseudo premium rates with the true rates and obtain the Mean Percentage Error (MPE) and the Root Mean Squared Error(RMSE).

Rate cross-validation demonstrates a smaller MPE and RMSE for the time-varying model. As is shown in table 1.4, when the true rate is derived from the conventional model (with an average rate equal to 0.0188), the mean squared error (MSE) of predicted rates of the time-varying model is 0.0087, which is 9.58% lower than the MSE (0.0097) obtained from the conventional model when the alternative (the average premium rate implied by the time-varying model is 0.0058) is true. In addition, the MPE is 0.45 for the time-varying model and 1.66 for the detrended model. Smaller MPE and MSE values indicate that the time-varying model is more accurate, flexible, and robust in terms of rate prediction. This prediction error can also be expressed in economic terms. For example, for a crop insurance contract with \$1000 liability per acre, the rate cross-validation error of the premium is \$8.68 for the time-varying model. The rate cross-validation error of the premium is \$9.60 for the conventional model. Therefore, the predicted premium error of the time-varying model is \$0.92 less than the detrended model per unit of insurance (\$1,000 of total liability in this example). In light of the fact that the total premium in the crop insurance program in 2008 is \$9 billion, the pricing errors can result in substantial aggregate losses. Consequently, the accuracy of insurance rates is improved by applying the time-varying yield distribution model.

1.6 Conclusions

This study has examined the accuracy of alternative methods for measuring conditional yield risk under technological change over time. We propose a method for incorporating trends in the yield distribution that may offer a more accurate and consistent method for estimating the distribution of crop yields than other approaches commonly used in the literature. This method involves simultaneously estimating the time trend effects and the parameters of the yield distribution and therefore overcomes possible

shortcomings associated with the more common approach of treating the detrended yields as "observed" data rather than data estimated from a prior detrending model. This may offer improvements in the accuracy of time trend and yield distribution estimates.

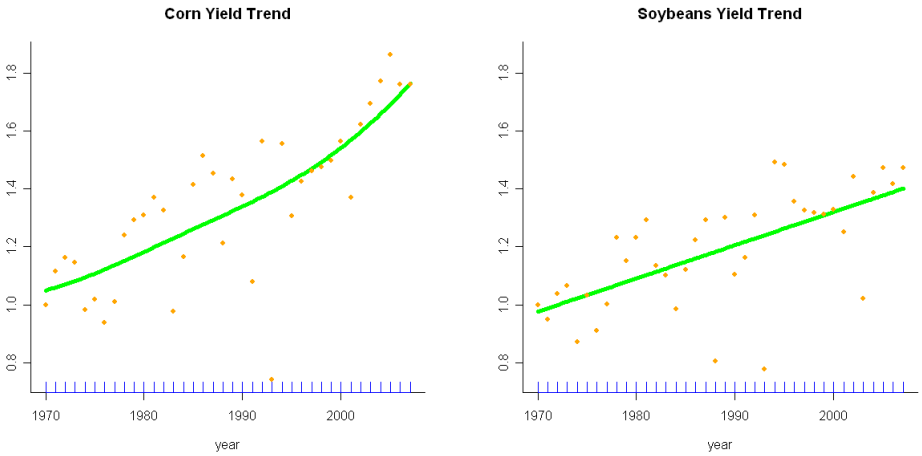
Several model selection tools are used to compare the in-sample goodness of fit and out-of-sample predictive power of the alternative models. The results show that the proposed time-varying model is superior to the conventional two-stage model in terms of providing a lower AIC and BIC and stronger out-of-sample predictive power for most of the major county/crop combinations. The results of out-of-sample prediction tests are consistent with prior expectations based on technological progress and biotechnology. In particular, multiple biotech traits and genetic improvements have occurred for corn and, to a lesser degree, for cotton. Much of the biotech innovations for soybeans have mainly involved herbicicle tolerance. The proposed time-varying method therefore works better for corn and cotton than soybeans in terms of capturing the technological effects on yield distributions.

In a rate simulation exercise, the premium rate derived from the time-varying model shows decreasing premium rates (from 0.83 percent in 1985 to 0.36 percent in 2006) over time, while the conventional model implies a constant premium rate (1.88 percent). A method of "rate cross-validation" is proposed to calculate the prediction error in terms of premium rates based on the two alternative models. This cross-validation of rates again demonstrates that the time-varying distribution model may offer significant advantages, even when the underlying yield trend process is misspecified. The time-varying model outperforms the conventional model in terms of both out-of-sample predictive power and the predictive accuracy of premium rates.

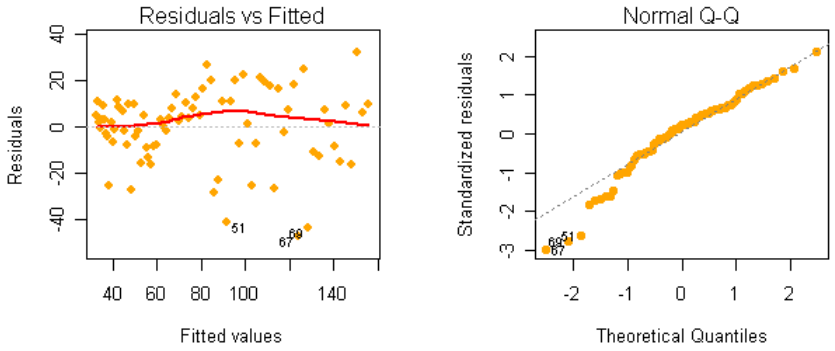
Overall, this analysis reveals a dynamic evolution of yield distributions under technological change for major U.S. crop yields. In our data, which represents county-

level yields for important crops in major growing areas, we find that the time-varying model provides a superior fit to the data. This study has policy implications that relate to improving the accuracy of assessing yield distributions in cases where parameters of the distribution evolve over time. When the distributions are changing, the premium rates can be adjusted to represent the most recent information. This offers the potential to improve the accuracy of models used in rating crop insurance contracts and thus may improve risk management mechanisms to protect producers from yield risk. The improved time-varying method has practical implications for the GRP and GRIP programs as well as the design of other insurance contracts. Our applications assume a Beta distribution for each year. Future research may benefit from relaxing this assumption by using more flexible models, such as a mixture of Beta distributions and nonparametric methods.

Figures and Tables

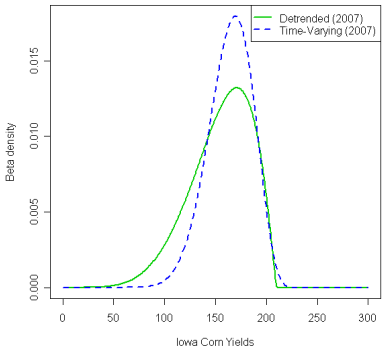


(a) Yield Trend of Different Crops (1970-2007)

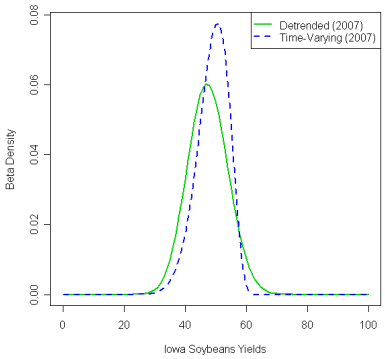


(b) Residual Plot of Annual Corn Yield, Adair County, Iowa

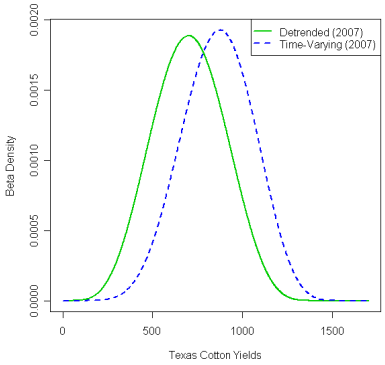
Figure 1.1: Scatter Plot and Residual Analysis



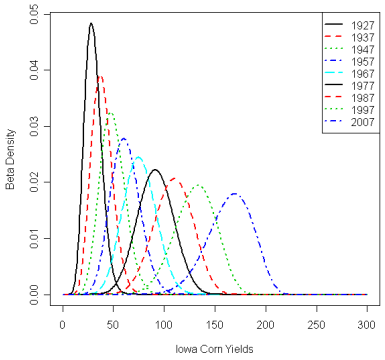
(a) Corn Yield Distribution of 2007: Detrended Beta vs. Time-Varying Beta



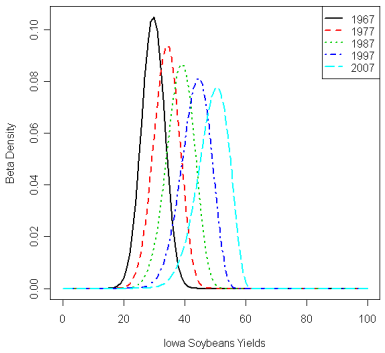
(b) Soybeans Yield Distribution of 2007: Detrended Beta vs. Time-Varying Beta



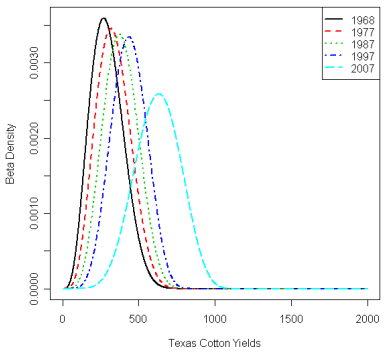
(c) Cotton Yield Distribution of 2007: Detrended Beta vs. Time-Varying Beta



(d) 10-year Overlay Beta Density Plot for Corn

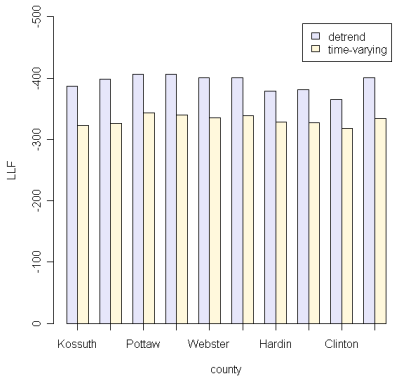


(e) 5-year Overlay Beta Density Plot for Soybeans

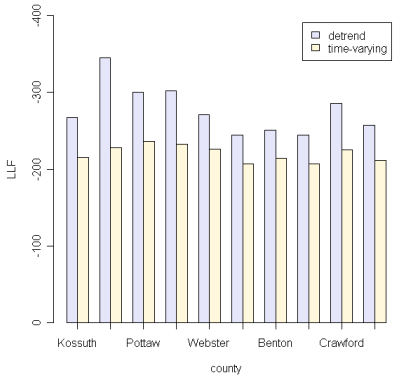


(f) 5-year Overlay Beta Density Plot for Cotton

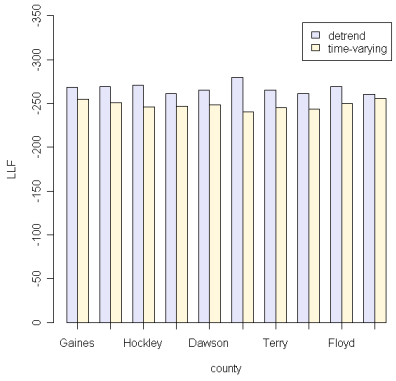
Figure 1.2: Estimated Time-Varying Beta Densities, Major Crop Yields in the U.S.



(a) LLF Comparison—Iowa Corn Yields



(b) LLF Comparison—Iowa Soybeans Yields



(c) LLF Comparison—Texas Cotton Yields

Figure 1.3: In-Sample Goodness-of-Fit Comparison of the Two Competing Models: LLF

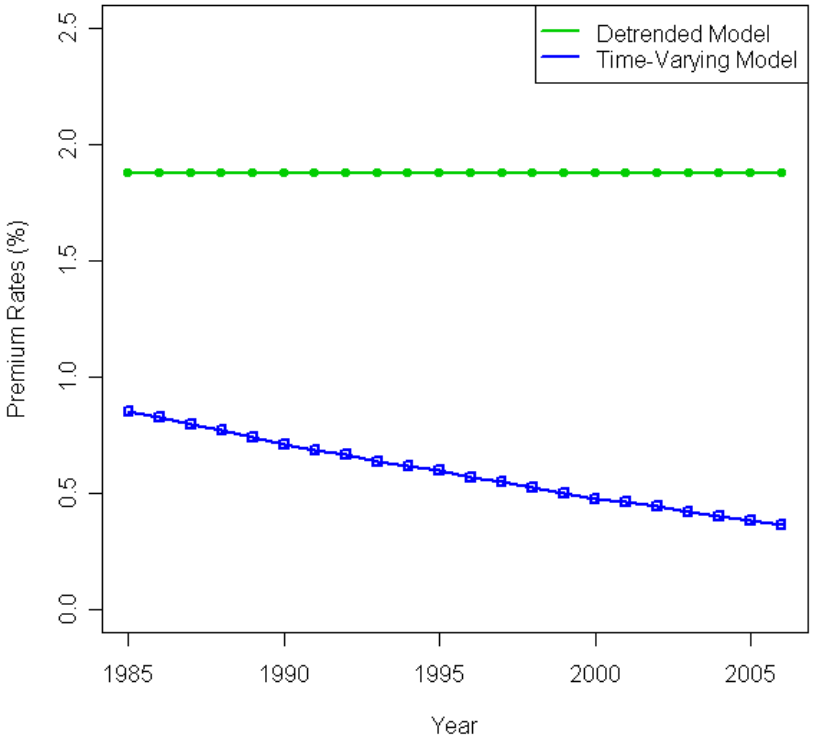


Figure 1.4: Premium Rates (for a 75% Coverage Level Crop Insurance Contract) for Time-Varying Model and Detrended Model (1985-2006)

Table 1.1: Maximum-Likelihood Parameter Estimates and Summary Statistics for Two-Stage Model and Time-Varying Models: Example for Adair County Corn Yields[c]

Two-Stage Model Based on Detrended Yield Data

Four-Parameter Beta (LLF = -378.69)			Three-Parameter Beta (LLF = -380.67)		
Parameters	**Estimates**	**Std. Error**	**Parameters**	**Estimates**	**Std. Error**
shape1(α)	5.99	0.21*	shape1(α)	5.99	0.19*
shape2(β)	2.10	0.23*	shape2(β)	2.07	0.23*
location(θ)	0.97	7.85	-	-	-
scale(δ)	203.43	1.04*	scale(δ)	204.13	1.07*

Time-Varying Models Based on Actual Yield Data

Linear Trend Structure[a] (LLF = -328.68)			Quadratic Trend Structure[b] (LLF = -326.62)		
Parameters	**Estimates**	**Std. Error**	**Parameters**	**Estimates**	**Std. Error**
b_1	2.38	0.32*	b_1	2.55	0.10*
b_2	0.43	0.75	b_2	0.16	0.40
b_3	—	—	b_3	-0.29	0.50
b_4	4.02	0.32*	b_4	2.95	0.10*
b_5	-2.71	1.29*	b_5	-1.63	0.30*
b_6	—	—	b_6	-2.61	4.8
b_7	7.47	14.99	b_7	12.26	117.70
b_8	-7.50	18.14	b_8	-15.27	138.15
b_9	—	—	b_9	-13.72	90.03

Time-Varying Models: **LLF(L): L1:** -328.68 **L2:** -326.62
LRT Statistics: $-2(L1 - L2) = 4.12$ χ^2_4 p-value = 0.39

Notes: An asterisk * denotes statistical significance at the $\alpha = 0.05$ or smaller level
[a] the Time-Varying Beta Model with a linear trend structure is defined as: $y_t \sim (\alpha_t, \beta_t, 0, \delta_t)$ $\alpha_t = exp(b_1 + b_2\tilde{t})$; $\beta_t = exp(b_4 + b_5\tilde{t})$ $\delta_t = exp(b_7 + b_8\tilde{t})$
[b] the Time-Varying Beta Model with a quadratic trend structure is defined as: $y_t \sim (\alpha_t, \beta_t, 0, \delta_t)$ $\alpha_t = exp(b_1 + b_2\tilde{t} + b_3\tilde{t^2})$; $\beta_t = exp(b_4 + b_5\tilde{t} + b_6\tilde{t^2})$; $\delta_t = exp(b_7 + b_8\tilde{t} + b_9\tilde{t^2})$
[c] Examples for other crops and counties are available from the author on request.

Table 1.2: Model Comparison Using In-Sample Goodness-of-fit Test and Non-nested Vuong's Test for Major Agricultural Yield

	Detrending Model–Model I				Time-Varying Model–Model II				
County	**K**	**LLF**	AIC/ΔAIC	BIC/ΔBIC	**K**	**LLF**	AIC/ΔAIC	BIC/ΔBIC	v^a
Iowa All-Practice Corn									
Kossuth	6	-386.90	785.80/129.17	785.28/129.17	6	-322.32	656.63/0	656.11/0	11.41*
Sioux	6	-398.085	808.17/143.62	807.65/143.62	6	-326.27	664.55/0	664.03/0	7.93*
Pottawattamie	6	-406.25	824.50/125.06	823.98/125.06	6	-343.72	699.44/0	698.92/0	12.61*
Plymouth	6	-406.47	824.94/133.44	824.43/133.44	6	-339.751	691.50/0	690.98/0	14.03*
Webster	6	-400.63	813.25/130.30	812.73/130.30	6	-335.48	682.95/0	690.98/0	13.15*
Pocohontas	6	-401.21	814.42/125.83	813.91/125.83	6	-338.30	688.60/0	688.08/0	12.95*
Hardin	6	-379.19	770.39/102.23	769.87/102.23	6	-328.08	668.16/0	667.64/0	10.76*
Franklin	6	-381.39	774.79/108.04	774.27/108.04	6	-327.37	666.75/0	666.23/0	8.55*
Clinton	6	-364.70	741.39/94.04	740.87/94.04	6	-317.67	647.35/0	646.83/0	8.99*
Woodbury	6	-401.22	814.45/133.34	813.93/133.34	6	-334.557	681.11/0	680.60/0	14.81*
Iowa Soybeans									
Kossuth	5	-267.22	544.44/102.06	544.01/102.15	6	-215.19	442.38/0	441.86/0	8.73*
Sioux	5	-345.65	701.3/233.88	700.87/233.97	6	-227.71	467.42/0	466.90/0	5.36*
Pottawattamie	5	-300.17	610.34/126.22	609.91/126.31	6	-236.06	484.12/0	483.60/0	8.27*
Plymouth	5	-302.13	614.26/136.72	613.83/136.81	6	-232.77	477.54/0	477.02/0	8.72*
Webster	5	-271.26	552.52/88.16	552.09/88.25	6	-226.18	464.36/0	463.84/0	8.79*
Woodbury	5	-244.77	499.54/73.28	499.11/73.37	6	-207.13	426.26/0	425.74/0	9.01*
Benton	5	-251.3	512.60/73.02	512.17/73.11	6	-213.79	439.58/0	439.06/0	9.57*
Grundy	5	-244.77	499.54/73.28	499.11/73.37	6	-207.13	426.26/0	425.74/0	9.01*
Crawford	5	-285.52	581.04/118.08	580.61/118.17	6	-225.48	462.96/0	462.44/0	6.07*
Tama	5	-257.38	524.76/89.44	524.33/89.53	6	-211.66	435.32/0	434.80	9.77*
Texas Upland Cotton									
Gaines	6	-268.23	548.46/27.38	547.94/27.38	6	-254.54	521.08/0	520.56/0	1.84
Lubbock	6	-269.56	551.12/38.22	550.60/38.22	6	-250.45	512.90/0	512.38/0	5.11*
Hockley	6	-270.87	553.74/50.14	553.22/50.14	6	-245.8	503.60/0	503.08/0	8.70*
Lynn	6	-261.55	535.1/29.56	534.58/29.56	6	-246.77	505.54/0	505.02/0	6.67*
Dawson	6	-264.87	541.74/32.74	541.22/32.74	6	-248.5	509/0	508.48/0	8.48*
Hale	6	-279.65	571.3/77.94	570.78/77.94	6	-240.68	493.36/0	492.84/0	2.33*
Terry	6	-264.96	541.92/40.34	541.40/40.34	6	-244.79	501.58/0	501.06/0	1.42
Crosby	6	-261.55	535.1/35.86	534.58/35.86	6	-243.62	499.24/0	498.72/0	1.29
Floyd	6	-268.92	549.84/37.26	549.32/37.26	6	-250.29	512.58/0	512.06/0	1.27
Martin	6	-260.08	532.16/8.86	531.64/8.86	6	-255.65	523.30/0	522.78/0	2.00*

Notes: An asterisk * denotes statistical significance at the $\alpha = 0.05$ or smaller level. K is the number of parameters in a model. "a" is the Vuong's test statistics for time-varying model vs. detrending model.

Table 1.3: Out-of-Sample Performance

	Detrending Model–Model I			Time-Varying Model–Model II		
County	$RMSE$	$RMSPE_2$	$RMSPE_{10}$	$RMSE$	$RMSPE_2$	$RMSPE_{10}$
Iowa All-Practice Corn						
Kossuth	211.54	493.66	512.06	210.97*	216.97*	105.27*
Sioux	182.59	256.96	554.73	188.87*	233.19*	212.06*
Pottawattamie	270.58	504.02	233.54	257.75*	477.47*	366.30*
Plymouth	242.45	378.34*	382.06*	242.32*	495.1	386.82
Webster	379.77	429.7	160.21	252.43*	313.94*	140.87*
Pocohontas	272.78*	489.72*	692.72*	275.09	492.25	851.23
Hardin	227.59	467.43	404.73	221.09*	365.9*	269.36*
Franklin	223.45	551.25*	342.77	210.33*	560.8	119.95*
Clinton	251.54	372.87*	584.10*	240.45*	383.02	708.91
Woodbury	217.84*	527.4	773.18	218.86	342.85*	259.8*
Iowa Soybeans						
Kossuth	4.12*	7.47	7.63	4.14	7.44*	7.60*
Sioux	4.18	5.13	5.67*	4.13*	4.82*	6.37
Pottawattamie	4.75	6.28	6.09*	4.73*	5.95*	6.27
Plymouth	4.74	4.73	5.06*	4.64*	4.06*	6.43
Webster	4.38	6.53*	6.36*	4.36*	6.61	6.92
Woodbury	3.74	6.09	5.82	3.69*	5.98*	5.71*
Benton	4.38	6.82	6.47	4.07*	5.61*	6.23*
Grundy	3.74	6.09	5.82	3.69*	5.98*	5.71*
Crawford	4.60	6.29	6.32	4.50*	6.12*	6.23*
Tama	3.99	6.67	6.20	3.96*	6.56*	6.18*
Texas Upland Cotton						
Gaines	130.72*	217.85	307.04*	130.90	217.23*	307.96
Lubbock	143.34	157.72	185.53	128.04*	182.61*	196.43*
Hockley	116.39	143.07*	194.23	100.13*	159.50	192.15*
Lynn	118.05	153.87	180.65	116.46*	136.78*	171.73*
Dawson	105.05*	96.13	84.69*	108.61	84.38*	163.53
Hale	155.32	187.42	239.46	113.24*	130.92*	116.32*
Terry	112.48*	174.85	277.63	129.23	133.25*	150.56*
Crosby	127.25	144.48*	153.71*	114.38*	165.91	161.32
Floyd	181.32	187.05	234.51	130.37*	158.82*	150.56*
Martin	146.23*	163.43	150.37*	148.57	153.27*	155.54

Note: an "*" indicates a smaller out-of-sample predicted error in the two competing models.

Table 1.4: Other Model Comparison Methods

Variable	Parameter Estimate	p Value[a]
Compared by Regression Method		
Intercept	-0.125	0.970
γ_1:Coefficient of Prediction Value of Detrended Beta	-0.065	0.890
γ_2:Coefficient of Prediction Value of Time-Varying Beta	1.068	0.034*

	Mean of True Rates from Conventional Model (0.01887)	Mean of True Rates from Time-varying Model (0.0058)
Rate Cross-Validation		
Root Mean Squared Error		
Conventional Predicted Rate (RMSE)	0	0.098
Time-varying Predicted Rate (RMSE)	0.093	0
Mean Percentage Error		
Conventional Predicted Rate (MPE)	0	1.66
Time-varying Predicted Rate (MPE)	0.45	0

Note: a: an "*" indicates statistical significance at the $\alpha = .10$ or smaller level.

Chapter 2

Modeling Dependence in the Design of Whole Farm Insurance Contract —A Copula-Based Approach

Abstract

The objective of this study is to model and evaluate the risks of corn and soybean production. This study focuses on the risk of revenue uncertainty that arises from variations in prices, yields shortfalls, or both. There are several parametric models for price and yield risk factors for corn and soybeans. For instance, yield risks can be modeled by a family of Beta distributions, whereas price shocks can be modeled by log-normal distributions. In order to develop a multivariate model that preserves a given set of marginal models, a copula approach can be used to characterize the joint yield and price risk of corn and soybeans, which are usually highly correlated. The copula

approach has been spurred by the recent developments in the whole farm insurance (WFI), resulting in an increasing need for the modeling of multivariate risk factors and their interaction. As a part of the study, various copula models are investigated for their suitability in modeling yield and price risks. Finally, the proposed copula approach is illustrated with simulated data to calculate the premium rate of whole farm insurance. Results show that WFI is superior to crop-specific insurance with a premium 36% cheaper than the latter.

2.1 Introduction

Federally regulated crop insurance has become a major risk management tool for U.S. farmers. This program offers protection against risks[1] in agricultural production such as yield shortfalls, price collapses, and revenue losses. Agricultural risk is typically assumed to originate from the unanticipated movements in prices, yields, and revenues. These risk factors do not exist in isolation of one another. Therefore, it is important to understand the distributions of several variables interacting simultaneously in the design and rating of crop insurance contracts. Ignoring dependencies among risk factors can lead to biased estimates of risk. For example, in the case of the natural hedge in which revenue is stabilized because of the negative relationship between crop yields and prices, if the negative relationship between price and yield is ignored, premiums will overestimate the risk of revenue insurance. Thus, it is important to be able to adequately model dependence and multivariate outcomes.

The federal crop insurance program currently insures in excess of $70 billion in crops and livestock commodities. There are 22 existing crop insurance products ranging from yield insurance to revenue insurance. The performance of the Federal Crop Insurance program is not very satisfactory. The main problem is the big loss ratio,[2] which means the federal crop insurance is losing money (Goodwin and Smith (1995), Goodwin (2001)). To guarantee an actuarially fair Federal Crop Insurance program, the calculation of an actuarially fair premium to be assigned to an insurance risk is crucial. This requires accurately estimated risk distributions. A risk management tool known as crop revenue insurance protects crop producers from declines in both crop prices and yields. Revenue insurance products currently represent approximately 80% of U.S. corn and soybean insurance acres (USDA/RMA), thus resulting in an increasing need for the modeling of

[1] The risk is deviation from the expected in terms of crop insurance.

[2] The loss ratio is defined as the ratio of indemnities paid out and the premium collected from farmers.

multivariate risk factors and their interactions in order to improve the performance of the risk management. There exist three revenue insurance (RI) programs: Crop Revenue Coverage (CRC), Revenue Assurance (RA) and Group Risk Income Protection (GRIP). These programs guarantee a certain level of crop revenue, rather than just production, for a given crop or for all insurable crops grown on a farm, and pay an indemnity if revenues fall beneath the guarantee. This indemnity payment scheme deals with both price and yield risk, and should be better correlated with a farm's need. Whole-farm insurance (WFI) is a type of revenue insurance which protects farmers against variation of multi-crop random prices and yields shocks. The rationale of WFI is thus to pool all of a farm's insurable risks into a single policy that provides coverage at lower rates against the farm's revenue losses.

The popularity of revenue insurance results in an increasing need for the modeling of multivariate risk factors and their interaction, in order to improve the performance of the risk management. One of the problems with pricing revenue insurance contracts or WFI is how to account for the dependence structure between the price and yield risks. It is difficult to calculate the premium rates for the revenue insurance when the joint distribution of two or more random variables need to be assessed. A large number of empirical and theoretical studies have addressed revenue risk by considering the degree of correlation between yields and prices. The common procedures utilized in the literature are to generate pairs of random variables from a given pair of marginals with a known degree of correlation. Babcock and Hennessy (1996) illustrate how to model and rate the revenue risk by using the Johnson and Tenenbein (1981) procedure to simulate correlated prices and yields from mixed marginal distributions. Hart, Hayes, and Babcock (2006) present how to rate the whole-farm insurance by using the Iman and Conover (IC) (1982) procedure. Cobel, Heifner, and Zuniga (2000) have investigated these correlations in revenue insurance and found that there is strong correlation between price and farm

yield and between farm and national yield in certain crops and regions. All these procedures are based on the simulation from a given correlation matrix. Spearman's rank correlation coefficient and Kendall's tau coefficient are two common measures of correlation which are invariant to monotone transformations. However, these procedures are limited in that only the first and second moments are matched (Goodwin and Ker, 2001). Also, these correlation structure measures characterize the dependence over the entire support of the risk variables, while many issues in risk management focus on the tail area behavior. Besides the above limitations, theoretical models of risk management imply that crop-specific revenue insurance and whole-farm insurance should be derived from the joint probability distribution of yields and prices of one crop or multiple crops. To know the joint loss probability, the joint distribution of crop yields and prices must be known. A joint probability distribution uniquely determines all lower dimensional marginal and conditional distributions. As a result, association between any set of random variables can be fully described by knowing the multivariate joint distribution. This multidimensional composition of several random variables calls for flexible models that can describe the major data properties, such as higher-order moments, fat tails, co-extreme movement and tail dependence.

The need to model the multivariate distribution of yields and prices motivated this study to use copula approach. The copula, which was originated by Sklar (1959), is a promising tool to study multivariate outcomes. Most of the related work in copulas was written after 1990. The basic idea of a copula is to link the marginal distributions together to form the joint distribution. It is a parametrically-specified, joint distribution generated from given marginals. The attractive feature of parametrically specified copulas is that estimation and inference is based on standard maximum likelihood procedures, which make it feasible to efficiently estimate the assumed copula model. Admittedly such efficiencies are lost when the true data generating mechanism is

significantly different from the assumed model.

The objective of this study is to model and evaluate multivariate risk factors (yield shortfalls and price shocks) and their interaction in corn and soybean production, with implications for evaluating the potential for establishing a combination insurance or whole-farm revenue insurance that would address the risks of corn and soybean production. The multivariate modeling is based on the copula approach to reconstruct a multivariate joint distribution of the risks based on a given set of yield and price marginal distributions, as well as to separate the dependence structure and the marginal distributions in a multivariate distribution. As a part of this study, various parametric copula models such as the Gaussian copula and t copula are investigated for their suitability in modeling yield and price risks.

This study focuses on the risk of whole-farm revenue variability that arises from changes in multi-crop random prices, random yields shortfalls or both. With whole farm coverage, all acres of a farm's crops, rather than an individual crop, are covered under one insurance unit. Since most crop risks are not perfectly correlated, whole farm insurance (WFI) provides more efficient coverage than insuring each crop with a specific policy. If a farm grows two crops, corn and soybeans, an insurance policy based on the farm's total revenue will be cheaper than the sum of the premium for corn and soybeans for the same expected revenue (Hennessy et al. 1997). The effects of an increase in dependence between multiplicative risks on the actuarially fair premium value of an insurance contract are analyzed in this study. The gains of moving from a combination revenue insurance delivered by crop-specific policies to WFI are also evaluated in this study.

Based on the copula WFI model approach, the rest of the paper is organized as follows. Section 2 explores the copula approach for WFI modeling by using a two-stage MLE method. Section 3 provides the empirical analysis of the copula model by using

corn and soybean data. A simulation study is conducted in Section 4 to compute the premium rate of the WFI and crop-specific insurance based on the copula estimates from Section 3. The conclusion is that premium of WFI are 36% cheaper at a 75% coverage level.

2.2 Copula Based Models for Dependent Random Variables

For any continuous multi-dimensional dependent random variables, a copula can be used to develop a multivariate model. The copula is a tool for understanding relationships among multivariate random variables. The usefulness of copulas in modeling dependence stems from a famous theorem of Sklar (1959).

Sklar's theorem states that any continuous multivariate distribution can be uniquely described by two parts: the univariate marginals and the multivariate dependence structure. The latter is represented by a copula. A m-dimensional copula $C(u_1, u_2, ..., u_m)$ is defined as any multivariate distribution function in the unit hypercube $[0, 1]^m$, with uniform $U[0, 1]$ marginal distributions. It can be shown (see Sklar, 1959) that any absolutely continuous joint distribution function $F(x_1, x_2, ..., x_m)$, whose marginal distribution functions are $F_1(x_1), F_2(x_2), ..., F_m(x_m)$, can be expressed as:

$$F(x_1, ..., x_m) = C(F_1(x_1), ..., F_m(x_m)) \tag{2.1}$$

where $C(\cdot)$ is a suitable copula function as defined above. Often the copula function is modeled as $C(u_1, \ldots, u_m; \theta)$, where θ is a vector of parameters of the copula which are called the *dependence parameters*, which measure dependence between the marginals.

Sklar's theorem implies that copulas can be used to express a multivariate distribution in terms of its marginal distributions. The copula $C(\cdot)$ is unique for a

given distribution $F(\cdot)$ if each marginal $F_j(x_j)$ is absolutely continuous. That is, the joint distribution of the random variables $X_1, \ldots, X_m$ can be expressed in terms of its respective marginal distributions $F_j(x_j)$ and the copula function $C(\cdot)$ that binds them together. This construction allows researchers to consider marginal distributions and dependence as two separate but related issues. As a result, joint distributions can be estimated with a two-step process. The first step is to estimate appropriate marginal distributions of each random variable. A substantial advantage of copula functions is that the marginal distributions may come from different families. Therefore, the copula construction does not constrain the choice of marginal distributions. The second step is to estimate a copula by using one of the copula models by estimating the finite dimensional parameters θ of the copula function, given the estimated marginal distributions from the first step. For empirical applications, the dependence parameter is the main focus of estimation. Alternatively, one can jointly estimate the marginals and the copula parameters using a maximum likelihood method. Equation (2.1) is a frequent starting point of empirical applications of copulas. In this equation, the joint distribution is expressed in terms of its respective marginal distributions and a function $C(\cdot)$ that binds them together. For the empirical applications, the dependence parameter θ is the main focus of estimation. This equation will be used in fitting copula-based models to agricultural data in the design of whole farm insurance contracts.

Once the marginal distributions are specified, an appropriate copula needs to be selected. Since copulas separate marginal distributions from dependence structures, the appropriate copula for a particular application is the one which best captures the dependence features of the data. A large number of parametric families of copulas are available in the literature (Nelson, 1999). The most common copulas in risk management are the "t copula" and the "Gaussian (normal) copula."

Gaussian (Normal) Copula

The Gaussian copula takes the form

$$C(u_1, \ldots, u_m; \theta) = \Phi_m(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_m); \boldsymbol{\theta}) \tag{2.2}$$

where Φ is the distribution function of the univariate standard normal distribution, Φ^{-1} is the inverse distribution function of the standard normal distribution and $\Phi_m(x_1, \ldots, x_m; \theta)$ is the multivariate normal distribution with mean zero and correlation matrix $\boldsymbol{\theta}$.

Student's t-copula

The t copula with γ degrees of freedom and correlation matrix $\boldsymbol{\theta}$ is

$$C(u_1, \ldots, u_m; \boldsymbol{\theta}) = T_{m\gamma}(T_{1\gamma}^{-1}(u_1), \ldots, T_{1\gamma}^{-1}(u_m); \boldsymbol{\theta}) \tag{2.3}$$

where $T_{1\gamma}$ is the distribution function of a univariate t-distribution with γ degrees of freedom and $T_{m\gamma}(x_1, \ldots, x_m; \boldsymbol{\theta})$ denotes the distribution function of a multivariate t-distribution with γ degrees of freedom. The two dependence parameters are $(\gamma, \boldsymbol{\theta})$. As $\gamma \to \infty$, the t copula approaches Gaussian copula.

Given a collection of marginal densities, the multivariate distribution can be defined by applying a copula to the prescribed marginals by using equation 2.1. There are several parametric models for price and yield risk factors for corn and soybeans. For instance, yield risks are usually modeled by a family of Beta distributions, whereas price shocks are usually modeled by log-normal distributions. Copulas allow us to piece together joint distributions of yield risks and price shocks when only parametric forms of the marginal distributions are known explicitly.

The density function of a multivariate distribution whose dependence structures is defined by a copula function can be obtained by differentiation of equation (1) and is given by

$$f(X_1, \ldots, X_m, \boldsymbol{\theta}) = c(F_1(X_1), \ldots, F_m(X_m)) \prod_{j=1}^{m} f_j(X_j) \tag{2.4}$$

where $f_j(x) = F_j^{'}(x)$ represents the marginal density function of X_j and $c(\cdot)$ is the density function of the copula function C. In other words,

$$c(u_1, \ldots, u_m) = \frac{\partial^m C(u_1, \ldots, u_m)}{\partial u_1, \ldots, \partial u_m} \tag{2.5}$$

where $u_j \in [0, 1], j = 1, \ldots, m$.

Suppose that the marginal distribution of X_j is modeled by a parametric family $F_j(x, \beta_j)$ where β_j is the vector of parameters for the $j - th$ marginal for $j = 1, \ldots, m$. Then the log likelihood function of $(\beta_1, \ldots, \beta_m, \boldsymbol{\theta})^T$ is given by

$$\begin{aligned} l(\beta_1, \ldots, \beta_m, \boldsymbol{\theta}) &= \sum_{i=1}^{n} log(c(F_1(X_{i1}; \beta_1), \ldots, F_m(X_{im}; \beta_m); \boldsymbol{\theta})) \\ &+ \sum_{i=1}^{n} \sum_{j=1}^{m} log(f_j(X_{ij}; \beta_j)) \end{aligned} \tag{2.6}$$

where $\boldsymbol{\theta}$ denotes the copula parameter.

2.3 Estimation Methods

2.3.1 Parametric Maximum Likelihood Method

The usual methodology used in the literature, as proposed by Joe and Xu (1996), requires one to find the MLE of the appropriate marginal in a two-stage approach. This method is called the inference function for margins (IFM) method. By using this method, the set of parameters of the model are estimated through a system of estimating equations, with each estimating equation being a score function (partial derivative of a log-likelihood) from some marginal distribution of the multivariate copula model. In the first step of the IFM method, the marginal parameter β_j is estimated by maximizing the loglikelihood of the m univariate margins separately

$$\hat{\beta}_{j,IFM} = \arg\max_{\beta_j} \sum_{i=1}^{n} \log(f_j(X_{ij}; \beta_j)) \tag{2.7}$$

where $j = 1, 2, \ldots, m$ stand for parameters from each marginal distribution F_j. In the second step, the copula form is chosen and the dependence parameter θ is estimated given the marginal estimates $\hat{\beta}_{IFM}$:

$$\hat{\boldsymbol{\theta}}_{IFM} = \arg\max_{\theta} \sum_{i=1}^{n} \text{logc}(F_1(X_{i1}; \hat{\beta}_{1,IFM}), \ldots, F_m(X_{im}; \hat{\beta}_{m,IFM}); \boldsymbol{\theta}) \tag{2.8}$$

An alternative method would be maximize the joint loglikelihood of $(\beta_1, \ldots, \beta_m, \boldsymbol{\theta})$, i.e.,

$$(\hat{\beta}_1, \ldots, \hat{\beta}_m, \hat{\boldsymbol{\theta}}) = \arg \max_{\beta_1, \ldots, \beta_m, \theta} l(\beta_1, \ldots, \beta_m, \boldsymbol{\theta}) \tag{2.9}$$

It may be noted that $l(\hat{\beta}_1, \ldots, \hat{\beta}_m, \hat{\boldsymbol{\theta}}) \geq l(\hat{\beta}_{1IFM}, \ldots, \hat{\beta}_{mIFM}, \hat{\boldsymbol{\theta}}_{IFM})$ and hence the two-stage method is sub-optimal unless $\hat{\beta}_j = \hat{\beta}_{jIFM}$ for $j = 1, \ldots, m$. In that case $\hat{\theta} = \hat{\theta}_{IFM}$.

2.3.2 Semi-Parametric Pseudo Maximum Likelihood Method

A semi-parametric method is carried out in this study to estimate the joint distribution of crop price and yield risks by evaluating a parametric copula function at their nonparametric kernel-based marginal distributions. This method estimates the kernel estimate of the marginal in the first step and estimates the dependence parameters in the second step.

To evaluate the multivariate distribution $F(x_1, x_2, ..., x_m)$
$= C(F_1(x_1), F_2(x_2), ..., F_m(x_m); \boldsymbol{\theta})$, one needs to estimate both the copula parameter θ and the marginal cdf's $F_j(x)$. Suppose $X_{ij}, i = 1, 2, ..., n, j = 1, 2, ..., m$ are observed, where j stands for the dimension of the data (the number of variables) and i stands for the number of the observation. $X_{1j}, ..., X_{nj}$ can be used to obtain a kernel density estimate (KDE) of $f_j(x) = F_j'(x)$ as follows.

$$\begin{aligned}
\hat{f}_1(x) &= \frac{1}{n}\sum_{i=1}^{n} K(\frac{x - X_{i1}}{\lambda_1})\frac{1}{\lambda_1} \\
\hat{f}_2(x) &= \frac{1}{n}\sum_{i=1}^{n} K(\frac{x - X_{i2}}{\lambda_2})\frac{1}{\lambda_2} \\
&\vdots \\
\hat{f}_m(x) &= \frac{1}{n}\sum_{i=1}^{n} K(\frac{x - X_{im}}{\lambda_m})\frac{1}{\lambda_m}
\end{aligned}$$

where λ_j $(j = 1, 2, \ldots, m)$ denotes the bandwidth for the jth variable. Now choose a copula model, say Gaussian or t copula as an example, $C(u_1, ..., u_m, \boldsymbol{\theta})$, and compute

$$c(u_1, ..., u_m, \boldsymbol{\theta}) = \frac{\partial^m C(u_1, ..., u_m, \boldsymbol{\theta})}{\partial u_1 ... \partial u_m}.$$

Compute the likelihood

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} c(\hat{F}_1(x_{i1}), ..., \hat{F}_m(x_{im}), \boldsymbol{\theta}),$$

where

$$\hat{F}_j(x) = \int_{-\infty}^{x} \hat{f}_j(t)dt = \frac{1}{n\lambda_j} \sum_{i=1}^{n} \int_{-\infty}^{x} K(\frac{s - X_{ij}}{\lambda_j})ds$$

and $K(u)$ can be chosen as Epanechnikov Kernel: $K(u) = \frac{3}{4}(1 - u^2)I(|u| < 1)$.[3] The CDF of this marginal is shown in the following.

$$\begin{aligned} \hat{F}_j(x) &= \frac{1}{n\lambda_j} \sum_{i=1}^{n} \int_{-\infty}^{x} K(\frac{s - X_{ij}}{\lambda_j})ds \\ &= \frac{1}{4n} \sum_{i=1}^{n} ((\frac{3(x - X_{ij})}{\lambda_j}) - (\frac{x - X_{ij}}{\lambda_j})^3 + 2), j = 1, \ldots, m. \end{aligned}$$

By subsituting $\hat{F}_j(x)$ back to $L(\theta)$ with the copula being a Gaussian or t copula, the dependence parameter $\hat{\theta}_{MLE}$ can be found by maximizing $L(\theta)$.

A kernel estimator assumes that the values of a smooth function stay approximately constant in a neighborhood. The bandwidth (λ) is the range of data that are effectively used when estimating at point x_{ij}. It controls the neighborhood size and thus the number of points in the estimator. As the neighborhood size increases, more points fall into the neighborhood and the estimated function would be smoother. The smoothing parameter determines the shape of the estimated density. Thus, it is critical to choose an appropriate value of smoothing parameter in the nonparametric analysis.

[3] In the discussion of choosing optimal kernel, it is shown that Epanechnikov kernel is optimal in terms of asymptotic MSE for Priestley-Chao estimator

2.4 Empirical Framework—A Copula-Based Whole Farm Insurance Model

The theory of finance states that idiosyncratic risk can be reduced by diversification. This idea can be applied to design Whole Farm Insurance (WFI). Under a Whole Farm Insurance (WFI) contract, a farmer purchases the insurance for all his acres under one insurance contract. The WFI provides risk diversification across different crops. In the WFI model, gross revenue is exposed to multiple risks, such as random yields and random prices for multiple crops, which may affect each other. It is important to model the distributions of these multivariate risk variables simultaneously by using a joint distribution model, which can capture the dependence structure between these risk factors. To model the risk of whole-farm revenue (WFI) variability that arises from changes in multi-crop prices, yields shortfalls or both, the copula method can be applied. Suppose a farm grows two crops, corn and soybeans, an insurance contract based on the farm's total revenue and the sum of the two separate premia of corn and soybeans for the same expected revenue can be compared based on a four-dimensional copula and two two-dimensional copulas, respectively.

2.4.1 Modeling Dependence in the Whole Farm Insurance—Price-Yield Dependence and Cross Price-Yield Dependence

The WFI contract has the standard indemnity scheme of the form

$$max[(\lambda R^e - R), 0]$$

where $R = Y_c P_c + Y_s P_s$ is the observed annual total revenue from corn and soybeans, $R^e = E(R)$ is the expected revenue and $\lambda \in (0, 1]$is the coverage level. P_i and $Y_i, i = c, s$ are the non-negative random variables representing price and yield risks of corn and

soybeans, respectively, with the joint probability distribution F defined by a certain copula function form

$$F(P_c, P_s, Y_c, Y_s) = C(F_1(P_c), F_2(P_s), F_3(Y_c), F_4(Y_s); \boldsymbol{\theta}). \tag{2.10}$$

If $R < \lambda R^e$, the insurer will pay the farmer $(\lambda R^e - R)$ as an indemnity. An actuarially fair premium for the WFI is equal to the expected loss of this contract. Suppose an insurance contract will insure some proportion λ of the mean crop revenue (R^e); the expected loss (in bushels) for this insurance contract that guarantees $\lambda \times 100\%$ of the predicted revenue (R^e) takes the form of

$$\mathrm{EL}(R) = E[(\lambda R^e - R) I(R \leq \lambda R^e)].$$

The expected value of the revenue risk R^e is with respect to the joint distribution of risk measures of price and yield risk. Premium calculation of a WFI contract thus requires knowing the joint distribution of $F(P_c, P_s, Y_c, Y_s)$. To construct the full joint distribution of yield and price risk, we have used a suitable class of copula functions.

We illustrate the copula modeling of WFI revenue risk by using four-dimensional copula models. The data used in this analysis are historical crop yield and price data. The yield data are obtained from the National Agricultural Statistics Service (NASS) on annual county-level corn and soybean yields of Iowa during the period of 1960 to 2006. The price data are the difference between the futures prices at harvest and at planting for corn and soybeans from the Chicago Board of Trade (CBOT) for the same time period.

2.4.2 Data Description

Current revenue insurance programs use CBOT futures market prices and the historical county average of the crop yields to compute the revenue coverage and guarantee. The CBOT's December corn futures contract and November soybean futures contract data from 1960 to 2006 are considered in our analysis.

A base market price is determined by averaging the new-crop futures prices during planting time (February) for both December corn futures contract and November soybeans futures contract. A delivery price is determined by averaging the new crop futures prices during delivery time for both corn (in December) and soybeans (in November). The data used in this study are February (planting time) corn futures price data (defined as $P_{c,t}^{2,12}$) and December (harvest time) corn futures price data (defined as $P_{c,t}^{12,12}$) of each year from 1960 to 2006; February soybean futures price data (defined as $P_{s,t}^{2,11}$) and November (harvest time) soybean futures price data (defined as $P_{s,t}^{11,11}$) of each year from 1960 to 2006.

By regressing the harvest-time futures price on the planting-time predicted futures price for both corn and soybeans, the estimates of the intercept and slope parameter can be obtained from the models $\log(P_{c,t}^{12,12}) = \gamma_{1c} + \gamma_{2c} * \log(P_{c,t}^{2,12}) + e_{c,t}$, and $\log(P_{s,t}^{11,11}) = \gamma_{1s} + \gamma_{2s} * \log(P_{s,t}^{2,11}) + e_{s,t}$. As shown in table 2.1, the estimated intercept γ_{1i} is not significantly different from 0 and the estimated slope γ_{2i} is close to 1 for both regressions for $i = c, s$. Thus, the logarithmic price shocks (log returns) for corn (defined as $\tilde{P}_{c,t}$) and soybeans (defined as $\tilde{P}_{s,t}$) are approximately equal to the difference between the logarithmic prices in harvest time and in planting time. That is, $\tilde{P}_{c,t} = \log(P_{c,t}^{12,12}) - \log(P_{c,t}^{2,12})$, $\tilde{P}_{s,t} = \log(P_{s,t}^{11,11}) - \log(P_{s,t}^{2,11})$. It is most often a common assumption that the price is distributed as log-normal. Therefore, the logarithmic price shocks ($\tilde{P}_{c,t}$ and $\tilde{P}_{s,t}$) are assumed to be normally distributed. The goodness-of-fit tests for normal distribution based on the data support this assumption. The summary statistics for the

price shocks are shown in table 2.2.

The yield data used here are the county-level annual yields data for corn and soybeans from 1960 to 2006 in Adair County, Iowa. By assuming that errors are proportional to the predicted mean, a proportional adjustment can be applied to obtain the detrended yield shocks. This detrending method is ad hoc but it is very common in practice. To do so, we regress the yields on a quadratic time trend and output the residuals using the model. $Y_t = \eta_1 + \eta_2 t + \eta_3 t^2 + e_t$, where Y_t is the observed crop yield data in year t. Then, we calculate the residual $\hat{e}_t$ and the predicted yield for 2006 as $\hat{Y}_{2006}$, the normalized and detrended data will be given by the following

$$\tilde{Y}_t = \hat{Y}_{2006}(1 + \frac{\hat{e}_t}{\hat{Y}_t}).$$

where $\hat{e}_t$ is the residual for time t by regressing the yields on a quadratic time trend, and $\hat{Y}_{2006}$ is the predicted yield for 2006.

2.4.3 Estimation Results—Fitting 4-dimensional Copula for WFI Model

By using the data on price and yield shock of corn and soybeans ($\tilde{p}_c$,$\tilde{p}_s$,$\tilde{y}_c$,$\tilde{y}_s$), the joint distribution of prices and yields shocks for corn and soybeans in any given year t can be estimated by using a 4-dimensional Gaussian or t copula (as defined in Section 2) with the correlation matrix

$$\Sigma = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ & 1 & \rho_{23} & \rho_{24} \\ & & 1 & \rho_{34} \\ & & & 1 \end{bmatrix}$$

which is an unstructured positive definite matrix and ρ_{ij}'s $(i < j)$ are dispersion parameters to be estimated. Since one of the good properties of copulas is invariance to monotonic transformations of the marginals, the dispersion matrix determines the dependence structure of the joint distribution of ($\tilde{p}_c$,$\tilde{p}_s$,$\tilde{y}_c$,$\tilde{y}_s$).

We use the two-step copula model fitting method. The initial step is to determine the appropriate marginals. The Chi-square goodness-of-fit statistics based on the data suggest that the yield marginals are Beta distributions and the logarithmic price shocks are distributed Normally. The Maximum Likelihood Estimate for corn logarithmic price shocks $\tilde{P}_c \sim N(-0.03, 0.20)$ and soybean logarithmic price shocks $\tilde{P}_s \sim N(0.02, 0.16)$ respectively, where $N(\mu, \sigma)$ denotes a normal distribution with mean μ and standard deviation σ. The marginal MLEs of yields shocks are estimated with corn yields $\tilde{y}_c \sim Beta(7.01, 2.09, 203.55)$, soybean yield $\tilde{y}_s \sim Beta(17.60, 7.66, 65.60)$, where $Beta(a, b, \delta)$ denotes a generalized Beta distribution[4] with shape parameter a and b, and scale parameter δ. In the second step, the dependence parameters for corn and soybeans are estimated given the marginal estimates from the first step. The estimated results of the 4-dimensional Gaussian copula ($LLF = 34.43631$) and t copula ($LLF = 36.5453$) are shown in tables 2.4 and 2.6. The corresponding correlation matrices for $(\tilde{p}_c, \tilde{p}_s, \tilde{y}_c, \tilde{y}_s)$ implied by the t and Gaussian copula with Normal and Beta marginals are shown as tables 2.5 and 2.7.

2.5 Copula Selection

Tables 2.5 and 2.7 show that Gaussian copula and t copula exhibit similar dependence structures among price and yield shocks of corn and soybeans. The within-crop and cross-crop correlations of price and yield are negative, while the cross-crop price correlation (0.73 in Gaussian copula and 0.74 in t copula) and cross-crop yield correlation

[4]The location parameter of this generalized Beta is assumed to be zero.

(0.68 in Gaussian copula and 0.71 in t copula) are positive. The magnitude of the negative correlations of prices and yields in t copula is higher, which means that the t copula implies higher overall dependence than the Gaussian copula.

The copula-based revenue risk model can be fitted by using other types of Copula functions. Some criteria need to be evaluated in order to select the "best" copula model among a given finite set of models. Some typical model selection methods can be used such as Akaike's Information Criterion (AIC) and the Bayesian Information Criterion of Schwarz (BIC). Various parametric copula models estimated in this section such as Gaussian copula and t copula, can be investigated for their suitability in modeling yield, price and revenue risks. From the estimation, the AIC of the t copula ($AIC = -59.1$) is slightly smaller than the AIC of the Gaussian copula ($AIC = -56.86$) and the log-likelihood value of the t copula is greater than the log-likelihood value of the Gaussian copula, which implies better goodness-of-fit of t copula than the Gaussian copula.

2.6 Simulation Study and Policy Implication

In this section, a Monte Carlo simulation method is used to simulate the 4-dimensional variables of price and yield shocks of corn and soybeans from the copula function estimated above, which preserves the rank correlations of these variables. The joint revenue loss and the premium rates of the revenue insurance contract can be obtained at certain coverage levels. Several scenarios are conducted in which gross crop revenue is guaranteed at the 75 and 85 percent level, respectively. The mean and variance of expected revenue, the expected loss, and the actuarially-fair premium rates for a whole-farm revenue insurance policy that guarantees a certain percent of the expected level of revenue are calculated.

Suppose that a representative Iowa farmer in Adair County grows corn and soybeans

in an equal proportion. The farm's guaranteed yields for crop insurance purposes for the two crops are equal to county trend yields. The dependence structure is imposed by the 4-dimensional copula as estimated in Section 3.3. One million revenue series were drawn from the 4-dimensional copulas which preserved the dependence structure among prices and yields of corn and soybean as well as the lognormal price shocks and Beta yields shocks. Those draws $(\tilde{p}_c, \tilde{p}_s, \tilde{y}_c, \tilde{y}_s)$ are used to calculate the expected loss and premium rates for the revenue insurance contract at different coverage levels at year 2006. The predicted harvest delivery price of corn is $\hat{P}^{12,12}_{c,2006} = exp(\tilde{P}_{c,2006} + log(P^{2,12}_{c,2006}))$, in which one million of realized predicted $\hat{P}^{12,12}_{c,2006}$ can be obtained. There are one million of predicted corn yields $\tilde{y}_{c,2006}$ of 2006 as well. The predicted revenue realization of corn is calculated as $\hat{R}_{c,2006} = \hat{P}^{12,12}_{c,2006} * \tilde{y}_{c,2006}$. Likewise, the predicted revenue realization of soybeans can be calculated as $\hat{R}_{s,2006} = \hat{P}^{11,11}_{s,2006} * \tilde{y}_{s,2006}$. Based on the simulated predicted revenue, the expected revenue loss and premium rate for crop-specific insurance contract and whole farm insurance contract can be calculated, as shown in tables 2.8 through 2.10.

Table 2.8 shows that the actuarially fair premium value of a WFI contract implied from a Gaussian copula at 75% coverage is \$4.44, while the premium implied from a t copula at the same coverage is \$3.63 as shown in table 2.9. This indicates that higher magnitude of dependence structure implied by a t copula tends to result in lower premium than that from a Gaussian copula given the same level of revenue insurance coverage (\$511 per unit of insurance).

Tables 2.8 and 2.9 show that the actuarially fair premium value of crop-specific revenue insurance and the WFI implied by a Gaussian copula and a t copula at 85% coverage. The results are consistent with those at 75% coverage. Table 2.10 presents actuarially fair crop-specific and whole-farm revenue assurance premiums for different types of copulas and for different coverage levels. Let D be the difference between crop

specific premiums and the whole-farm premium. The difference D ranges from \$2.46 in the Gaussian copula and \$2.06 in the t copula at 75% coverage, which implies the premium of WFI is 36% cheaper than the sum of the premiums of two separate crop-specific revenue insurance. At 85% coverage level, the premium of WFI is 39% cheaper than the combination of the two crop-specific revenue insurance plans, which shows that the portifolio effect increases as the coverage level increases. It also indicates a big difference between crop-specific coverage and whole-farm coverage revenue insurance. These results suggest that there is a reduction in the actuarially fair premium for whole-farm revenue insurance relative to single crop-specific insurance. The risk reduction is driven by portifolio effects of pooling across crops if a farmer compares single crop-specific and whole-farm crop insurance at the same coverage level for all crops. The WFI on farmer's gross revenue is less costly than the sum of crop-specific insurance policies at the same level of insurance coverage. Therefore, it is cheaper to provide coverage to the combined crops rather than to insure each revenue separately.

2.7 Conclusion

This study has taken a close look at the efficiency of whole-farm insurance when a corn and soybean producer faces joint yields and price risk by using a copula model approach. Systemic price and yield risks of different crops have been studied as it is very important in crop revenue insurance. The empirical analysis in this study shows that the whole-farm contracts are more attractive as a risk management tool than the combination of the crop-specific contracts since the whole-farm contracts recognize the portfolio effects of different risk factors in a single insurance contract.

This study contributes to the assessment of the new whole-farm revenue insurance which provides overall coverage to all farms' crops. The idea of whole-farm insurance is

to pool all of a farm's insurable risks into a single insurance policy that provides cheaper premium rates at the same protection level against the gross farm revenue losses. The results in this analysis support this idea. Simulation results indicate that a WFI scheme that guarantees a certain expected revenue to producers could provide approximately the same level of coverage as the sum of crop-specific programs at as little as 64% ($3.63/$5.69) the cost at 75% coverage level, and at as little as 76% the cost at 85% coverage level. This shows that the accuracy and efficiency of modeling cross-crop yield and price correlations and rating of whole-farm insurance contracts are improved by using this copula approach. The result will have implications for accurately pricing whole farm insurance products which protect farmers from gross revenue losses. These results are also crucial for conducting better risk management program to producers from the whole-farm crop revenue risks.

Tables and Figures

Table 2.1: Equation Estimates: Corn and Soybean Price shocks

Variable	**Parameter Estimate**	**Standard Error**	t **Ratio**[a]
..... Regression with intercept			
Intercept(logarithmic Corn Futures Price)	0.84	0.42	2.01
$\log(P_{c,t}^{2,12})$	0.84	0.07	10.68*
Intercept(logarithmic Soybean Futures Price)	0.69	0.35	1.99
$\log(P_{s,t}^{2,11})$	0.89	0.06	15.87*
..... Regression without intercept			
$\log(P_{c,t}^{2,12})$	0.99	0.01	180.96 *
$\log(P_{s,t}^{2,11})$	1.003	0.003	262.17 *

[a] An "*" indicates statistical significance at the $\alpha = .05$ or smaller level.

Table 2.2: Variable Descriptions and Summary Statistics[a] (1960-2006)

Variable	**Description**	**Mean**[b]	**Standard Deviation**
$P_{c,t}^{2,12}$	Predicted December Futures Price of Corn at February planting time of year t	2.22	0.71
$P_{c,t}^{12,12}$	Realized December Futures Price of Corn of year t	2.14	0.72
$P_{s,t}^{2,11}$	Predicted November Futures Price of Soybeans at February planting time of year t	5.12	1.78
$P_{s,t}^{11,11}$	Realized November Futures Price of Corn of year t	5.20	1.81
$\tilde{P}_{c,t}$	Logarithm Price shocks for corn	-0.034	0.20
$\tilde{P}_{s,t}$	Logarithm Price shocks for soybeans	0.02	0.16

[a]Data source: Chicago Board of Trade (CBOT). Number of observation is 47.
Year $t = 1960, \ldots, 2006$.
[b]The units of the futures price are dollars/bushel (5000 bushels per contract)

Table 2.3: Equation Estimates: Detrending Corn Yield

Variable	**Parameter Estimate**	**Standard Error**	t **Ratio**[a]
.....	Adair County, Iowa (Fips = 19001)	.....	
Intercept	32.03	5.47	5.85*
t	0.47	0.31	1.52
t^2	0.012	0.004	3.43*

[a]An "*" indicates statistical significance at the $\alpha = .10$ or smaller level.

Table 2.4: The Estimates of the Dependence Parameters of the Gaussian Copula

Gussian	**Estimate**	**Std. Error**	**P-value**
ρ_{12}	0.73	0.055	< 0.001
ρ_{13}	-0.16	0.14	0.25
ρ_{14}	-0.17	0.14	0.23
ρ_{23}	-0.27	0.13	0.04
ρ_{24}	-0.29	0.13	0.03
ρ_{34}	0.68	0.065	< 0.001

Table 2.5: The Dependence Structure Implied by the Gaussian Copula

	Corn Price	**Soybean Price**	**Corn Yield**	**Soybean Yield**
Corn Price	1	0.73	-0.16	-0.17
Soybean Price		1	-0.27	-0.29
Corn Yield			1	0.68
Soybean Yield				1

Table 2.6: The Estimates of the Dependence Parameters of the t Copula

t	**Estimate**	**Std. Error**	**P-value**
ρ_{12}	0.74	0.062	< 0.001
ρ_{13}	-0.31	0.15	0.042
ρ_{14}	-0.29	0.15	0.06
ρ_{23}	-0.31	0.14	0.026
ρ_{24}	-0.26	0.15	0.084
ρ_{34}	0.71	0.068	< 0.001
d.f.	7	3.68	0.05

Table 2.7: The Dependence Structure Implied by the t Copula

	Corn Price	**Soybean Price**	**Corn Yield**	**Soybean Yield**
Corn Price	1	0.74	-0.31	-0.29
Soybean Price		1	-0.31	-0.26
Corn Yield			1	0.71
Soybean Yield				1

Table 2.8: The Revenue Insurance Implied from a Gaussian-copula

	Liability	**Expected Loss**	**Premium Rate**	**Premium**
..... 75% Coverage				
corn(Revenue)	294.65	5.83	0.0198	5.83
soybean(Revenue)	216.60	1.08	0.0050	1.08
WFI(Revenue)	511.25	4.44	0.0087	4.44
..... 85% Coverage				
corn(Revenue)	333.90	13.62	0.041	13.62
soybean(Revenue)	245.52	4.53	0.018	4.53
WFI(Revenue)	579.42	14.30	0.025	14.30

Table 2.9: The Revenue Insurance Implied from a t-copula

	Liability	**Expected Loss**	**Premium Rate**	**Premium**
..... 75% Coverage				
corn(Revenue)	294.46	4.43	0.0152	4.43
soybean(Revenue)	216.56	1.26	0.0059	1.26
WFI(Revenue)	511.01	3.63	0.0072	3.63
..... 85% Coverage				
corn(Revenue)	333.83	11.18	0.033	11.18
soybean(Revenue)	245.49	4.60	0.0187	4.60
WFI(Revenue)	579.32	12.01	0.0207	12.01

Table 2.10: Premium Comparison

	Crop Type			
Copula	Corn	Soybeans	Total Premium for Crop Specific	Total Premium for Whole-Farm
Gaussian copula (75%)	5.83	1.08	6.91	4.44
t copula (75%)	4.43	1.26	5.69	3.63
Gaussian copula (85%)	13.62	4.53	18.15	14.30
t copula (85%)	11.18	4.60	15.78	12.01

Chapter 3

Dynamic Spatial Models for Estimating Systemic Yield Risk

Abstract

Systemic risk is one of the leading causes for the failure of the crop insurance market. This study analyzes systemic crop yield risk by exploring the spatial dependence structure of crop yields. Various dynamic spatial autoregressive models are explored to account for county-level dependence of crop yields. The time-varying dynamic spatial models allow smooth estimation of not only the trend but also the variability and correlation across neighboring sites. The information of the spatial dependence of yield risk both in size and direction will help the construction of a reinsurance plan and better risk management programs for protecting producers from systemic yield risks.

3.1 Introduction

Yield risk for a given crop can differ systematically over space due to the changing agronomic conditions, such as climate and soil type. The concept of spatial correlation is important in crop insurance due to systemic risk in agriculture, which comes from the fact that adverse weather events usually induce substantial yield correlations among counties. For example, drought risk will affect most if not all farmers in a state. As a result of spatially correlated weather pattern, yields are usually spatially correlated and yield shortfalls in a particular area such as a county are likely to be correlated with yield shortfalls in the neighboring counties (Ozaki, et al (2008)).

Usually the precondition of insurability is that individual risks are independent or the covariance among risks is small. This assumption does not hold for some types of spatially correlated weather patterns and systemic production risk. This may present challenges for agricultural insurers. It's been argued that the causes of crop insurance market failure is systemic risk (Miranda and Glauber (1997)). Miranda and Glauber state that U.S. crop insurers face portfolio risk due to a high level of systemic risk, which undermines a crop insurer's ability to diversify risk across space and prevents it from performing the function of pooling risk across farmers. Appropriate measuring and quantifying of spatial risk is important for risk management in order to cope with the systemic nature of production risk. The relationship between production risk at different locations is not only relevant when calculating joint losses from the viewpoint of the insurer, it is also crucial for the reinsurer when diversifying risk among different regions. Modeling the spatial correlation between yield losses in different counties and measurement of the systemic crop yield risk will give implications for insurability, risk management, reinsurance and crop insurance pricing, as well as helping to conduct better management of systemic risk and maintain consistency of premium rates among contiguous counties.

In order to identify appropriate measures for coping with systemic production risk, we may need to use a spatial model to quantify the risk. Economic motivation of using spatial models comes from the fact that there is systemic risk in agricultural insurance. Clearly, the systemic nature of the yield risk depends not only on the neighboring areas but may also depend on the directional neighbors. At locations within a county or within a state, the production/weather risks are highly correlated whereas these spatial dependencies vanish/vary with direction or geographic aggregate level. Thus, the question arises of how the dependence structure of these production risks varies as a function of geographic neighbors and directions.

In view of the relevance of spatial dependence of production risk for insurance issues, there are many attempts to address this topic. The usual approach is based on correlation coefficients between risk variables measured at different locations. The methods are based on correlations depending on distances. Examples of this kind of approach can be found in Goodwin (2001), which applied the spatial technique to the U.S. yield data and calculated the spatial correlations based on distances. Wang and Zhang (2003) also addressed the spatial characteristics of crop insurance. Using the concept of finite range positive dependence of spatial variables, they calculated the effectiveness of risk pooling for cropping areas in the US. We notice that the distance alone may not be a good factor to measure the spatial correlation of county-level yield. For example, the distance between the centroids of two neighboring regions may not necessarily capture the spatial correlation between the entire area's yields, especially when the the counties are large in area. Therefore, the spatial weighting matrix based only on distance may not be an ideal indicator of the spatial dependence for county-level data.

To model the spatial correlation of yields and capture the neighborhood-level spatial effects, a spatial autoregressive model (AR) holds promise. AR models include the

conditional autoregressive (CAR) model and the simultaneous autoregressive (SAR) model. An AR model incorporates the discrete nature of neighborhood information. It pools innovations from contiguous regions and reduces the sampling variance of the estimated conditional distribution of the central county by pooling information from its neighboring regions. In the AR model, a neighborhood structure based on the shape of the lattice is defined. Thus, instead of measuring the distance between centroids of regions, a system is used to define regions to be neighbors based on whether they share common borders. Once the neighborhood structure is defined, models similar to time series autoregressive models are considered. There are many theoretical and empirical studies discussing the AR model (e.g., see Gelfand et al. (1998) and Wall (2004)). Although spatial statistics have been successfully used in other fields and in agriculture for other purposes, the spatial statistics approach has been rarely employed in the study of crop insurance. Ozaki et al. (2008) used CAR models to investigate the spatiotemporal structure in crop insurance with an autogressive structure in both time and space dimensions. This approach assigned equal weights of spatial correlations to all directions and does not take into account the directional effect.

Due to the anisotropic nature of weather and soil patterns, direction is very relevant in the specification of the spatial structure of yield risk. Thus, it calls for more flexible models which can account for the directional effect of spatial dependence associated with agricultural production. An anisotropic/directional spatial autoregressive method holds promise to capture the spatial, directional effect and to faciliate interpretation of the spatial pattern. Kyung and Ghosh (2009) evaluate and estimate the directional CAR model using a Bayesian approach, which however doesn't account for temporal effects. This study would extend the concept of directional spatial dependence into a panel data framework (directional spatio-temporal dimensions) by allowing spatial parameters to change with both time and direction in the setting of modeling agricultural production

risk. We address the spatial effects of yields, while allowing structural parameters to vary over time. Furthermore, the estimated degree of spatial correlation may depend not only on the way the regions are connected, but also on the direction by which they are separated.

This represents a potential improvement over more commonly used methods to model lattice data. This study proposes a directional, dynamic, spatial autoregressive (DAR) model to estimate the spatial pattern and the correlation parameters between crop yields across space and over time. In contrast to the Ozaki et al. (2008) study, we do not model temporal autocorrelation using temporal AR models; instead, we model the spatial dependence structure for all time periods simultaneously. At the same time, we take into account the directional effect within a multiple-time period structure which is an extension of Kyung and Ghosh's static directional CAR model. Our model accepts the notion that spatial relationships may change over time and vary in directions in response to (exogenous) systemic weather changes/patterns. Elaborating on this model by capturing spatial effects over years allows us to examine the evolution of spatial patterns in yield risk. The estimation results will have implications for how strong the spatial dependence is for observations from nearby regions as compared to observations from regions farther apart. This information of the variation of the spatial correlation across space can provide insights into the design of area-yield insurance and reinsurance contracts and for agricultural risk management policy.

The rest of this chapter is organized as follows. Section 2 outlines an autoregressive spatial model approach to modeling yield distributions. Section 3 discusses the estimation methods. Section 4 presents the estimation results, inference, and evaluates the models by model selections methods. Section 5 elaborates the AR approach in temporal and directional dimensions for econometric modeling of yield distributions and illustrates the empirical application of the directional spatial-temporal model to

county-level agricultural yield data. A simulation study is conducted in Section 6 to compute the insurance premium rate of these spatial models, and their implications are discussed. The final section offers concluding remarks.

3.2 Statistical Modeling Framework

Regression analysis is frequently used in traditional statistics. The same idea can be used in spatial analysis to incorporate the neighborhood relationships and spatially correlated error terms. In the presence of spatially correlated error terms, residual analysis becomes complicated and needs to be treated carefully (see Haining (1994), Christensen et al. (1992) and Haslett (1998)). Many researchers describe approaches for exploring spatial autocorrelation based on OLS residuals. Given a set of counties, yields observations collected nearer to each other tend to have similar characteristics compared to when they are further apart. In geography, this feature is known as *Tobler's First Law* (Miller 2004). The autocorrelation between pairs of regions tends to be higher for nearer neighbors than for those that are far apart. These spatial effects are often modeled by using a simultaneous autoregressive (SAR) or conditional autoregressive (CAR) model (see Clayton and Kaldor (1987) and Cressie and Chan (1989), Besag (1974)). The first step to quantifying the structure of spatial dependency in a data set is to define the spatial relationships that exist between objects. In crop insurance, these objects could be farm land units and county level yields. After we define the area objects, we can measure the spatial dependence in different ways. The CAR and SAR models are spatial models which include the spatial correlation parameters to control for the strength of spatial correlation. These models can be used in modeling county-level regional yield data. In many settings, average yields over geographically defined region such as a county or a state, are observed, and regression or classification analysis can

be performed.

To measure systemic crop yield risk, we need to explore the spatial dependence structure of crop yields. This is to account for the spatial correlation between crop yields across space (both in contiguity and direction). We assume that indemnity payments directly or indirectly depend on yield risk measured at several locations. The underlying question is to what extent yield risk exposures at different places can be diversified by increasing the selling area of the insurance and reinsurance contracts. For that purpose, the probability distribution of the joint losses is estimated using a spatial model. Once the joint distribution of the yield risk has been determined, the total loss of the insurers can be calculated by simulation. This study focuses on improved modeling of spatial and temporal components of aggregate production risk. We consider both spatially fixed effect and random effects, while allowing structural parameters to vary over time.

A linear model can be used for the aggregated crop yields Y.

$$\mathbf{Y}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \text{with } \boldsymbol{\epsilon}_t \sim N(0, \boldsymbol{\Lambda}_t), t = 1, \ldots, T. \tag{3.1}$$

where $\mathbf{Y}_t = (Y_{1t}, \ldots, Y_{It})'$ denotes the $I \times 1$ vector of observations of the yield variables from I different counties in certain state at year t. $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \ldots, \epsilon_{It})'$ denotes the $I \times 1$ vector of error terms with mean 0 and variance-covariance $\boldsymbol{\Lambda}_t$.

Starting from this general yield risk model, all the models in the following subsections are illustrated via the Figure (3.1). For each non-spatial or spatial model, the main difference is the specification of $\boldsymbol{\Lambda}_t$, coefficients and the parameter spaces.

3.2.1 Model 0—Non-Spatial Random Coefficients Model

Using regression analysis as a starting point, we will first address the joint county-level yield risk estimation without taking into account spatial correlation. This would serve as

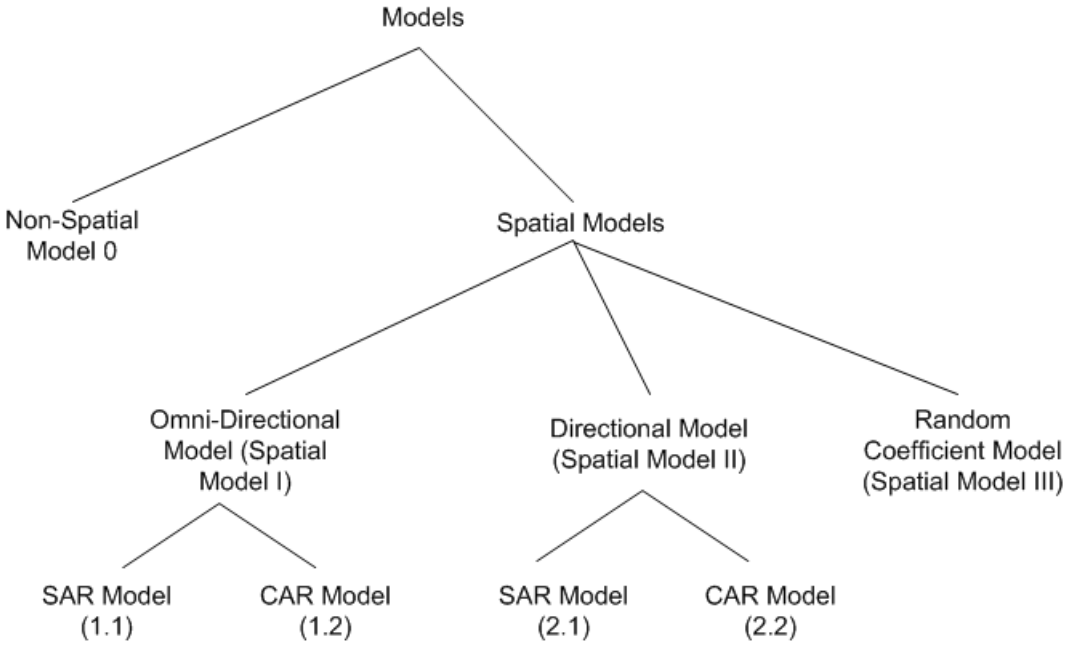


Figure 3.1: Different Models to Capture the Systemic Yield Risk

a base model. Here we choose a non-spatial random coefficients model instead of a fixed coefficients model, since a random coefficients model can better capture the differences in yield risk across counties through time. In a random coefficients model, each county yield has its own underlying inherent trend. This modeling strategy will allow us to address the yield risk trajectories[1] for each individual county. We may think that each county yield risk has its own intercept and slope and all the county yields observed are arising from a population of all such county-level yields. In the design of crop insurance, it may be of interest for insurers and policy makers to have an estimate of an individual farm or county trajectory or inherent trend.

[1]The county risk trajectories are defined as county-specific mean yield parameters estimation.

Focusing on the *ith* county, the mean yield can be modeled as

$$\mu_{it} = \sum_{l=1}^{L} \Psi_l(t)' \beta_{il} + \boldsymbol{\delta}_t' \mathbf{x}_{it} = \boldsymbol{\Psi}(t)' \boldsymbol{\beta_i} + \mathbf{x}_{it}' \boldsymbol{\delta}_t \tag{3.2}$$

where $\mathbf{x}_{it}$ is the vector of weather variables for the *ith* county in year t, $\boldsymbol{\Psi}(t)' = (\Psi_1(t), \ldots, \Psi_L(t))$ and $\boldsymbol{\beta}_i = (\beta_{1i}, \ldots, \beta_{Li})'$. Notice that if we choose $\Psi_l(t) = t^{l-1}, l = 1, \ldots, L$, we get a polynomial. We may also use splines by choosing $\Psi_l(t) = (t - t_l)_+^3, l = 1, \ldots, L$, where $t_1 < t_2 < \ldots < t_L$ are some chosen knots. For example, if $\Psi_l(t)$ can be modeled as a polynomial function in t plus some other covariate x_{it}, then this says that the county i has its own intercept and slope, β_{i1} and β_{i2} (β_{i3}), respectively that determine this trend. This intercept and slope are unique to county i. δ_t is the coefficient(s) of weather which is assumed to vary with t. In this model, $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_T)'$.

The error term $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \ldots, \epsilon_{It})$ in this model may be individual-specific and time-varying. Specifically, there are two sources of variation in this model—within-unit variation and among-unit variation, which can be acknowledged explicitly in this random coefficient model approach. Consequently, the random coefficient model may be thought of as a two-stage hierarchy. The first stage captures the within-unit variation. Suppose the yield observations on the *ith* county follow a quadratic trend with a county-specific intercept β_{i1} and slope parameters β_{i2}, β_{i3}, and the weather coefficients $\boldsymbol{\delta}$, that actual yields vary about this regression line due to within-unit sources (measurement error and individual fluctuations about the trend) is represented explicitly by the deviation ϵ_{ij} with mean 0, and $var(\varepsilon_{it}) = \sigma_{it}^2$. The structure of the $var(\varepsilon_t) = \Sigma_{\varepsilon_t} \equiv \Lambda_\varepsilon$ is a diagonal matrix; that is,

$$\Lambda_t = diag(\sigma_{1t}^2, \ldots, \sigma_{It}^2) \tag{3.3}$$

where $\sigma_{it}^2 = exp\boldsymbol{\Psi}(t)'\gamma_i)$ with $\boldsymbol{\Psi}(t)' = (\Psi_1(t), \ldots, \Psi_L(t))$ and $\boldsymbol{\gamma}_i = (\gamma_{1i}, \ldots, \gamma_{Li})'$.

The second stage captures the among-unit variation; that is, the regression coefficient

$\boldsymbol{\beta}_i$'s variations among individual counties. We may think that there are I numbers of random vectors $\beta_{1(3\times1)}, \ldots, \beta_{I(3\times1)}$ corresponding to each of I counties. Each of which has multivariate probability distribution features summarizing the way in which all coefficients at county $i = 1, \ldots, I$ take on their possible values jointly.

For the ith such vector, the coefficient $\boldsymbol{\beta}_i$ follows a Normal distribution such that

$$\boldsymbol{\beta}_i \overset{\text{iid}}{\sim} N(\boldsymbol{\mu}_\beta, \Sigma_\beta), \tag{3.4}$$

where $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, \beta_{i3})$, $E(\boldsymbol{\beta}_i) = \boldsymbol{\mu}_\beta$ and covariance matrix $var(\boldsymbol{\beta}_i) = \boldsymbol{\Sigma}_\beta$.

These two stages give a complete description of what we believe about each county yield and the population of the counties' yields, acknowledging the two sources of variation explicitly.

This model implies $cov(y_{it}, y_{jt'}) = 0$ for all $i \neq j$ and $t \neq t'$ because

$$\begin{aligned} cov(y_{it}, y_{jt'}) &= cov(E(y_{it}|\boldsymbol{\beta}_i), E(y_{jt'}|\boldsymbol{\beta}_j)) + E(cov(y_{it}, y_{jt'}|\boldsymbol{\beta}_i, \boldsymbol{\beta}_j))) \\ &= cov(\beta_i^{'}\Psi(t), \beta_j^{'}\Psi(t)) + 0 \\ &= \Psi(t)'cov(\beta_i, \beta_j)\Psi(t) + 0 \\ &= 0 \end{aligned}$$

The parameter vector $\boldsymbol{\theta}$ in this model is $\boldsymbol{\theta} = (\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta, \boldsymbol{\mu}_\gamma, \boldsymbol{\Sigma}_\gamma, \boldsymbol{\delta})$. The total number of parameters to be estimated is $p = L + \frac{L(L+1)}{2} + L + \frac{L(L+1)}{2} + T$

The random effects model allows the county-specific random vector $\boldsymbol{\beta}_i$ to describe the trajectories of each individual county and their typical yield risk features. This helps estimation of each individual county's yield risk but ignores any possible spatial correlation in the data. It is necessary to seek other models that can capture the county-specify yield risk as well as the spatial interaction between different counties.

3.2.2 Spatial Model I—Omni-directional Autoregressive Model

Dynamic Simultaneous Autoregressive Model

When there is a need to incorporate the neighborhood relationship and spatial correlations between county-level yield risks in a regression analysis, a Simultaneous Autoregressive (SAR) model can be used to capture spatial dependence among response variables that may evolve over time. The SAR model was first introduced by Whittle (1954).The SAR model has been used extensively for various data observed over irregular lattices in applied econometrics (e.g., Anselin and Florax 1995, Kelejian and Prucha 1999, and Anselin 2002).

Here we extend the SAR model to a dynamic SAR model by allowing the spatial parameters to change over time. Suppose a cross section of I counties is observed over T time periods, the interaction between those county yields over T periods of time can be described by a SAR model. In this dynamic SAR model, the I autoregressions occur simultaneously at each data location at each time and can be written as the following regression form:

$$Y_{it} = \mu_{it} + \rho_t \sum_{j=1}^{I} (Y_{jt} - \mu_{jt}) w_{ij} + \varepsilon_{it}, \text{with } \varepsilon_{it} \sim N(0, \sigma_{it}^2) \tag{3.5}$$

where $w_{ii} = 0$ and w_{ij} denotes the weight attached to jth neighbor of ith county and ρ_t measures the strength of correlation between the neighboring counties that may evolve over time. Here Y_{it} is the crop yield at ith county in year t, where $i = 1, \ldots, I$ and $t = 1, \ldots T$. The rationale of the spatial autoregressive model is that if the locations that are neighbors of location i have negative residuals (yield shocks) which indicate that they are of below-average yields for this location, then the yield in location i is likely to be below average as well. This suggests using the average of residuals from a location's neighbors as a factor to explain the mean of y_{it} as shown in equation (3.5).

In matrix form:

$$\mathbf{Y}_t = \boldsymbol{\mu}_t + \rho_t \mathbf{W}(\mathbf{Y_t} - \boldsymbol{\mu}_t) + \boldsymbol{\epsilon}_t$$

and hence,

$$\mathbf{Y_t} = \boldsymbol{\mu}_t + (I - \rho_t \mathbf{W})^{-1} \boldsymbol{\varepsilon}_t$$

Here $\mathbf{Y}_t$ denotes the $I \times 1$ vector of observations of the dependent variable. $\mathbf{W} = ((w_{ij}))$ is a $I \times I$ non-stochastic matrix which is typically referred as a spatial weighting matrix and ρ_t are unknown parameters which are typically called spatial autoregressive parameters. $\boldsymbol{\epsilon}_t$ is an $I \times 1$ vector of innovations with $E(\boldsymbol{\varepsilon}_t) = 0$ and $var(\boldsymbol{\varepsilon}_t) = D_t = diag(\frac{\sigma_t^2}{n_i}, \ldots, \frac{\sigma_t^2}{n_I})$, where $n_i = \sum_{j \neq i} w_{ij}$ is the number of neighbors at each location i. $\mathbf{N_i}$ is the set of neighbors of location i. This implies that

$$\Lambda_t = var(\mathbf{Y}_t) = \Sigma_{SAR,t} = (I - \rho_t \mathbf{W})^{-1} diag(\frac{\sigma_t^2}{n_1}, \ldots, \frac{\sigma_t^2}{n_I})(I - \rho_t \mathbf{W}')^{-1} \tag{3.6}$$

Thus it follows that,

$$\mathbf{Y_t} \sim N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_{SAR,t}), t = 1, \ldots, T \tag{3.7}$$

We will use NSAR (Normal SAR) to define equation (3.7) as $\mathbf{Y_t} \sim NSAR(\boldsymbol{\mu}_t, \boldsymbol{\sigma}_t, \rho_t, \mathbf{W})$. The mean vector is $\boldsymbol{\mu}_t = (\mu_{1t}, \ldots, \mu_{It})'_{(I \times 1)}$ with the element μ_{it} representing mean yield at region i in time t, $i = 1, \ldots, I$and $t = 1, \ldots, T$. Similar to equation (3.2), it can be modeled as a function of some explanatory variables such as time, weather and other area level covariates with coefficients $\boldsymbol{\beta}_i$ and $\boldsymbol{\delta}_t$. That is,

$$\mu_{it} = \sum_{l=1}^{L} \Psi_l(t)\beta_{il} + x_{it}\delta_t = \boldsymbol{\Psi(t)}'\boldsymbol{\beta_i} + \mathbf{x}_{it}\boldsymbol{\delta}_t \tag{3.8}$$

where x_{it} is the vector of weather variables for the ith county in year t. $\boldsymbol{\Psi}(t)' =$

$(\Psi_1(t), \ldots, \Psi_L(t))$ and $\boldsymbol{\beta}_i = (\beta_{1i}, \ldots, \beta_{Li})'$. We can choose $\Psi_l(t) = t^{l-1}, l = 1, \ldots, L$ to get a polynomial. We may also use splines by choosing $\Psi_l(t) = (t - t_l)_+^3, l = 1, \ldots, L$, where $t_1 < t_2 < \ldots < t_L$ are some chosen knots.

For a well defined model, the matrix of $(I - \rho_t \mathbf{W})$ in equation (3.7) is required to be non-singular. This imposes the following conditions: $a(\mathbf{W}) = \frac{1}{v_{min}(\mathbf{W})} < \rho_t < \frac{1}{v_{max}(\mathbf{W})} = b(\mathbf{W})$and $v_{min}(\mathbf{W}) < 0 < v_{max}(\mathbf{W})$ (Haining, 1990, P.82), where $v_{max}(\mathbf{W})$ and $v_{min}(\mathbf{W})$ are the largest and smallest eigenvalues of $\mathbf{W}$, respectively. To impose this condition on the model, the ρ_t can be estimated as free parameters under the constraints of the lower and upper bounds. The SAR model can be estimated by using maximum likelihood estimation methods. The parameter space $\boldsymbol{\theta}$ in this model is $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_I, \boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_T, \boldsymbol{\sigma}^2, \boldsymbol{\rho})$, where $\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_T^2)', \boldsymbol{\rho} = (\rho_1, \ldots, \rho_T)'$. The total number of parameters is $p = IL + T + T + T = IL + 3T$.

The moments of yields from different counties are:

$$
\begin{aligned}
E(\mathbf{Y}_t) &= \boldsymbol{\mu}_t \\
Cov(\mathbf{Y}_t) &= (I - \rho_t \mathbf{W})^{-1} D_t (I - \rho_t \mathbf{W}')^{-1}, \text{with } D_t = diag(\frac{\sigma_t^2}{n_1}, \ldots, \frac{\sigma_t^2}{n_I})
\end{aligned}
$$

3.2.3 Dynamic Conditional Autoregressive Model

The simultaneous autoregressive approach provides one multivariate model that describes the spatial dependence among the data. We may find it is also interesting to use the dynamic conditional autoregressive (DCAR) model to model the conditional probability distributions of each yield risk observation Y_{it} given the yield risk of all the other observations. By using the notation of the neighborhood matrix, the conditional distribution of Y_{it} in the CAR model is defined as follows:

$$
Y_{it} | Y_{j \neq i,t} \sim N(\mu_{it} + \rho_t \sum_{j \in N_i} w_{ij}(Y_{jt} - \mu_{jt}), \frac{\sigma_t^2}{n_i}) \tag{3.9}
$$

where N_i is the set of neighbors of location i. Locations $i = 1, \ldots, I$ forms a lattice of counties in a state. $n_i = \sum_{j \neq i} w_{ij}$ is the number of neighbors at each county i and w_{ij} denotes the weight attached to jth neighbor of the ith county with $w_{ii} = 0$. We will use NCAR (Normal CAR) to define equation (3.9) as $Y_{it}|Y_{j\neq i,t} \sim NCAR(\mu_{it}, \rho_t, \sigma_t; w_{ij})$. The conditional mean $E(y_{it}|y_{j,t}) = \mu_{i,t} + \rho_t \sum_{j \in N_i}(y_{j,t} - \mu_{j,t})$, and the conditional variance is $\frac{\sigma_t^2}{n_i}$. ρ_t is referred to as the 'spatial correlation or spatial dependence' parameter and controls the strength of spatial association.

By Brook's lemma (Brook (1964)), the joint distribution of $\mathbf{Y}_t$ in matrix notation is as follows:

$$\begin{aligned} \mathbf{Y}_t = (Y_{1t}, \ldots, Y_{It})' &\sim N(\boldsymbol{\mu}_t, (I - \rho_t \mathbf{W})^{-1} D_t) \\ &\sim N(\boldsymbol{\mu}_t, \Sigma_{CAR,t}) \\ &\sim NCAR(\boldsymbol{\mu}_t, \rho_t, \boldsymbol{\sigma}_t, \mathbf{W}, N) \end{aligned} \tag{3.10}$$

where $D_t = diag\left(\frac{\sigma_t^2}{n_1}, ..., \frac{\sigma_t^2}{n_I}\right)$ and $n_i = \sum_{j \neq i} w_{ij}$.

In the CAR model, the covariance of $\mathbf{Y}_t$ is $\boldsymbol{\sigma}_t^2 (I - \rho_t \mathbf{W})^{-1} diag\left(\frac{1}{n_1}, ..., \frac{1}{n_I}\right)$. We need to guarantee that $(I - \rho_t \mathbf{W})^{-1} diag\left(\frac{1}{n_1}, ..., \frac{1}{n_I}\right)$ is symmetric.

$$\begin{aligned} & (I - \rho_t \mathbf{W})^{-1} diag\left(\frac{1}{n_1}, ..., \frac{1}{n_I}\right) \\ = \; & [diag\,(n_1, ..., n_I)\,(I - \rho_t \mathbf{W})]^{-1} \\ = \; & [diag\,(n_1, ..., n_I) - \rho_t diag\,(n_1, ..., n_I)\,\mathbf{W}]^{-1} \end{aligned}$$

Therefore we need $diag\,(n_1, ..., n_I)\,W$ to be symmetric, that is,

$$n_i W_{ij} = n_j W_{ji}$$

This does not hold for a symmetric matrix W. For the CAR model we let the matrix

$\tilde{\mathbf{W}} = ((\tilde{w}_{ij} = \frac{w_{ij}}{n_i}))$ be scaled matrix, and $\mathbf{W} = ((w_{ij}))$ is an unscaled symmetric matrix. One way to specify which to use 0 and 1 is defined below (see Wall (2004) pages 314 and 319).

$$w_{ij} = \begin{cases} 1 & \text{if region } i \text{ shares a common edge/border with region } j, \\ 0 & \text{if i=j,} \\ 0 & \text{otherwise} \end{cases}$$

The parameter vector $\boldsymbol{\theta}$ in this model is $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_I, \boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_T, \boldsymbol{\sigma}^2, \boldsymbol{\rho})$, where $\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_T^2)', \boldsymbol{\rho} = (\rho_1, \ldots, \rho_T)'$. The total number of parameters is $p = IL + T + T + T = IL + 3T$.

The moments of yields from different counties are:

$$\begin{aligned} E(\mathbf{Y}_t) &= \boldsymbol{\mu}_t \\ Cov(\mathbf{Y}_t) &= (I - \boldsymbol{\rho}_t \mathbf{W})^{-1} \mathbf{D}_t, \text{with } \mathbf{D}_t = diag(\frac{\sigma_t^2}{n_1}, \ldots, \frac{\sigma_t^2}{n_I}) \end{aligned}$$

3.2.4 Spatial Model II—Anisotropic/Directional Dynamic Autoregressive Model—Two-Parameter Autoregressive Models

One of the limitations of both DSAR and DCAR models is that the neighbors are formed using a neighborhood matrix and it assumes the same spatial correlation in all directions by assigning equal weight to all directions. Thus, the effect of direction is completely ignored. This is counter-intuitive for the spatial correlation of crop yields. Usually, due to the larger weather similarity in the West-East (W-E) direction than the South-North (S-N) direction, the yields distribution between the west and east should be more similar than that of the north-south. Thus, we can expect that the spatial correlation between W-E should be larger than that of S-N. In spatial statistics, the directional spatial dependence is called anisotropy. In order to incorporate the idea of

directional effects in spatial correlation, an anisotropic DSAR/DCAR model is proposed in this study. In some recent literature, there have been some attempts to use different CAR models for different parts of the region. White and Ghosh (2009) develop a CAR model with stochastic parameters to determine effects of the neighborhood. Most of the Gaussian Markov Random fields (GMRFs) in the current literature do not specifically take into account the anisotropic nature of the areal data. Reich et al. (2007) present a model for periodontal disease and use separate CAR models for separate jaws. White and Ghosh (2008) used a stochastic parameter within the CAR framework to determine effects of the neighbors. Kyung and Ghosh (2009) point out these model can not capture the inherent anisotropy in spatial data and present a directional CAR (DCAR) model to accommodate spatial variations by using different weights to neighbors in different directions in a Bayesian framework.

Here we adopt the DCAR approach (Kyung and Ghosh (2009), Waller and Gotway (2004)) to estimate the spatial correlation of crop yields in different regions in the U.S. We extended the DCAR model into multiple time periods and model the dynamic directional CAR (DDCAR) model. At the same time, an analogous version of dynamic directional SAR model (DDSAR) model is proposed.

Suppose that we want to allow the spatial dependence to vary with direction. One way to do this is to consider two spatial dependence parameters, one for each direction. For example, if we consider a spatial locations and the first nearest-neighbor structure in a SAR or a CAR model, we could allow the strength of spatial dependence in the east-west direction to be different from the strength of spatial dependence in the north-south direction and specify the different spatial dependence parameters to different regions based on directions. Here we adopt Kyung and Ghosh's (2008) definition of a neighbor structure that depends on the directions between centroids for any pair of sub-regions. Let $s_i = (s_{1i}, s_{2i})$ be a centroid of the county S_i, where s_{1i} corresponds to the longitude

(x-axis) and s_{2i} is the latitude (y-axis). The angle in radians between S_i and S_j is defined as

$$\alpha_{ij} = \alpha(S_i, S_j) = \begin{cases} |tan^{-1}(\frac{s_{2j}-s_{1j}}{s_{2i}-s_{1i}})| & if s_{2j} - s_{1j} \geq 0, \\ -(\pi - |tan^{-1}(\frac{s_{2j}-s_{1j}}{s_{2i}-s_{1i}})| & if s_{2j} - s_{1j} \leq 0. \end{cases}$$

Let N_{ik} represent a set of neighboring counties for the ith county S_i that are based on the whether their borders touch or not. The subset based on different directions can be denoted as N_{ik}, where k stands for a different direction. Each pair of the diagonally opposite neighbor sets can form a neighborhood. For example, if there are two directions (N-S and W-E), N_{i1} and N_{i2} can be constructed based on the associated N-S and W-E neighborhoods as defined as follows:

$$\begin{aligned} N_i^{(1)} = N_i^{NS} &= j : j \in N_i \text{ and } (\frac{\pi}{4} \leq \alpha_{ij} < \frac{3}{4}\pi \text{ or } \frac{5}{4}\pi \leq \alpha_{ij} < \frac{7}{4}\pi) \\ N_i^{(2)} = N_i^{WE} &= j : j \in N_i \text{ and} \\ &\quad (\frac{7}{4}\pi \leq \alpha_{ij} < 2\pi \text{ or } 0 \leq \alpha_{ij} < \frac{\pi}{4} \text{ or } \frac{3}{4}\pi \leq \alpha_{ij} < \frac{5}{4}\pi) \end{aligned} \quad (3.11)$$

For example, in Figure 3.2, S_j is in the north-south (NS) region of S_i since $\alpha(S_i, S_j)$ is in $(\frac{\pi}{4}, \frac{3}{4}\pi)$. S_k s in the west-east (WE) region of S_i since $\alpha(S_i, S_k)$ is in $(0, \frac{\pi}{4})$. It can be shown that $N_i = N_i^{NS} \cup N_i^{WE}$. By this definition, we can define the directions of neighbors for each county. Based on subsets of the directional neighborhoods, $N_i = N_i^{NS} \cup N_i^{WE}$, we can construct the directional neighborhood matrices as $W^{(1)} = ((\omega_{ij}^1))$ and $W^{(2)} = ((\omega_{ij}^2))$, respectively. That is, $w_{ij}^k = \begin{cases} 1 & \text{if } j \in N_i^k, \\ 0 & \text{if i=j}, \\ 0 & \text{otherwise} \end{cases}$ where $k = 1$ and 2 stand for N-S and W-E directions, respectively. Notice that $W = W^{(1)} + W^{(2)}$ are the same neighborhood matrix as in the regular unidirectional spatial

autoregressive model.

Based on the directional neighborhood matrix W, a DDCAR model can be developed to account for anisotropy. Let $\rho_t^{(1)}$ and $\rho_t^{(2)}$ denote the directional spatial effects corresponding to N_{i1} and N_{i2}. $\rho_t^{(1)}$ and $\rho_t^{(2)}$ are the S-N and W-E spatial correlation respectively.

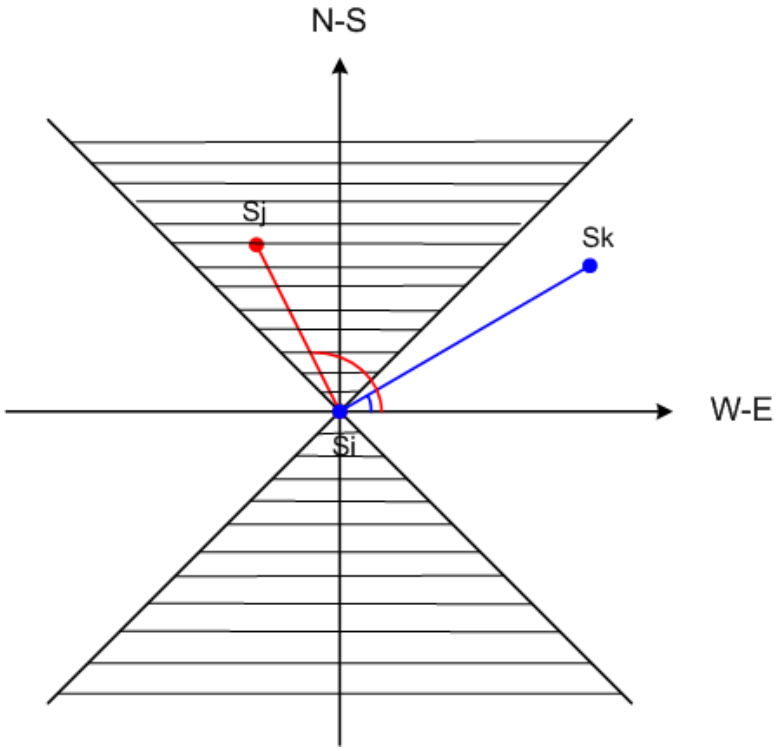


Figure 3.2: Directional Neighborhood Illustration

The joint distribution of Y is:

$$\begin{aligned} Y_t = (y_{1t}, \ldots, y_{It})' &\sim N(\boldsymbol{\mu}_t, (I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)})^{-1} D_t) \\ &\sim (\boldsymbol{\mu}_t, \Sigma_{DCAR,t}), t = 1, \ldots, T \end{aligned} \tag{3.12}$$

where $\boldsymbol{\mu}_t$ is the same as the mean specification in SAR and CAR model (equation 3.8). $D_t = diag(\frac{\sigma_t^2}{n_1}, \ldots, \frac{\sigma_t^2}{n_I})$. The parameter space $\boldsymbol{\theta}$ in this model is $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_I, \boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_T, \boldsymbol{\sigma}^2, \boldsymbol{\rho}^{(1)}, \boldsymbol{\rho}^{(2)})$, where $\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_T^2)', \boldsymbol{\rho}^{(k)} = (\rho_1, \ldots, \rho_T)'$ for $k = 1, 2$. The total number of parameters is $p = IL + T + T + 2T = IL + 4T$.

Similarly, a Directional SAR model can be defined as:

$$\mathbf{Y_t} = \boldsymbol{\mu}_t + (I - \rho_t^{(1)}\mathbf{W}^{(1)} - \rho_t^{(1)}\mathbf{W}^{(1)})^{-1}\boldsymbol{\epsilon}_t.$$

The joint distribution of $\mathbf{Y_t}$ in DSAR model is:

$$\begin{aligned} Y_t = (y_{1t}, \ldots, y_{It})' &\sim N(\boldsymbol{\mu}_t, (I - \rho_t^{(1)}W^{(1)} - \rho_t^{(2)}W^{(2)})^{-1}D_t(I - \rho_t^{(1)}W^{(1)\prime} - \rho_t^{(2)}W^{(2)\prime})^{-1}), \\ &\sim (\boldsymbol{\mu}_t, \Sigma_{DSAR,t}), t = 1, \ldots, T \end{aligned}$$

where$D_t = diag(\frac{\sigma_t^2}{n_1}, \ldots, \frac{\sigma_t^2}{n_I})$.

The parameter space $\boldsymbol{\theta}$ in this model is $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_I, \boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_T, \boldsymbol{\sigma}^2, \boldsymbol{\rho}^{(1)}, \boldsymbol{\rho}^{(2)})$, where $\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_T^2)', \boldsymbol{\rho}^{(k)} = (\rho_1, \ldots, \rho_T)'$ for $k = 1, 2$. The total number of parameters is $p = IL + T + T + 2T = IL + 4T$.

For a proper Normal distribution, the variance-covariance matrix needs to be symmetric and positive definite. It is not hard to see that $\Sigma_{DSAR,t}$ is symmetric and positive definite matrix. Kyung and Ghosh (2009) shows that $\Sigma_{DCAR,t}$ is symmetric and positive definite if imposing the condition of $max_k|\rho^{(k)}| < 1$ and if using scaled weighting matrix $\tilde{W}^{(k)} = ((\tilde{w}_{ij}^{(k)} = \frac{w_{ij}^{(k)}}{n_i}))$, where $k = 1, 2, n_i$ is the total number of neighbors of county i.

3.2.5 Spatial Model III—Random Coefficient Model with Spatial Dependence on Beta (Slope and Intercept)

Another way to introduce spatial correlation is to define the spatial relationship between the regression coefficients of all counties' yields. This model is similar to model 0 but the time coefficients that go into one county yield and the coefficients that go into another county level yield are correlated through a conditional autoregressive model structure. The mean function is the same as defined in equation (3.2), but the coefficients become spatial random coefficients given by vector $\boldsymbol{\beta}^{(l)} = (\beta_{l1}, \ldots, \beta_{lI})', l = 1, \ldots, L$, which stands for the lth coefficient of all counties' yields regression coefficients:

$$\boldsymbol{\beta}^{(l)} \sim NCAR(\boldsymbol{\mu}_\beta^{(l)}, \sigma_\beta^{(l)}, \rho_\beta^{(l)}; W), l = 1, \ldots, L$$

where $\boldsymbol{\beta}_{(l)} = (\beta_1^{(l)}, \ldots, \beta_I^{(l)})$ denotes the lth coefficient in all counties of which the spatial correlations are captured by the joint Normal Conditional Autoregressive ($NCAR$) distribution.[2]. There are L independent CAR models in this specification. The parameter space $\boldsymbol{\theta}$ in this model is $\boldsymbol{\theta} = (\boldsymbol{\mu}_\beta^{(1)}, \ldots, \boldsymbol{\mu}_\beta^{(L)}, \boldsymbol{\sigma}_\beta^{(1)}, \ldots, \boldsymbol{\sigma}_\beta^{(L)}, \rho_\beta^{(1)}, \ldots, \rho_\beta^{(L)}, \boldsymbol{\delta}_t)$. The total number of parameters is $p = IL + L + L + T = (I + 2)L + T$.

This model implies at any given t, the covariance between any two county-level yield as follows:

$$\begin{aligned} cov(y_{it}, y_{jt}) &= cov(E(y_{it}|\beta_{it}), E(y_{jt}|\beta_{jt})) + E(cov(y_{it}, y_{jt}|\beta_i, \beta_j))) \\ &= cov(\beta_i^{'}\Psi(t), \beta_j^{'}\Psi(t)) \\ &= \Psi(t)' cov(\beta_i, \beta_j)\Psi(t) \\ &\neq 0 \end{aligned}$$

[2]The $NCAR$ distribution is specified in section (3.2.3) and equation (3.10)

Table 3.1: Model Summary

Model	Model Description	Specification	Parameter
	general multivariate normal	$\mathbf{Y}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_t \sim N(0, \Lambda_t)$	
0	Random Coefficient Model	$\mu_{it} = \sum_{l=1}^{L} \psi_l(t)\beta_{il} + \delta_t x_{it}, \Lambda_t = \sigma_t I_I$ $\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}_\beta, \Sigma_\beta)$	$\boldsymbol{\mu}_\beta, \Sigma_\beta, \boldsymbol{\sigma}, \boldsymbol{\delta}$
1.1	SAR	$\mu_{it} = \sum_{l=1}^{L} \psi_l(t)\beta_{il} + \delta_t x_{it}$, $\Lambda_t = (I - \rho_t W)^{-1} D_t (I - \rho_t W')^{-1}$	$\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_I, \boldsymbol{\sigma}^2, \boldsymbol{\rho}, \boldsymbol{\delta}$
1.2	CAR	$\mu_{it} = \sum_{l=1}^{L} \psi_l(t)\beta_{il} + \delta_t x_{it}$, $\Lambda_t = (I - \rho_t W)^{-1} D_t$	$\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_I, \boldsymbol{\sigma}^2, \boldsymbol{\rho}, \boldsymbol{\delta}$
2.1	Directional SAR	$\mu_{it} = \sum_{l=1}^{L} \psi_l(t)\beta_{il} + \delta_t x_{it}$, $\Lambda_t = \Sigma^*_{DSAR,t}$	$\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_I, \boldsymbol{\sigma}^2, \boldsymbol{\delta}, \boldsymbol{\rho}^{(1)}, \boldsymbol{\rho}^{(2)}$
2.2	Directional CAR	$\mu_{it} = \sum_{l=1}^{L} \psi_l(t)\beta_{il} + \delta_t x_{it}$, $\Lambda_t = \left(I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)}\right)^{-1} D_t$	$\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_I, \boldsymbol{\sigma}^2, \boldsymbol{\delta}, \boldsymbol{\rho}^{(1)}, \boldsymbol{\rho}^{(2)}$
3	Random Coefficient Model with CAR	$\mu_{it} = \sum_{l=1}^{L} \psi_l(t)\beta_{il} + \delta_t x_{it}, \Lambda_t = \sigma_t I_I$, $\boldsymbol{\beta}^l \sim CAR(\boldsymbol{\mu}_\beta^{(l)}, \sigma_\beta^{(l)}, \rho_\beta^{(l)}; w)$	$\boldsymbol{\mu}_\beta^{(1)}, \ldots, \boldsymbol{\mu}_\beta^{(L)}, \sigma_\beta^{(1)}, \ldots, \sigma_\beta^{(1)}$, $\rho_\beta^{(1)}, \ldots, \rho_\beta^{(L)}, \boldsymbol{\delta}, \boldsymbol{\sigma}^2$

From this equation we can see that the correlation between the yield risks of any two counties come from the correlation of the random coefficients $\boldsymbol{\beta}^{(l)}$s.

All the models in this section are summarized in the tables 3.1.

3.3 Statistical Estimation Method—Bayesian and MLE

Once a valid model is specified, parameter estimates can be obtained using maximum likelihood methods, or the posterior distribution of the parameters (Schabenberger and Gotway 2005). The estimation methods for these models specified in the previous section will be discussed in this section. One method that has been employed to estimate these models is the maximum likelihood estimation procedure, where the likelihood corresponds to a multivariate Normal distribution in our specifications. Given that the

likelihood function involves the determinant of a matrix whose dimensions depend on the sample size, there can be significant difficulties in the practical computation of these estimators, especially if we need to extend the spatial model into dynamic panel model and include multiple time periods. In this case, there are a large number of parameters due to the dynamic features of the spatial model. In developing this MLE, we will develop a modular approach which can solve these multiple parameters problems.

From the statistical estimation perspective, a Gaussian CAR model has the merit that under fairly general regularity conditions, lower dimensional conditional Gaussian distributions uniquely determine joint Gaussianity of the spatial CAR processes. Thus, the maximum likelihood and the Bayesian estimates can be easily obtained. The statistical estimation method of SAR model, CAR model, Directional SAR and CAR models, and Random Coefficient Models will be discussed and the estimation results will be presented in this section.

3.3.1 SAR and DSAR Model Parameter Estimation

The likelihood function of the dynamic SAR model is given by

$$\begin{aligned} L = & \prod_t \{ |\sigma_t^2 (I - \rho_t W)^{-1} D \left(I - \rho_t W'\right)^{-1} |^{-\frac{1}{2}} \\ & \exp(-\frac{1}{2\sigma_t^2} (\mathbf{Y}_t - \boldsymbol{\mu}_t)' (I - \rho_t W') D^{-1} (I - \rho_t W)(\mathbf{Y}_t - \boldsymbol{\mu}_t)) \} \end{aligned} \tag{3.13}$$

where $D = diag(\frac{1}{n_1}, \ldots, \frac{1}{n_I})$. The log-likelihood function $-2\log(L)$

$$\begin{aligned} LLF(\theta) &= \sum_t \left\{ I \cdot \log\left(\sigma_t^2\right) - 2\log|(I-\rho_t W)| + \log|D| \right\} \\ &\quad + \sum_t \frac{1}{\sigma_t^2} (\mathbf{Y}_t - \mu_t)'(I-\rho_t W')D^{-1}(I-\rho_t W)(\mathbf{Y}_t - \mu_t) \\ &= T \cdot (\log|D|) + \sum_t \left\{ I \cdot log\sigma_t^2 - 2\log|(I-\rho_t W)| \right\} \\ &\quad + \sum_t \exp\left(-\log\sigma_t^2\right) (\mathbf{Y}_t - \left(\Phi(t)' \otimes I_I\right) vec(\beta))'(I-\rho_t W')D^{-1} \\ &\quad (I-\rho_t W)(\mathbf{Y}_t - \left(\Phi(t)' \otimes I_I\right) vec(\beta)) \end{aligned}$$

where β is $L \times I$ and Φ is $L \times 1$. Note that $vec(\beta)$ is a vector that stacks the columns of matrix β. The vectorization property implies

$$\beta'\Phi(t) = vec\left(\beta'\Phi(t)\right) = \left(\Phi(t)' \otimes I_I\right) vec(\beta),$$

where $\otimes$ is the Kronecker product.

Let $\Phi(t) = (1, t^*, (t^*)^2)$, where $t^* = [t - 0.5(T+1)]/T$. We use the polynomial of order 2 to fit the μ_{it}. In specific, $\mu_{it} = \beta_{0i} + \beta_{1i}t^* + \beta_{2i}(t^*)^2$. The details of the algorithm are described in the Appendix A.

Similarly, the DSAR model can be estimated with the following likelihood and log-likelihood functions.

$$\begin{aligned} L &= \prod_t \Big\{ |\sigma_t^2 (I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)})^{-1} D \left(I - \rho_t^{(1)} W^{(1)\prime} - \rho_t^{(2)} W^{(2)\prime}\right)^{-1}|^{-\frac{1}{2}} \\ &\quad \exp(-\frac{1}{2\sigma_t^2} (\mathbf{Y}_t - \boldsymbol{\mu}_t)'(I - \rho_t^{(1)} W^{(1)\prime} - \rho_t^{(2)} W^{(2)\prime}) D^{-1} \\ &\quad (I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)})(\mathbf{Y}_t - \boldsymbol{\mu}_t)) \Big\} \end{aligned}$$

where $D = diag(\frac{1}{n_1}, \ldots, \frac{1}{n_I})$. The log-likelihood function $-2\log(L)$

$$\begin{aligned} LLF(\theta) &= \sum_t \left\{ I \cdot \log\left(\sigma_t^2\right) - 2\log|(I - \rho_t^{(1)}W^{(1)} - \rho_t^{(2)}W^{(2)}| + \log|D| \right\} \\ &\quad + \sum_t \frac{1}{\sigma_t^2}(\mathbf{Y}_t - \mu_t)'(I - \rho_t^{(1)}W^{(1)\prime} - \rho_t^{(2)}W^{(2)\prime})D^{-1} \\ &\quad (I - \rho_t^{(1)}W^{(1)} - \rho_t^{(2)}W^{(2)})(\mathbf{Y}_t - \mu_t) \\ &= T \cdot (\log|D|) + \sum_t \left\{ I \cdot log\sigma_t^2 - 2\log|(I - \rho_t^{(1)}W^{(1)} - \rho_t^{(2)}W^{(2)})| \right\} \\ &\quad + \sum_t \exp\left(-\log\sigma_t^2\right)(\mathbf{Y}_t - \left(\Phi(t)' \otimes I_I\right) vec(\beta))' \\ &\quad (I - \rho_t^{(1)}W^{(1)\prime} - \rho_t^{(2)}W^{(2)\prime})D^{-1} \\ &\quad (I - \rho_t^{(1)}W^{(1)} - \rho_t^{(2)}W^{(2)}) \cdot (\mathbf{Y}_t - \left(\Phi(t)' \otimes I_I\right) vec(\beta)) \end{aligned}$$

3.3.2 CAR and DCAR Model Parameter Estimation

The maximum likelihood estimation method can be carried out to estimate the CAR model. The likelihood function $L(\theta)$ is defined as follows.

$$L = |\sigma_t^2(I - \rho_t W)^{-1}D|^{-\frac{1}{2}} \exp(-\frac{1}{2}(\mathbf{Y}_t - \mu_t)'(\sigma_t^2(I - \rho_t W)^{-1}D)^{-1}(\mathbf{Y}_t - \mu_t))$$

where $\mu_t = (\mu_{1t}, \ldots, \mu_{It})'$, $D = diag(\frac{1}{n_1}, \ldots, \frac{1}{n_I})$ and $n_i = \sum_{j \neq i} w_{ij}$. Then the log-likelihood function $LLF(\theta)$ is $(-2\log(L(\theta))$:

$$\begin{aligned}
LLF &= \sum_t \left\{ I \cdot \log \sigma_t^2 - \log |I - \rho_t W| + \frac{1}{\sigma_t^2} (\mathbf{Y}_t - \mu_t)' D^{-1} (I - \rho_t W)(\mathbf{Y}_t - \mu_t) \right\} \\
&= \sum_t \left\{ I \cdot \log \sigma_t^2 - \log |I - \rho_t W| + \frac{1}{\sigma_t^2} (\mathbf{Y}_t - \beta' \Phi(t))' D^{-1} (I - \rho_t W)(\mathbf{Y}_t - \beta' \Phi(t)) \right\} \\
&= \sum_t \left\{ I \cdot \log \sigma_t^2 - \log |I - \rho_t W| \right\} \\
&+ \sum_t \Bigg\{ \exp\left(-\log \sigma_t^2\right) \Big(\mathbf{Y}_t - \left(\Phi(t)' \otimes I_I\right) vec(\beta) \Big)' D^{-1} \\
&(I - \rho_t W) \Big(\mathbf{Y}_t - \left(\Phi(t)' \otimes I_I\right) vec(\beta) \Big) \Bigg\} \qquad (3.14)
\end{aligned}$$

where β is $L \times I$ and Φ is $L \times 1$. Note that $vec(\beta)$ is a vector that stacks the columns of matrix β. The vectorization property implies

$$\beta' \Phi(t) = vec\left(\beta' \Phi(t)\right) = \left(\Phi(t)' \otimes I_I\right) vec(\beta),$$

where $\otimes$ is Kronecker product. The details of the algorithm for CAR model estimation are described in the Appendix B.

Similarly, the DCAR model can be estimated with the following likelihood and log-likelihood functions.

$$\begin{aligned}
L &= |\sigma_t^2 (I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)})^{-1} D|^{-\frac{1}{2}} \\
&\quad \exp(-\frac{1}{2} (\mathbf{Y}_t - \mu_t)' (\sigma_t^2 (I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)})^{-1} D)^{-1} (\mathbf{Y}_t - \mu_t))
\end{aligned}$$

The log-likelihood function $LLF(\theta)$ is $(-2\log(L(\theta))$:

$$
\begin{aligned}
LLF = & \sum_t \Big\{ I \cdot \log \sigma_t^2 - \log |I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)}| \\
& + \frac{1}{\sigma_t^2} (\mathbf{Y}_t - \mu_t)' D^{-1} (I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)}) (\mathbf{Y}_t - \mu_t) \Big\} \\
= & \sum_t \Big\{ I \cdot \log \sigma_t^2 - \log |I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)}| \\
& + \frac{1}{\sigma_t^2} (\mathbf{Y}_t - \beta' \Phi(t))' D^{-1} (I - \rho_t W) (\mathbf{Y}_t - \beta' \Phi(t)) \Big\} \\
= & \sum_t \Big\{ I \cdot \log \sigma_t^2 - \log |I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)}| \Big\} \qquad (3.15) \\
& + \sum_t \Bigg\{ \exp\left(-\log \sigma_t^2\right) \Big(\mathbf{Y}_t - \left(\Phi\left(t\right)' \otimes I_I\right) vec\left(\beta\right) \Big)' D^{-1} (I - \rho_t^{(1)} W^{(1)} - \rho_t^{(2)} W^{(2)}) \\
& \Big(\mathbf{Y}_t - \left(\Phi\left(t\right)' \otimes I_I\right) vec\left(\beta\right) \Big) \Bigg\}
\end{aligned}
$$

3.3.3 Random Coefficient Model Estimation

In the random coefficient models, because the likelihood function may involve high-dimensional integration, posterior estimation is not easy to achieve analytically. Therefore, Markov chain Monte Carlo (MCMC) methods need to be used to obtain samples from the path of Markov chains whose stationary density is the posterior density. We can use Bayesian inference about the unknown parameters when the likelihood function involves high-dimensional integration.

Non-spatial Random Effect model Estimation

In the random coefficient model, $\boldsymbol{\theta} = (\boldsymbol{\mu}_\beta, \Sigma_\beta, \boldsymbol{\sigma}^2, \boldsymbol{\delta})$, and the prior is $\pi(\boldsymbol{\theta}|\mathbf{Y})$. By using MCMC methods, we can obtain samples from the path of MC to estimate posterior density and inference for the parameters.

This model has the form of a regression model unique to the ith county. Each county

has such a model.

$$\begin{aligned} y_{it} &\sim N(\mu_{it}, \sigma_{it}) \\ \mu_{it} &= \beta_{i1} + \beta_{i2}t^* + \beta_{i3}(t^*)^2 + \delta_t x_{it} \end{aligned}$$

where $t^* = [t - 0.5(T+1)]/T$. We estimate this model by using the Bayesian method. The coefficient $\boldsymbol{\beta}_i$ follows a Normal distribution such that:

$$\begin{aligned} \boldsymbol{\beta}_i &\overset{\text{iid}}{\sim} N(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta) \\ \sigma_t &= exp(-\eta_t) \\ \eta_t &\sim N(\mu_{\eta_t}, \sigma_\eta) \\ \mu_{\eta_t} &= \gamma_1 + \gamma_2 \tilde{t} \end{aligned}$$

The prior:

$$\begin{aligned} \mu_\beta &\sim N(\mu_{0,\beta}, \sigma_{\beta_0}) \\ \boldsymbol{\Sigma}_\beta &\sim Wishart(R_{\beta[,]}, 4) \\ \gamma_i &\sim N(\gamma_0, \tau_{\gamma_0}), i = 1, 2 \\ \sigma_\eta &\sim Gamma(0.1, 0.1) \end{aligned}$$

where $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, \beta_{i3})$, $E(\boldsymbol{\beta}_i) = \boldsymbol{\mu}_\beta$ and covariance matrix $var(\boldsymbol{\beta}_i) = \boldsymbol{\Sigma}_\beta$. Gamma$(\alpha, \theta)$ stands for a Gamma distribution with a shape parameter α and scale parameter θ. The covariance matrix $\boldsymbol{\Sigma}_\beta$ follows a Wishart distribution (a multivariate scaled χ^2 distribution) with scale matrix $R_{\beta[,]}$ and degrees of freedom 4.

Random Coefficient Spatial Model III Estimation (Correlations From Slope and Intercept)

In the random coefficient spatial model, the coefficients that go into one county yield and the coefficients that go into another county level yield are correlated through a CAR model structure. That is,

$$\begin{aligned} y_{it} &\sim N(\mu_{it}, \sigma_{it}) \\ \mu_{it} &= \beta_{i1} + \beta_{i2}t^* + \beta_{i3}(t^*)^2 + \delta_t x_{it} \end{aligned}$$

We can estimate this model by using Bayesian method.

$$\begin{aligned} \sigma_t &= exp(-\eta_t) \\ \eta_t &\sim N(\mu_{\eta_t}, \sigma_\eta) \\ \mu_{\eta_t} &= \gamma_1 + \gamma_2 \tilde{t} \end{aligned}$$

The Spatial Random Coefficients:

$$\boldsymbol{\beta}^{(l)} \sim NCAR(\boldsymbol{\mu}_\beta^{(l)}, \sigma_\beta^{(l)}, \rho_\beta^{(l)}; W), l = 1, \ldots, L$$

The prior:

$$\begin{aligned}
\mu_\beta &\sim N(0, 0.01) \\
\sigma_\beta &\sim GAMMA(0.1, 0.1) \\
\rho_\beta &\sim UNIF(0, 1) \\
\gamma &\sim MNORM(\gamma_0, \sigma_{\gamma_0}) \\
\sigma_\eta &\sim GAMMA(0.1, 0.1) \\
\sigma_\eta &= \frac{1}{\sqrt{\tau_\eta}}
\end{aligned}$$

$Gamma(\alpha, \theta)$ stands for a Gamma distribution with a shape parameter α and scale parameter θ.

There is no closed form for the posterior distribution of $\boldsymbol{\theta}$. In order to obtain samples from the path of Markov Chain, we need specify starting values of each parameters for different chains and run parallel chains. After 1000 burn-in replications, we obtain approximate samples from the posterior density $\pi(\boldsymbol{\theta}_t|Y_t)$.

3.4 Estimation Results and Inference

In this section we apply various spatial model procedures to county level corn yield data and weather data in Illinois. The application is motivated by the fact that there is systemic risk in agricultural insurance. Modeling the spatial pattern of yield risks in different counties will help us understand and quantify the systemic yield risk and conduct better design of area-based insurance and reinsurance. Various spatial models are estimated and the parameter estimates are interpreted in the crop insurance setting. This section ends with some model comparison methods to select the best spatial model to capture the agricultural systemic risk.

3.4.1 Application to U.S. Agricultural Production Data

This study focuses on the location, time and spatio-temporal components associated with county-level yield data to improve the prediction of the yields and identify the systemic yield risk. A sample of 102 county-level corn yields over a 82-year period in Illinois is used in this study. The lattice is formed by the counties of Illinois. The centroid of each county is used to calculate the direction of the spatial correlation. We also use weather data from the National Climatic Data Center (NCDC). The variable used here is the Palmer "Z" Index (ZNDX) and July's temperature (TEMP). The July's temperature is selected since this month is the critical growing season for corn. The ZNDX is the generated monthly Z values, and they can be expressed as the "Moisture Anomaly Index." Each monthly Z value is a measure of the departure from normal of the soil moisture climate for that month. It is an indicator of severe wetness if the ZNDX values is above 2.5, while it is severe drought if the ZNDX values fall below -2. The annual plots of ZNDX values of Iowa and Illinois (Figure (3.3)) show small negative values in year 1991 and 2005 in Illinois, which correspond to drought years in that state. A large positive ZNDX value in 1993 corresponds to the flood year in Illinois. In most other regular years, the ZNDX values are winthin the normal ranges. Illinois is in the central Corn Belt where climate and soil are ideal for corn production, so the corn yield volatility will tend to be low in this state.

Using traditional likelihood and Bayesian methods, we fit the following nonspatial models and spatial autoregressive models of SAR, CAR, DSAR and DCAR for the county-level yield risk data:

- Fixed effect SAR/CAR model
- Directional DSAR/DCAR model
- Non-spatial Random Effect Model as defined in Section 3.2.1

- Random Coefficient Model with Spatial Dependence on Slope and Intercept as defined is Section 3.2.5.

3.4.2 Model Fitting and Choice

This first part of the estimation is about the mean function parameters. The time coefficients are county-specific for all the models. Most counties show strong time trends and the significance of quadratic time coefficients vary by counties. Since there are 102 counties in Illinois, we will not present the county-specific time coefficient estimates for each individual county. On average, the intercept term is 81.69, the time coefficient is 0.8, and the coefficient on quadratic time is 0.01. We first estimate the model with the time-varying weather coefficients (δ_t) of the two weather variables—ZNDX and July Temperature. The MLE estimation results from SAR and CAR model for the time-varying weather parameters δ_t are illustrated in Figure 3.4 and 3.5, respectively, which indicate that the effect of ZNDX in most years is positive and this effect is varying by years. Similarly, the effect of July temperature on yield risk is time varying but not as significant as the effect of ZNDX. The results are also presented in Table 3.10 and show that the ZNDX plays a stronger role in drought year (for example, $\delta^{(CAR)}_{2005,zndx} = 6.16$ in 2005) than in regular year ($\delta^{(CAR)}_{2007,zndx} = 0.146$ in 2007), which can be interpreted as the existence of strong systemic yield risk under severe weather conditions and severe weather can affect mean yield more significantly.

Alternatively, we can estimate the mean function without allowing the weather effects δ to vary with time. The parameter estimations of all candidate models under this mean function specification are presented in Table 3.11. When the weather coefficients are assumed to be constant over time, all models indicate positive ZNDX effects, while most of the parameter estimates for July Temperature (TEMP) are negative. The estimated weather coefficients $\hat{\boldsymbol{\delta}}$ in the mean function (equation 3.8) refer to the marginal impacts

that the conditioning weather variables have on the county-level yield risk. For example, the weather coefficient estimate in the DSAR model means an increase from normal value of ZNDX by 1 is associated with an increase in production by 0.33. As for the TEMP variable, the estimate shows that an increase in Temperature in July is associated with a decrease in production by 0.697. This estimation results support the idea that the moisture and cool summer weather is helpful to increase the average corn production.

Also in Table 3.11, we present the estimates of spatial dependence parameters in year 2005. The Omni-directional SAR and CAR models yield spatial dependence parameters of $\rho_s = 0.74$ and $\rho_c = 0.99$. We found directional spatial effects in the data by using directional SAR and CAR model. Both the directional models indicates a stronger spatial dependence in the W-E direction than in N-S. For example, in the DSAR model, the spatial dependence in W-E direction is $\rho^{(2)} = 0.85$ which is 54% higher than that in N-S direction ($\rho^{(1)} = 0.55$). In the DCAR model, the spatial dependence in W-E direction $\rho^{(2)} = 0.999$ is greater than that in the N-S direction $\rho^{(1)} = 0.954$. The time-varing spatial dependence parameter estimates from the autoregressive models are illustrated in Figure 3.6. The spatial dependence parameters in CAR model are very close to zero and show less variation with time compared to those in the SAR model. Both the directional DSAR and DCAR model show a strong directional effect and indicate an overall stronger spatial dependence in W-E direction than in N-S direction over the 82 years as illustrated in Figure 3.6(d) and 3.6(b).

What do these spatial autoregressive models say about the spatial correlations? In the model specifications section (section 3.2), we examine the correlation structure of yield risk $\mathbf{Y}_t$ for the Omni-directional SAR, CAR and Directional SAR and CAR models. In general, the variance-covariances of joint yield distribution of I counties at year t are a function of their spatial dependence parameters ρ_s, ρ_c, or $\rho^{(1)}, \rho^{(2)}$ as defined in equations 3.6,3.10,3.12 and 3.13. These variance-covariances can be standardized to

define a correlation matrix and the implied correlations between any pair of counties can be calculated based on the estimated ρ_s, ρ_c, or $\rho^{(1)}, \rho^{(2)}$ and the neighbor matrix $\mathbf{W}$. Figure 3.7 shows histograms of all the neighbor correlations implied by the spatial autoregressive models in 2005. The correlations range from 0.01 to 0.95 for SAR and DSAR models, while the correlations range from 0.18 to 0.95 for CAR and DCAR models.

Here we select seven contiguous counties in the western part of Illinois to illustrate the idea of spatial correlation and systemic yield risk. Woodford county is among Illinois top ten corn producers and has six neighboring counties, among which La Salle and Marshall counties are in its N-S neighborhood; Peoria, Tazewell, Mclean, Livingston counties are in its W-E neighborhood. The pairwise spatial correlations of these seven contiguous counties are presented in the first panels of Table 3.2, 3.3, 3.4,3.5. For comparison purpose, seven random counties in Illinois are selected from further apart and the pairwise correlations are calculated for these counties Adams, Alexander, Bond, Boone, Brown, Bureau and Calhoun. As an example, the pairwise correlation implied from the DSAR model as presented in table 3.3 shows that the smallest correlation in 2005 is 0.578 among the seven contiguous counties, which occurs between La Salle county (FIPs ID: 17009) and Tazewell (FIPs ID: 17179) as they are separated by Woodford county from N-S direction. The largest correlation among the seven contiguous counties in 2005 is 0.832, which occurs between W-E directional neighbors of Marshall county (FIPs ID: 17123) and La Salle county (FIPs ID: 17009).

The pairwise spatial correlations in all these tables implied from SAR, CAR, DSAR and DCAR models demonstrate similar spatial dependence patterns. First, the implied pairwise correlation between any two random counties is found to be smaller than those of contiguous counties. Second, the overall spatial correlations in a regular year such as 2007 are smaller than those of drought years such as 2005, which implies that the yield

shortfalls of different spatially correlated counties are most likely to happen together when there are more severe weather conditions. This is the justification for the existence of systemic risk in agricultural production and crop insurance.

Furthermore, Figures (3.8) and (3.9) illustrate that the extent of spatial correlation may be state-dependent which is consistent to the study of Goodwin (2001). In particular, yields tend to be more highly correlated in severe weather years such as the 1988 drought, 1991 drought, 1993 flood and 2005 drought, in which the correlations decay slowly with distance and reach zero with a distance of 300 miles between counties. In a normal yield year such as 2007, the correlations reach zero with a distance of 100 miles between counties, which is a much faster decay in spatial correlation. This is evidence that the extent of systemic risk may arise in catastrophic yield years.

We need to choose the range of suggested models by model comparison methods. Notice that when $\rho^{(1)} = \rho^{(2)} = \rho$, the Directional autoregressive model reduces to the Omni-directional autoregressive model. The regular SAR and CAR models are thus nested in the DSAR and DCAR models respectively. In this case, we can use model comparison methods such as AIC, BIC to select the best model. Table 3.11 shows the model comparison results of AIC, BIC and LLF which suggest DSAR is the best fit.

3.5 Economic Implication

The primary goal of modeling spatial dependence of county-level yield risk is to construct models which can measure the systemic yield risk of agricultural production due to agronomic conditions, and then use these joint probabilities of systemic risk to determine actuarially fair insurance rates and design reinsurance contracts between different spatial correlated counties. From the estimation results of previous sections, we found substantial spatial correlation among counties in adverse weather events. The

presence of spatial systemic risk has important agricultural insurance and reinsurance market implications. In particular, the government insurance and reinsurance agencies can use this geographic information to rate the insurance contracts and offer reinsurance across spatialy heterogeneous regions to transfer and diversify risk among geographic regions to reduce the effects of systemic risk. In our analysis, an actuarially-fair premium and indemnity payments directly or indirectly depend on yield risk measured at several locations. Appropriate measuring and quantifying of the spatially correlated yield risk will help design instruments for transferring risk in crop insurance based on the spatial information of the yield risk. Also, accurately modeled contiguous county information will help construct insurance contracts and will help define the reasonable range and shape of the geographical area in the design of area-based insurance and reinsurance contract. For that purpose, the probability distribution of the joint yield losses estimated from a spatial model can be used to determine the total loss of the insurers.

3.5.1 Rate Calculation

This study explores the spatial dependence structure of crop yields between counties, which is important for the development of the group risk plan (GRP)[3] of crop insurance. Yield–based insurance policies in the federal crop insurance program include the individual, farm-level multiple peril crop insurance (MPCI) and the county-level Group Risk Plan (GRP), which is based upon county-average yields from NASS. An important policy parameter in the GRP program is the premium rate, which is based on the county-average yield distribution. In this section, we evaluate the economic impacts of adopting rates based on the spatial dependence models. The premium rates from the spatial model approach are illustrated with simulated data.

[3]The Group Risk Plan (GRP) insurance protects producers against a widespread crop failure. It insures against county production risk. If the average yield for the county in which the insured crop is located falls below the trigger level chosen, the producer receives a payment, regardless of the farm's individual yield.

Standard crop yield insurance pays an indemnity at a predetermined price to replace yield losses. Under the GRP, insured farmers collect an indemnity when the county average yield falls beneath a guarantee, regardless of the farmers' actual yields. An actuarially-fair premium is determined from the expected loss, which is calculated by using the joint probabilities of $\mathbf{Y}_t$ as defined in the previously defined spatial autoregressive models. The expected loss at time t is as follows:

$$\text{E(Loss)} = E[(\lambda y_t^e - y_t)I(y_t \leq \lambda y_t^e)]p = E[(\lambda y_t^e - y_t)^+]p \tag{3.16}$$

where $a^+ = \max(0, a)$ for a number $a \in \mathbb{R}$. In the preceding discussion, $\mathbf{Y}_t$ denotes the observed annual county level yield at year t and y_t^e represents the predicted (guaranteed) yield. The expected loss can be calculated from the joint distribution of yields estimated from the models of SAR, CAR, DSAR , DCAR and the non-spatial random coefficient model, and the random coefficient spatial model.

By running the simulation from the estimated spatial autoregressive models, the expected loss and premium rates can be calculated. For example, under the DSAR distribution, simulated yields are generated from $N((\boldsymbol{\mu}_t, \Sigma_{DSAR,t}))$ for any time t. The probability of joint yield loss and county-specific yield loss, the expected loss, and the actuarially fair premium rate associated with a contract that guarantees 90 percent of the expected yield are calculated and presented in the last two panels of Table 3.6. The premium rates calculated for the seven contiguous counties in the west Illinois range from 0.11% to 0.39%. If we design an area-based yield insurance by grouping these seven contiguous counties together, the premium rate to cover all seven regions is 0.007%. As illustrated in the previous section and Table 3.2, there is strong evidence in the spatial dependence in this seven contiguous counties (pairwise correlations range from 0.538 to 0.806). Therefore, it is hard to diversify the yield risks within the neighborhood due to

the systemic correlated risk. In contrast, the premium rates calculated for the randomly selected counties have higher individual county-specific risk (ranging from 0.042% to 0.653%) but a lower group premium rate (0.07%) than those of contiguous counties, as the spatial dependence is lower among these random selected counties. This significant difference within an individual county and across counties suggests that the premiums rates strongly depend on the spatial dependence of the underlying risk.

Ignoring systemic spatial correlation may result in under-estimating the yield risk as illustrated in the rate simulation from the non-spatial random coefficient model (Table 3.8). We found that for each individual county-specific risk estimation, the premium rates range from 0.12% to 0.13%, which are similar to those from the DSAR model for the same seven contiguous counties, while the grouping premium rate (0.00092%) is much smaller than that from the DSAR model (0.007%). Area-based insurance and reinsurance design based on non-spatial model structure will have losses especially the systemic spatial risk is actually strong.

3.5.2 Design of the Insurance and Reinsurance Contract by Directional Spatial Information

If we want to design a crop insurance contract with large spatial regions, this Directional autoregressive model can be used to capture the systemic spatial correlations which may be varying by direction. The estimated spatial effect on neighborhood and directions in which the area crop yields are significantly spatially correlated will give implications to the range and the shape of the geographical area to be considered in measuring yield risk and design of reinsurance and the area-yield index for the GRP. The implied spatial correlation between different counties should help accurately measure the systemic yield risk and thus improve the efficiency of risk management.

Appropriate measuring and quantifying of spatial risk can make it possible for

designing the area-yield insurance, options and reinsurance based on the degree of the systemic spatial heterogeneity in yield risk. Based on the spatial information of yield risk, there are several possible instruments that can be considered to allow handling of systemic risk. For example, spatial information will help in making the decision of geographic reinsurance. It can also be applied to model weather risk and make inferences for weather derivatives. By taking into account the spatial dependence of weather risk and geographical basis risk, the hedging effectiveness of weather derivatives can be improved. Also, an insurance company may try to spatially diversify systemic weather risk by optimal design of their trading area both in size and by direction.

3.6 Conclusion

Appropriately measuring and quantifying spatial risk is important for risk management due to the systemic nature of agricultural production shock. When selling insurance contracts in the U.S, an insurer will face systemic risk in crop yields, which is a big problem for the provision of crop insurance by the private sectors. Area-based insurance and reinsurance offer a possible solution. It is crucial for both the insurer and reinsurer to understand and quantify systemic spatial risk in order to evaluate the insurability, reinsuring across spatially heterogeneous areas.

Location, time, and direction are all important to explain the systemic change of yield risk. In this study, we have examined the interrelationships among these components of yield risk and explored and quantified systemic risk by using various spatial autoregressive models and non-spatial random coefficient models. We extended the SAR model into dynamic SAR and CAR models by allowing the spatial parameters to change over time and direction. We also develop and evaluate the directional effect of systemic risk by allowing the spatial dependence in the dynamic SAR and CAR models

to vary with direction.

The temporal effect is considered and the spatial correlation and premium in regular and drought years are compared to justify the different degree of systemic risk under different conditions. We found that the yield shortfalls of different spatially correlated counties are most likely to happen together when there are more severe weather conditions. This is the justification for the existence of systemic risk in agricultural production and crop insurance.

A strong directional effect is also found in the crop yield and indicates an overall stronger spatial dependence in the W-E direction than in the N-S direction over years. In particular, the direction of W-E appears to be more important than the direction of N-S in explaining the spatial dependence of crop yield risk. The rate simulation based on the proposed spatial models indicates that premium rates based on the underlying yield risk show moderate spatial dependence at different regions in the U.S. Thus, it is possible to reduce overall risk exposure by reinsurance or grouping more counties across different regions based on the spatial correlation information. This spatial systemic risk estimation offers the potential to improve the global and regional reinsurance and transferring production risk between regions so that the systemic risk problem can be alleviated.

An extension of the spatial analysis will go to the spatial correlation for different crops, e.g.,corn and soybeans, by using a copula approach in the Dynamic spatial autoregressive model. The bivariate distribution obtained in the copula analysis allows for an analysis of the higher-order spatial relationships and higher-order cross-moments of crop yields. These can be used in the design of area-yield insurance contracts that account for the spatial correlation of different crop yields.

Figures and Tables

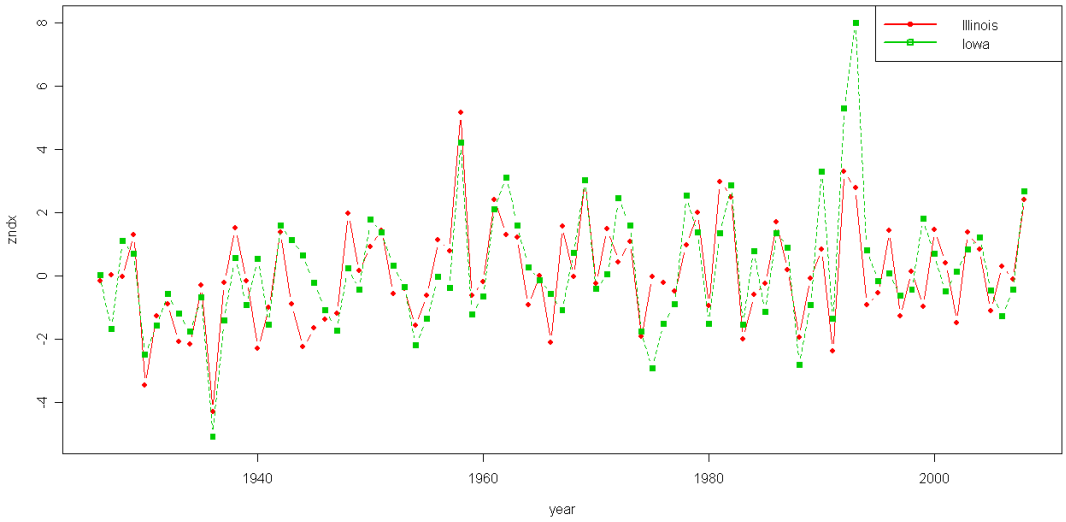


Figure 3.3: Illinois and Iowa Temporal Zndx

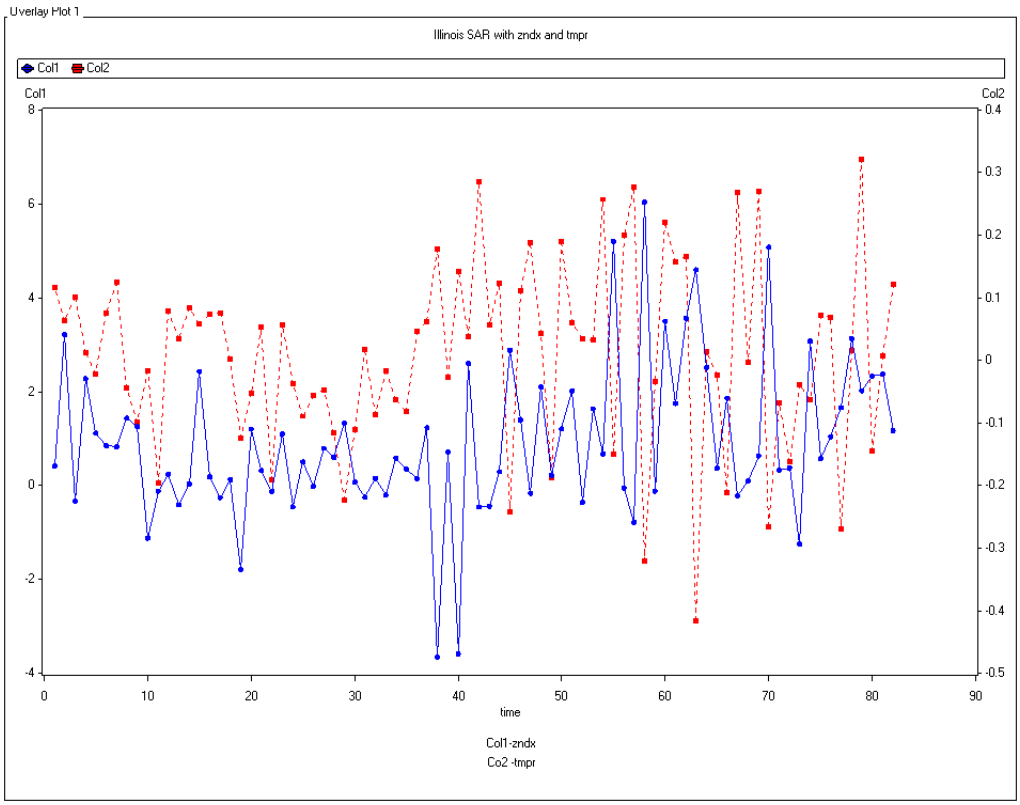


Figure 3.4: Coefficients of Weather implied from DSAR Model
(IL County-Level Data)

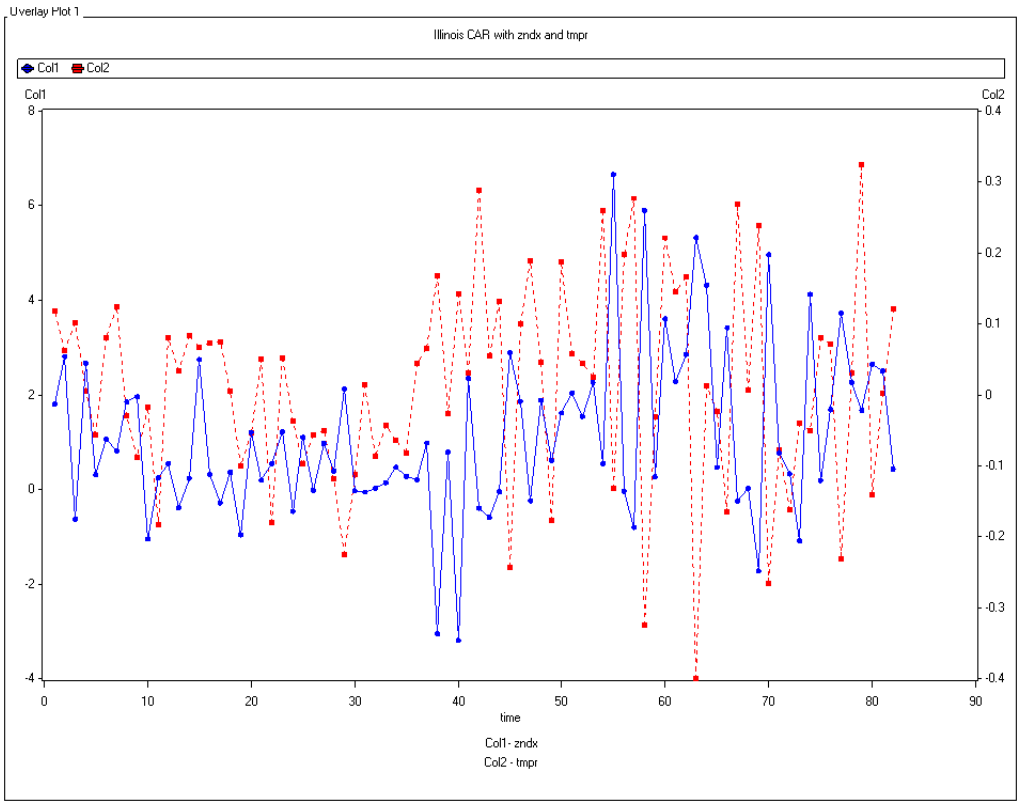
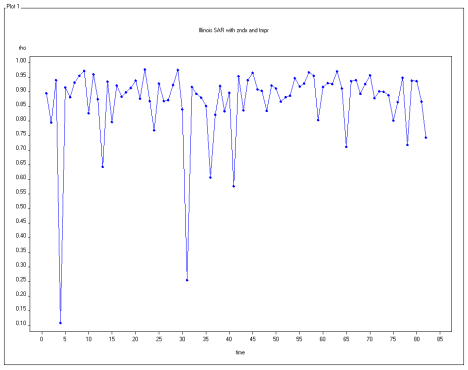
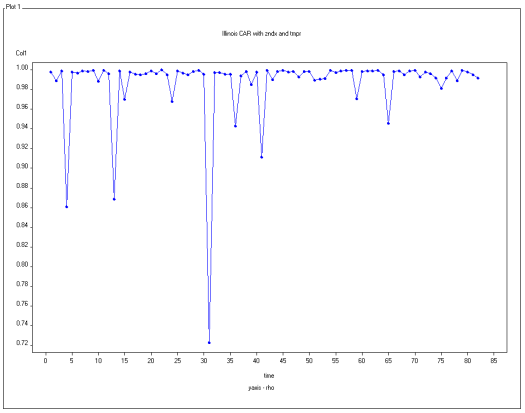


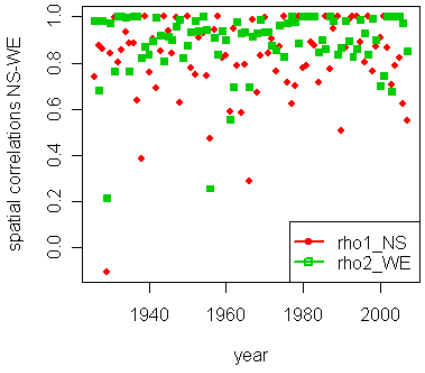
Figure 3.5: Coefficients of Weather implied from DCAR Model(IL County-Level Data)



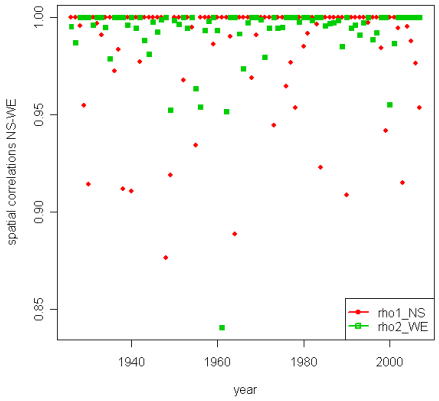
(a) Temporal Plot of the Dependence Parameters Implied by dynamic SAR Model—MLE Estimates (IL County-Level Data



(c) Temporal Plot of the Dependence Parameters Implied by dynamic CAR Model—MLE Estimates (IL County-Level Data

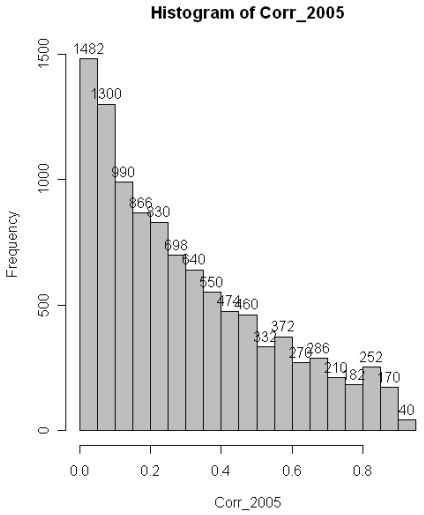


(b) Temporal Plot of the Dependence Parameters Implied by the Directional DSAR Model—MLE Estimates (IL County-Level Data

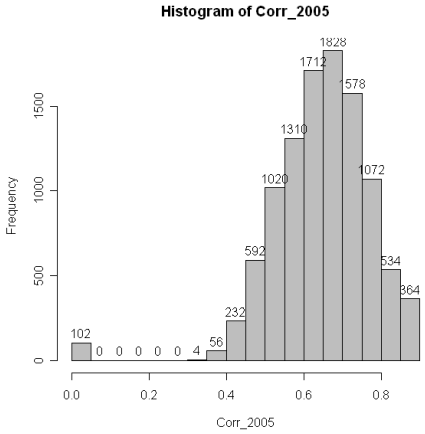


(d) Temporal Plot of the Dependence Parameters Implied by the Directional DCAR Model—MLE Estimates (IL County-Level Data

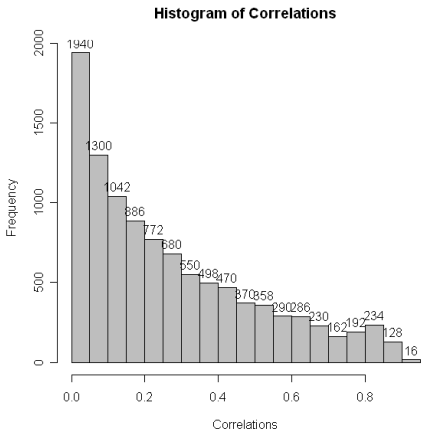
Figure 3.6: Temporal Plot of the Dependence Parameters Implied by Different Spatial Models



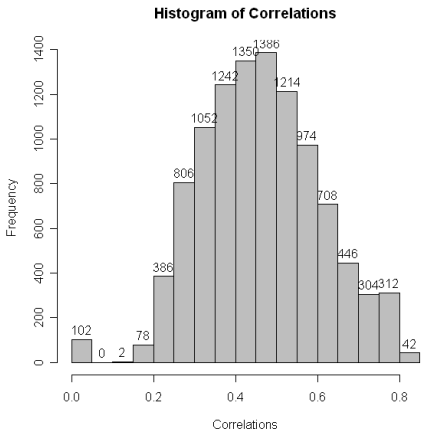
(a) Histogram of SAR Implied Neighbor Correlations (2005)



(c) Histogram of CAR Implied Neighbor Correlations (2005)



(b) Histogram of DSAR Implied Neighbor Correlations (2005)



(d) Histogram of DCAR Implied Neighbor Correlations (2005)

Figure 3.7: Histogram of Implied Neighbor Correlations (IL County-Level Corn Yield Data)

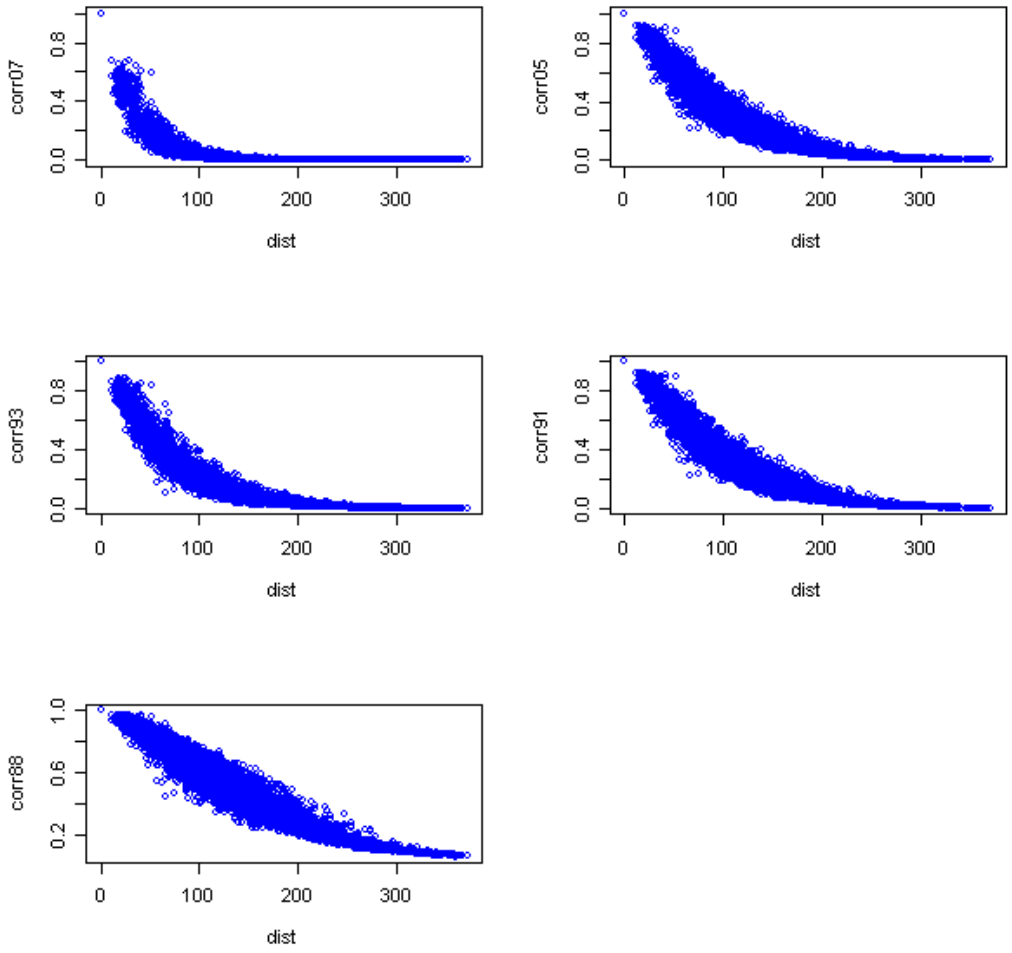


Figure 3.8: Spatial Correlations vs. Distance (IL County-Level Data)—Normal and Extreme Yield Years

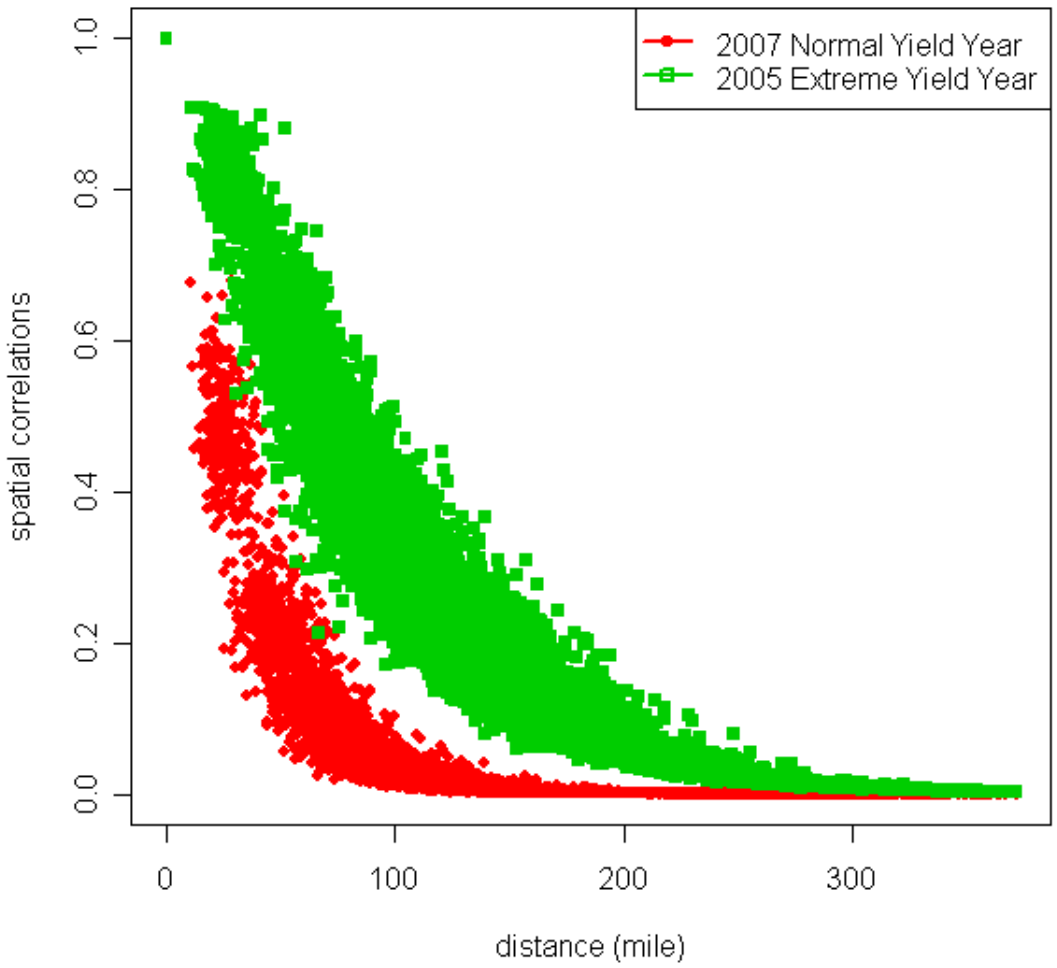


Figure 3.9: State-Dependent Spatial Correlations vs. Distance (IL County-Level Data)

Table 3.2: The Dependence Structure of the Neighborhood Implied by the SAR Model with Constant Weather Coefficients

	La Salle 17099	Livingston 17105	Mclean 17113	Marshall 17123	Peoria 17143	Tazewell 17179	Woodford 17203
..... **Contiguous Counties during Regular Year (2007)**							
17099	1	0.394	0.174	0.467	0.231	0.151	0.400
17105	0.394	1	0.399	0.203	0.158	0.187	0.409
17113	0.174	0.399	1	0.146	0.186	0.387	0.387
17123	0.467	0.203	0.146	1	0.456	0.212	0.438
17143	0.231	0.158	0.186	0.456	1	0.428	0.438
17179	0.151	0.187	0.387	0.212	0.428	1	0.419
17203	0.400	0.409	0.387	0.438	0.438	0.419	1
..... **Counties Counties during Drought Year (2005)**							
17099	1	0.780	0.579	0.832	0.680	0.578	0.791
17105	0.780	1	0.766	0.651	0.596	0.617	0.789
17113	0.579	0.766	1	0.555	0.601	0.757	0.752
17123	0.832	0.651	0.555	1	0.819	0.647	0.813
17143	0.680	0.596	0.601	0.819	1	0.794	0.808
17179	0.578	0.617	0.757	0.647	0.794	1	0.786
17203	0.791	0.789	0.752	0.813	0.808	0.786	1

..... **Random Selected Counties during Drought Year (2005)** .....

	ADAMS 17001	Alexander 17003	BOND 17005	BOONE 17007	BROWN 17009	BUREAU 17011	CALHOUN 17013
17001	1	0.042	0.306	0.141	0.867	0.290	0.615
17003	0.042	1	0.209	0.008	0.051	0.015	0.070
17005	0.306	0.209	1	0.062	0.361	0.114	0.434
17007	0.141	0.008	0.062	1	0.130	0.600	0.073
17009	0.867	0.051	0.361	0.130	1	0.265	0.681
17011	0.290	0.015	0.114	0.600	0.265	1	0.150
17013	0.615	0.070	0.434	0.073	0.681	0.150	1

Table 3.3: The Spatial Dependence Structure of the Neighborhood Implied by the Directional SAR Model with Constant Weather Coefficients

	La Salle 17099	Livingston 17105	Mclean 17113	Marshall 17123	Peoria 17143	Tazewell 17179	Woodford 17203
..... **Contiguous Counties during Regular Year (2007)**							
17099	1	0.322	0.139	0.453	0.217	0.131	0.325
17105	0.322	1	0.364	0.176	0.160	0.193	0.414
17113	0.139	0.364	1	0.134	0.198	0.422	0.419
17123	0.453	0.176	0.134	1	0.462	0.197	0.376
17143	0.217	0.160	0.198	0.462	1	0.404	0.453
17179	0.131	0.193	0.422	0.197	0.404	1	0.454
17203	0.325	0.414	0.419	0.376	0.453	0.454	1
..... **Contiguous Counties during Drought Year (2005)**							
17099	1	0.731	0.530	0.814	0.651	0.538	0.744
17105	0.731	1	0.736	0.613	0.571	0.595	0.766
17113	0.530	0.736	1	0.524	0.591	0.758	0.748
17123	0.814	0.613	0.524	1	0.806	0.619	0.780
17143	0.651	0.571	0.591	0.806	1	0.780	0.800
17179	0.538	0.595	0.758	0.619	0.780	1	0.785
17203	0.744	0.766	0.748	0.780	0.800	0.785	1

..... **Random Selected Counties during Drought Year (2005)** .....

	ADAMS 17001	Alexander 17003	BOND 17005	BOONE 17007	BROWN 17009	BUREAU 17011	CALHOUN 17013
17001	1	0.033	0.276	0.114	0.854	0.257	0.558
17003	0.033	1	0.180	0.005	0.040	0.010	0.060
17005	0.276	0.180	1	0.043	0.329	0.089	0.422
17007	0.114	0.005	0.043	1	0.105	0.554	0.053
17009	0.854	0.040	0.329	0.105	1	0.236	0.628
17011	0.257	0.010	0.089	0.554	0.236	1	0.120
17013	0.558	0.060	0.422	0.053	0.628	0.120	1

Table 3.4: The Dependence Structure of the Neighborhood Implied by the CAR Model with Constant Weather Coefficients

	La Salle 17099	Livingston 17105	Mclean 17113	Marshall 17123	Peoria 17143	Tazewell 17179	Woodford 17203
..... **Counties Counties during Drought Year (2005)**							
17099	1	0.862	0.818	0.875	0.830	0.806	0.864
17105	0.862	1	0.854	0.806	0.796	0.801	0.852
17113	0.818	0.854	1	0.794	0.810	0.855	0.853
17123	0.875	0.806	0.794	1	0.857	0.803	0.856
17143	0.830	0.796	0.810	0.857	1	0.852	0.857
17179	0.806	0.801	0.855	0.803	0.852	1	0.851
17203	0.864	0.852	0.853	0.856	0.857	0.851	1

	ADAMS 17001	Alexander 17003	BOND 17005	BOONE 17007	BROWN 17009	BUREAU 17011	CALHOUN 17013
..... **Random Selected Counties during Drought Year (2005)**							
17001	1	0.450	0.653	0.595	0.844	0.681	0.713
17003	0.450	1	0.555	0.383	0.471	0.432	0.454
17005	0.653	0.555	1	0.546	0.683	0.616	0.664
17007	0.595	0.383	0.546	1	0.604	0.767	0.532
17009	0.844	0.471	0.683	0.604	1	0.690	0.744
17011	0.681	0.432	0.616	0.767	0.690	1	0.605
17013	0.713	0.454	0.664	0.532	0.744	0.605	1

Table 3.5: The Dependence Structure of the Neighborhood Implied by the Directional DCAR Model with Constant Weather Coefficients

	La Salle 17099	Livingston 17105	Mclean 17113	Marshall 17123	Peoria 17143	Tazewell 17179	Woodford 17203
..... **Counties Counties during Drought Year (2005)**							
17099	1	0.757	0.679	0.784	0.705	0.662	0.761
17105	0.757	1	0.742	0.669	0.649	0.656	0.742
17113	0.679	0.742	1	0.645	0.668	0.743	0.740
17123	0.784	0.669	0.645	1	0.757	0.664	0.752
17143	0.705	0.649	0.668	0.757	1	0.743	0.752
17179	0.662	0.656	0.743	0.664	0.743	1	0.741
17203	0.761	0.742	0.740	0.752	0.752	0.741	1

	ADAMS 17001	Alexander 17003	BOND 17005	BOONE 17007	BROWN 17009	BUREAU 17011	CALHOUN 17013
..... **Random Selected Counties during Drought Year (2005)**							
17001	1	0.246	0.454	0.385	0.748	0.489	0.561
17003	0.246	1	0.371	0.180	0.266	0.221	0.258
17005	0.454	0.371	1	0.320	0.493	0.394	0.485
17007	0.385	0.180	0.320	1	0.393	0.628	0.315
17009	0.748	0.266	0.493	0.393	1	0.496	0.603
17011	0.489	0.221	0.394	0.628	0.496	1	0.394
17013	0.561	0.258	0.485	0.315	0.603	0.394	1

Table 3.6: The Yield Insurance Implied from the Yield Distributions of SAR and DSAR(2005)

	Mean Yield	Liability	E(Loss)	Premium Rate (%)
..... **Contiguous Counties during Drought Year (2005)—SAR**				
17099	168.450	151.605	1.859	1.226
17105	166.753	150.077	1.837	1.224
17113	177.127	159.415	1.327	0.832
17123	169.278	152.351	2.109	1.384
17143	168.378	151.541	1.875	1.237
17179	174.401	156.961	1.451	0.924
17203	174.200	156.780	1.543	0.984
all together*	1198.589	1078.730	7.924	0.735
..... **Random Selected Counties during Drought Year (2005)—SAR**				
17001	157.0462	141.342	2.079	1.471
17003	151.345	136.211	2.166	1.590
17005	129.6049	116.644	2.222	1.905
17007	154.1794	138.762	2.424	1.747
17009	155.8036	140.223	2.154	1.536
17011	169.8676	152.881	1.564	1.023
17013	149.5823	134.624	2.011	1.494
all together*	1067.429	960.686	10.030	1.044
..... **Contiguous Counties during Drought Year (2005)—DSAR**				
17099	169.873	152.886	0.041	0.027
17105	167.367	150.631	0.035	0.023
17113	178.137	160.323	0.018	0.011
17123	170.412	153.370	0.057	0.037
17143	169.444	152.500	0.059	0.039
17179	175.302	157.772	0.035	0.022
17203	175.177	157.659	0.029	0.018
all together*	1205.711	1085.140	0.075	0.007
..... **Random Selected Counties during Drought Year (2005)—DSAR**				
17001	156.992	141.293	0.165	0.117
17003	152.796	137.516	0.897	0.653
17005	130.649	117.584	0.247	0.210
17007	155.903	140.313	0.328	0.234
17009	156.106	140.495	0.153	0.109
17011	171.182	154.063	0.065	0.042
17013	149.654	134.688	0.394	0.293
all together*	1073.281	965.953	0.049	0.005

* Note: "all together" indicates the joint risk of all seven counties

Table 3.7: The Yield Insurance (90% Coverage) Implied from the Yield Distributions of CAR and DCAR(2005)

	Mean Yield	Liability	E(Loss)	Premium Rate(%)
..... **Contiguous Counties during Drought Year (2005)—CAR**				
17099	163.0835	146.7751	2.31211	1.575
17105	161.6593	145.4934	2.407724	1.655
17113	172.099	154.8891	1.982064	1.280
17123	164.1024	147.6921	2.454017	1.662
17143	163.2861	146.9575	2.389018	1.626
17179	169.4581	152.5123	2.149053	1.409
17203	169.2422	152.318	2.196987	1.442
all together*	1162.931	1046.637	13.087	1.250
..... **Random Selected Counties during Drought Year (2005)—CAR**				
17001	154.0916	138.6824	0.05042813	0.036
17003	150.3496	135.3146	0.062	0.046
17005	131.7023	118.532	0.0883656	0.075
17007	149.6531	134.6878	0.08010004	0.059
17009	152.3384	137.1045	0.06472221	0.047
17011	164.8597	148.3738	0.03466763	0.023
17013	147.9971	133.1974	0.06562378	0.049
all together*	1050.992	945.8926	0.163542	0.017
..... **Contiguous Counties during Drought Year (2005)—DCAR**				
17099	163.0835	146.7751	0.03577361	0.024
17105	161.6593	145.4934	0.0432868	0.030
17113	172.099	154.8891	0.0176799	0.011
17123	164.1024	147.6921	0.04858031	0.033
17143	163.2861	146.9575	0.04391998	0.030
17179	169.4581	152.5123	0.0284479	0.019
17203	169.2422	152.318	0.03068695	0.020
all together*	1162.931	1046.637	0.076	0.007
..... **Random Selected Counties during Drought Year (2005)—DCAR**				
17001	154.0916	138.6824	0.1099998	0.079
17003	150.3496	135.3146	0.501	0.370
17005	131.7023	118.532	0.1805277	0.152
17007	149.6531	134.6878	0.1768469	0.131
17009	152.3384	137.1045	0.0883672	0.064
17011	164.8597	148.3738	0.04540845	0.031
17013	147.9971	133.1974	0.2036184	0.153
all together*	1050.992	945.8926	0.1030064	0.011

* Note: "all together" indicates the joint risk of all seven counties

Table 3.8: The Yield Insurance (90% Coverage) Implied from the Nonspatial Model

	Mean Yield	Liability	E(Loss)	Premium Rate (%)
..... **Contiguous Counties during Regular Year (2007)—Non-spatial**				
17099	168.7	151.8	0.162	0.1069
17105	168	151.1	0.1606	0.1065
17113	179.6	161.8	0.1483	0.0918
17123	170.7	153.8	0.1573	0.1025
17143	168.8	152	0.1649	0.1087
17179	175.9	158.4	0.1525	0.0965
17203	174.8	157.3	0.1509	0.0961
all together*	1206.5	1086	0.004303	0.000397
..... **Random Selected Counties during Regular Year (2007)—Non-spatial**				
17001	155.5	139.9	0.1835	0.1314
17003	153.2	137.9	0.1856	0.1349
17005	133.8	120.4	0.2144	0.1785
17007	162.4	146.2	0.1681	0.1152
17009	155.6	140.1	0.1806	0.1291
17011	173.2	155.9	0.1525	0.0981
17013	147.3	132.5	0.1915	0.1447
all together*	1081	972.9	0.008758	0.000901
..... **Contiguous Counties during Drought Year (2005)—Non-spatial**				
17099	156.5	140.8	0.1855	0.1321
17105	156.2	140.5	0.1835	0.1308
17113	163.6	147.4	0.1743	0.1185
17123	156.1	140.6	0.1846	0.1316
17143	154.4	139	0.1862	0.1341
17179	161.1	144.9	0.1758	0.1215
17203	159.9	143.9	0.1776	0.1236
all together*	1107.8	997.1	0.009121	0.0000915
..... **Random Selected Counties during Drought Year (2005)—Non-spatial**				
17001	142.3	128.1	0.2079	0.1624
17003	153.3	132	0.112	0.0850
17005	133.8	111.3	0.09718	0.0874
17007	162.4	140	0.09945	0.0710
17009	155.5	128.3	0.05833	0.0455
17011	159.9	144	0.182	0.1266
17013	138.2	124.4	0.2143	0.1727
all together*	1045.4	908.2	0.001515	0.000167

* Note: "all together" indicates the joint risk of all seven counties

Table 3.9: The Yield Insurance (90% Coverage) Implied from the Random Coefficient Spatial Model

	Mean Yield	Liability	E(Loss)	Premium Rate (%)
..... **Contiguous Counties during Regular Year (2007)**				
17099	171	154.1	0.174	0.1128
17105	168	151.6	0.142	0.0938
17113	180.5	162.6	0.142	0.0873
17123	173.7	155.7	0.144	0.0926
17143	170.5	153.6	0.144	0.0940
17179	176.5	159.3	0.138	0.0868
17203	175.2	158.2	0.156	0.0987
all together*	1215.4	1095	0.002	0.00018
..... **Random Selected Counties during Regular Year (2007)**				
17001	157.7	142	0.168	0.1128
17003	159.2	143.3	0.19	0.0938
17005	129.6	117.3	0.248	0.0873
17007	161.4	145.1	0.14	0.0926
17009	157.5	141.6	0.162	0.0940
17011	176.9	158.5	0.102	0.0868
17013	149	134.5	0.196	0.0987
all together*	1091.3	982.3	0.002	0.000180
..... **Contiguous Counties during Drought Year (2005)**				
17099	162.6	146	0.198	0.1361
17105	162.5	145.8	0.186	0.1279
17113	170.3	152.3	0.184	0.121
17123	160.4	146.8	0.246	0.1679
17143	160	144.1	0.202	0.1401
17179	164.2	148.7	0.22	0.148
17203	166	148.7	0.194	0.131
all together*	1146	1032	0.016	0.00156
..... **Random Selected Counties during Drought Year (2005)**				
17001	145.1	132.2	0.252	0.1904
17003	153	137.5	0.23	0.1673
17005	124.3	111.6	0.276	0.2471
17007	154.6	140.1	0.226	0.1609
17009	143.9	130.7	0.27	0.2064
17011	165.6	149.4	0.202	0.135
17013	143.3	129	0.226	0.1754
all together*	1029.8	930.6	0.034	0.00364

* Note: "all together" indicates the joint risk of all seven counties

Table 3.10: Maximum-Likelihood Parameter Estimates (Time-Varying Weather Effect) and Summary Statistics for Candidates Models: Example for Illinois Corn Yields*

	SAR	**CAR**	**DSAR**	**DCAR**
			Illinois 2007	
zndx	1.21	0.43	1.21	0.38
temp	0.13	0.12	0.13	0.24
spatial ρ	0.750	0.985	$\rho^{(1)} = 0.592, \rho^{(2)} = 0.861$	$\rho^{(1)} = 0.983, \rho^{(2)} = 0.999$
variance	469.66	434.40	92.21	86.45
			Illinois 2005	
zndx	2.51	2.65	2.51	2.44
temp	-0.19	-0.14	-0.19	-0.17
spatial ρ	0.839	0.991	$\rho^{(1)} = 0.999, \rho^{(2)} = 0.749$	$\rho^{(1)} = 0.999, \rho^{(2)} = 0.979$
variance	313.65	341.49	62.84	72.75
			Illinois 1993	
zndx	0.11	0.02	0.11	-0.17
temp	-0.04	0.01	-0.04	-1.31
spatial ρ	0.874	0.995	$\rho^{(1)} = 0.982, \rho^{(2)} = 0.794$	$\rho^{(1)} = 0.999, \rho^{(2)} = 0.983$
variance	246.22	251.26	51.18	54.55
			Illinois1991	
zndx	1.85	3.42	1.85	3.47
temp	-0.19	-0.16	-0.19	0.09
spatial ρ	0.877	0.990	$\rho^{(1)} = 0.916, \rho^{(2)} = 0.825$	$\rho^{(1)} = 0.996, \rho^{(2)} = 0.999$
variance	464.03	521.56	86.49	90.09
			Illinois 1988	
zndx	6.64	5.33	6.64	7.80
temp	-0.37	-0.40	-0.37	-0.34
spatial ρ	0.852	0.986	$\rho^{(1)} = 0.337, \rho^{(2)} = 0.999$	$\rho^{(1)} = 0.679, \rho^{(2)} = 0.999$
variance	377.04	399.87	75.84	86.95
obs. #	8364	8364	8364	8364
param. #	574.00	574.00	656.00	738.00
AIC/n	17323.87	24210.89	17192.39	17207.78
BIC/n	17323.94	24210.96	17192.47	17207.86
log likelihood	-17323.80	-24210.82	-17192.31	-17207.69

* Notes: ρ_s is the spatial dependence parameter in DSAR model.
ρ_c is the spatial dependence parameter in DCAR model.
$\rho^{(1)}$ is the spatial dependence parameter in the North-South direction.
$\rho^{(2)}$ is the spatial dependence parameter in the West-East direction.

Table 3.11: Maximum-Likelihood Parameter Estimates (Constant Weather Effect) and Summary Statistics for Candidates Models: Example for Illinois Corn Yields*

	SAR	CAR	DSAR	DCAR	Random Spatial	Random Nonspatial
ZNDX	0.32	0.40	0.33	1.21	2.02	1.15
TEMP	-0.76	-0.81	-0.70	-1.50	-0.02	0.27
spatial ρ	$\rho_s = 0.74$	$\rho_c = 0.99$	$\rho^{(1)} = 0.55$, $\rho^{(2)} = 0.85$	$\rho^{(1)} = 0.95$, $\rho^{(2)} = 0.99$	$\rho_{zndx} = 0.99$,$\rho_{temp} = 0.94$	—
s.e.	0.075	0.011	(0.189),(0.140)	(0.167),(0.119)	(0.010),(0.096)	—
variance	459.565	429.966	90.077	84.174	1.246, 72.65	—
obs. #	8364	8364	8364	8364	8364	8364
param. #	412	412	494	494	—	—
AIC/n	17662.31	24399.40	17541.81	17679.69	$DIC = 62035.8$	$DIC = 62674.6$
BIC/n	17662.36	24399.45	17541.87	17679.75	—	—
log likelihood	-17662.26	-24399.35	-17541.75	-17679.63	-31017.90	-31337.30

* Notes: ρ_s is the spatial dependence parameter in DSAR model.
ρ_c is the spatial dependence parameter in DCAR model.
$\rho^{(1)}$ is the spatial dependence parameter in the North-South direction.
$\rho^{(2)}$ is the spatial dependence parameter in the West-East direction.

Conclusions

Agricultural risks are typically assumed to originate from unanticipated movements in prices, yields and revenues. Risk averse farmers may employ strategies to deal with risks by managing their agricultural practices. The federally regulated crop insurance programs are very important risk management tools for farmers and have played an important role to convey economic benefits to the agricultural sector. Problems exist in US crop insurance programs due to asymmetric information and systemic risk. Losses attributed to asymmetric information and systemic risk can be potentially reduced by properly designed risk management instruments. However, if the underlying risk is not accurately measured in the construction of premium rates, the problems in the crop insurance market may not be mitigated. In order to better understand and evaluate agricultural risk management mechanisms and government policies to facilitate access to risk management instruments, it is important to measure the risk in a quantitative fashion; that is, to estimate, assess, and manage risk in the agricultural production process. The primary goal of this dissertation was to accurately measure the underlying risk and predict the premium rates, thus improving agricultural risk management. Noticing that these agricultural risk factors are not isolated from each other, the focus of this research is on investigating the dependence structure and the interaction of the agricultural production risks.

This dissertation explored the dependence structure of both yield and revenue risk in agricultural production. Using time-series and panel data from the Corn-Belt, we investigate the role that different dependence structures might play in crop insurance markets by evaluating the risk dependence structures from different dimensions (as specified in Chapters 1, 2 and 3). All three approaches used statistical methods to model dependence structures of the distributions. The first method employed a

time-varying distribution method to capture the evolution of distributions under the technology changes; the second method employed a copula method to capture the joint distribution of multivariate risks. The third approach used a spatial autoregressive model to investigate the anisotropic and time-varying features of the systemic spatial correlation. In addition, NASS county-level yields, prices and the relationships between spatial heterogeneity in crop yields and the relevance of correlation relationships among prices and reported crop yields are investigated. The role of modeling dependence and the county-level price and yield forecasts in the federal crop insurance program is also evaluated.

Chapter 1 evaluates and models the yield risk associated with agricultural production by applying a new time-varying distribution method. Yield insurance contract development requires properly estimating the distribution of crop yields. Proper representation of the yield distributions using data observed over time may be complicated by technological progress. Advances in technology, germplasm, breeding techniques, and the development of new hybrids may significantly affect the distributions of crop yields. Hence, various moments of the distributions may evolve over time as technology progresses. This study investigated how changes in technology affect the evaluation of yield risk by employing a time-varying modeling approach to capture dynamic changes in yield distributions. The estimates showed that the time-varying model can improve the accuracy of yield distribution estimates compared to the more conventional detrended model. This analysis extends a long line of research in this area and offers significant improvements in the accuracy of the models used in rating yield insurance contracts.

The second chapter investigated the dependence structure between price and yield of multiple agricultural commodities. To estimate the joint loss probability of crop yields and prices, the joint distribution of crop yields and prices must be known. Motivated by

this observation, a copula approach is used to investigate yield and price dependence of corn and soybeans under a Whole Farm Insurance policy. This study demonstrates the practical implementation and estimation of a joint price and yield distribution by fitting a copula to crop prices and yields. Consistent with diversification/portfolio theory, the estimation results implied that the Whole Farm Insurance contracts are more efficient as risk management tools than crop-specific contracts. Evidence of the portfolio effect that exists in the design of revenue insurance contracts is presented and its potential in hedging revenue risk among different commodities is discussed.

Catastrophic events induce systemic correlation in agricultural yields and thus potentially prevent diversification in insurance markets. Private insurers often lack adequate capacity to underwrite systemic risks. This calls for the reinsurance and other risk transfer methods, which are sometimes provided by the federal government. To address and measure the spatial dependence structure of systemic yield risk in order to make implications for risk management, in Chapter 3, we develop a spatial autoregressive model by using less restrictive simultaneous autoregressive (SAR) model and conditional autoregressive (CAR) model which allow the spatial dependence parameters to change over time and by direction. The evidence of systemic risk under different production scenarios is presented. We also found that state-contingent spatial correlation exists, and yields tend to demonstrate higher correlation in extreme low yield years than in normal yield years. Based on the appropriate measure and quantification of systemic spatial risk, this study showed that it is possible to design efficient and affordable instruments for transferring the risk among geographic regions to reduce the effects of systemic risk.

A contribution of this study is the investigation and evaluation of dependence structures in the design of crop insurance contracts; highlighting their use at the dynamic, portfolio, spatial level for risk transfer and management. By modeling and measuring dependence, it is possible to improve risk management instruments

that take advantage of dependencies between different products. In this manner, temporal and spatial markets may transfer the correlated risks associated with high-risk products/periods/regions to low-risk products/periods/regions.

Future research should focus on applying the models proposed in this study to farm-level data. Future research could also explore alternative policy structures for implementing new risk management tools for agriculture. For example, the spatial model can be used to model weather index insurance.

BIBLIOGRAPHY

Akaike, H. 1974. "A New Look at Statistical Model Identification." *IEEE Transactions on Automatic Control* 19:716–723.

Anselin, L. 2002. "Under the hood Issues in the Specification and Interpretation of Spatial Regression Models." *Agricultural Economics* 27:247–267.

Anselin, L., and R. Florax. 1995. *New Directions in Spatial Econometrics*. Springer, Berlin.

Atwood, J., S. Shaik, and M. Watts. 2003. "Are Crop Yields Normally Distributed? A Re-examination." *American Journal of Agricultural Economics* 85(4):888–901.

Brook, D. 1964. "On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems." *Biometrika* 51:481–483.

Burnham, K.P., and D. Anderson. 2003. *Model Selection and Multi-Model Inference*. Springer.

Chen, S.L., and M. Miranda. Aug. 2004. "Modeling Multivariate Crop Yield Densities with Frequent Extreme Events." *AAEA meeting*, pp. 1–16.

Coble, K., R. Heifner, and M. Zuniga. 2000. "Implications of Crop Yield and Revenue Insurance for Producer Hedging." *Journal of Agricultural and Resource Economics* 25(2):432–452.

Day, R. 1965. "Probability Distributions of Field Crop Yields." *Journal of Farm Economics* 47:713–741.

Frees, E.W., and E.A. Valdez. 1999. "Understanding Relationships Using Copulas." *North American Actuarial Journal* 2:1–25.

Gallagher, P. 1987. "U.S. Soybean Yields: Estimation and Forecasting with Nonsymmetric Distributions." *American Journal of Agricultural Economics* 69:798–803.

Genest, C., and J. MacKay. 1986. "The Joy of Copulas: Bivariate Distribution with Uniform Marginals." *The American Statistician* 40(4):280–283.

Genest, C., and L.P. Rivest. 1993. "Statistical inference procedures for bivariate Archimedean copulas." *American Statistics Association* 88:1034–1043.

Genest, C., and F. Verret. 2005. "Locally Most Powerful Rank Tests of Independence for Copulas Models." *Nonparametric Statistics* 17:521–535.

Goodwin, B.K. 2001. "Problems with Market Insurance in Agriculture." *American Journal of Agricultural Economics* 83(3):643–649.

Goodwin, B.K., and A.P. Ker. 2001. "Modeling Price and Yield Risk." *Chapter in A Comprehensive Assessment of the Role of Risk in U.S. Agriculture*, pp. 289–323.

—. 1998. "Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk (GRP) Crop Insurance Contracts." *American Journal of Agricultural Economics* 80:139–53.

Goodwin, B.K., M.C. Roberts, and K.H. Coble. 2000. "Measurement of Price Risk in Revenue Insurance: Implications of Distributional Assumptions." *Journal of Agricultural and Resource Economics* 25(1):195–214.

Gyorfi, L., M. Kohler, A. Krzyzak, and H. Walk. 2002. *A Distribution-Free Theory of Nonparametric Regression*. Springer.

Harri, A., C. Erdem, K.H. Coble, and T.O. Knight. 2009. "Crop Yield Normality: A Reconciliation of Previous Research and Statistical Tests for Normality." *Review of Agricultural Economics* 31:163–182.

Hennessy, D., B. Babcock, and D. Hayes. 1997. "Budgetary and Producer Welfare Effects of Revenue Insurance." *American Journal of Agricultural Economics* 79:1024–34.

Huard, D., G. Évin, and A.C. Favre. 2006. "Bayesian Copula Selection." *Computational Statistics & Data Analysis* 51:809–822.

Iman, R.L., and W.J. Conover. 1982. "A Distribution-free Approach to Inducing Rank Correlation Among Input Variables." *Communications in Statistics* 11:311–334.

Johnson, M., and A. Tenenbein. 1981. "A Bivariate Distribution Family with Specified Marginals." *Journal of the American Statistical Association* 76:198–201.

Just, R.E., and Q. Weninger. 1999. "Are Crop Yields Normally Distributed?" *American Journal of Agricultural Economics* 81:287–304.

Kelejian, H., and I. Prucha. 1999. "A Generalized moments Estimator for the Autoregressive Parameter in a Spatial Model." *International Economic Review* 40(2):509–533.

Ker, A.P., and B.K. Goodwin. 2000. "Nonparametric Rating of Group Risk Insurance Contracts Revisited." *American Journal of Agricultural Economics* 82:463–478.

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

Morduch, J. 1995. "Income Smoothing and Consumption Smoothing." *Journal of Economic Perspectives* 3:103–14.

Nelsen, R.B. 2006. *An Introduction to Copulas*. Springer.

Nelson, C. 1990. "The Influence of Distribution Assumptions of the Calculation of Crop Insurance Premia." *North Central Journal of Agricultural Economics* 12:71–78.

Nelson, C., and P. Preckel. 1989. "The Conditional Beta Distribution as a Stochastic Production Function." *American Journal of Agricultural Economics* 71(2):370–378.

Ozaki, V.A., S.K. Ghosh, B.K. Goodwin, and R. Shirota. 2008. "Spatio-Temporal Modeling of Agricultural Yield Data with an Application to Pricing Crop Insurance Contracts." *American Journal of Agricultural Economics* 90(4):951–961.

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

Ramirez, O., S. Misra, and J. Field. 2003. "Crop-yield Distributions Revisited." *American Journal of Agricultural Economics* 85(1):108–120.

Robinson, P. 1987. "Asymptotically efficiency estimation in the presence of heteroscedasticity of unknown form." *Econometrika* 55:875–891.

Schabenberger, O., and C.A. Gotway. 2005. *Statistical methods for spatial data analysis*. Chapman & Hall/CRC.

Schwarz, G. 1978. "Estimating the dimension of a model." *The Annals of Statistics* 6:461–464.

Sherrick, B., F. Zanani, G. Schnitkey, and S. Irwin. 2004. "Crop Insurance Valuation under Alternative Yield Distribution." *American Journal of Agricultural Economics* 86(2):406–419.

Skees, J.R., J.R. Black, and B.J. Barnett. 1997. "Designing and Rating an Area Yield Crop Insurance Contract." *American Journal of Agricultural Economics* 79(2):430–438.

Sklar, A. 1959. "Fonctions de rpartition n dimensions et leur marges." *Publications de lInstitut de Statistique de lUniversit de Paris* 8:229–231.

Swinton, S., and R. King. 1991. "Evaluating Robust Regression Techniques for Detrending Crop Yield Data with Non-Normal Errors." *American Journal of Agricultural Economics* 73:446–461.

Taylor, C. 1990. "Two Practical Procedures for Estimating Multivariate Nonnormal Probability Density Functions." *American Journal of Agricultural Economics* 72:210–217.

Wall, M.M. 2004. "A Close Look at the Spatial Structure Implied by the CAR and SAR Models." *Journal of Statistical Planning and Inference* 12:311–324.

Whittle, P. 1954. "On stationary processes in the plane." *Biometrika* 41:434449.

Wilks, S. 1938. "The Large-sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses." *The Annals of Mathematical Statistics* 9:554–560.

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

APPENDICES

APPENDIX A—Statistics Estimation Algorithm for SAR Model

The algorithm of SAR model estimation is described as follows:

(a) Let $\hat{\sigma_t}^2 = c$ for any constant number $c > 0$

(b) Solve β and σ_t^2 by

$$vec\left(\hat{\beta}\right) = \left(\sum_t \frac{1}{\hat{\sigma}_t^2}\left(\Phi\left(t\right)' \otimes I_I\right)'(I - \rho_t W')D^{-1}(I - \rho_t W)\left(\Phi\left(t\right)' \otimes I_I\right)\right)^{-1} \cdot \left(\sum_t \frac{1}{\hat{\sigma}_t^2}\left(\Phi\left(t\right)' \otimes I_I\right)'(I - \rho_t W')D^{-1}(I - \rho_t W)Y_t\right) \quad (17)$$

and

$$\hat{\sigma}_t^2 = \frac{1}{I}(\mathbf{Y}_t - \left(\Phi\left(t\right)' \otimes I_I\right) vec\left(\hat{\beta}\right))'(I - \rho_t W')D^{-1}(I - \rho_t W)(\mathbf{Y}_t - \left(\Phi\left(t\right)' \otimes I_I\right) vec\left(\hat{\beta}\right))$$

Solve ρ_t by maximize

$$\sum_t -2\log|(I - \rho_t W)| + \sum_t \frac{1}{\hat{\sigma}_t^2}(\mathbf{Y}_t - \left(\Phi\left(t\right)' \otimes I_I\right) vec\left(\beta\right))'(I - \rho_t W')D^{-1}(I - \rho_t W) (\mathbf{Y}_t - \left(\Phi\left(t\right)' \otimes I_I\right) vec\left(\beta\right))$$

within the interval $(1/v_{min}(\mathbf{W}), 1/v_{max}(\mathbf{W}))$.

(c) iterate between (b) and (c) until convergence or stablility.

APPENDIX B—Statistics Estimation Algorithm for CAR Model

The algorithm of CAR model estimation is described as follows:

Suppose $\{\rho_t\}_{t=1}^T$ and $\{\sigma_t\}_{t=1}^T$ are given, we solve for β. Taking the derivative w.r.t. $vec\,(\beta)$, we get

$$\begin{aligned} 0 &= \sum_t \exp\left(-\log\sigma_t^2\right)\left(\Phi\,(t)'\otimes I_I\right)'D^{-1}(I-\rho_t W)\left[\mathbf{Y}_t-\left(\Phi\,(t)'\otimes I_I\right)vec\,(\beta)\right] \\ &= \sum_t \exp\left(-\log\sigma_t^2\right)\left(\Phi\,(t)'\otimes I_I\right)'D^{-1}(I-\rho_t W)Y_t \\ &\quad -\sum_t \exp\left(-\log\sigma_t^2\right)\left(\Phi\,(t)'\otimes I_I\right)'D^{-1}(I-\rho_t W)\left(\Phi\,(t)'\otimes I_I\right)vec\,(\beta) \end{aligned}$$

Solving for $vec\left(\hat{\beta}\right)$:

$$\begin{aligned} vec\left(\hat{\beta}\right) = &\left(\sum_t \exp\left(\log\sigma_t^2\right)\left(\Phi\,(t)'\otimes I_I\right)'D^{-1}(I-\rho_t W)\left(\Phi\,(t)'\otimes I_I\right)\right)^{-1} \\ &\cdot\left(\sum_t \exp\left(-\log\sigma_t^2\right)\left(\Phi\,(t)'\otimes I_I\right)'D^{-1}(I-\rho_t W)Y_t\right) \end{aligned} \tag{18}$$

If we set $\sigma_t^2 = c$ for any constant $c > 0$, then β and σ can be estimated by

$$\begin{aligned} vec\left(\hat{\beta}\right) = &\left(\sum_t \frac{1}{\hat{\sigma}_t^2}\left(\Phi\,(t)'\otimes I_I\right)'D^{-1}(I-\rho_t W)\left(\Phi\,(t)'\otimes I_I\right)\right)^{-1} \\ &\cdot\left(\sum_t \frac{1}{\hat{\sigma}_t^2}\left(\Phi\,(t)'\otimes I_I\right)'D^{-1}(I-\rho_t W)Y_t\right) \end{aligned} \tag{19}$$

and

$$\hat{\sigma}_t^2 = \frac{1}{I}\Big(\mathbf{Y}_t-\left(\Phi\,(t)'\otimes I_I\right)vec\,(\beta)\Big)'D^{-1}(I-\rho_t W)\Big(\mathbf{Y}_t-\left(\Phi\,(t)'\otimes I_I\right)vec\,(\beta)\Big)$$

ρ_t can be estimated by minimizing

$$\{-\log|I-\rho_t W|\}$$
$$+\left\{\frac{1}{\hat{\sigma}_t^2}\Big(\mathbf{Y}_t-\left(\Phi\left(t\right)'\otimes I_I\right)vec\left(\beta\right)\Big)'D^{-1}(I-\rho_t W)\Big(\mathbf{Y}_t-\left(\Phi\left(t\right)'\otimes I_I\right)vec\left(\beta\right)\Big)\right\}$$