2011

Three essays on weather and crop yield

Tian Yu
Iowa State University

Three essays on weather and crop yield

by

Tian Yu

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Bruce A. Babcock, Major Professor
Dermot J. Hayes
John C. Beghin
Catherine L. Kling
Chad E. Hart

Iowa State University

Ames, Iowa

2011

TABLE OF CONTENTS

LIST OF TABLES v

LIST OF FIGURES vii

ACKNOWLEDGEMENTS ix

ABSTRACT x

CHAPTER 1. GENERAL INTRODUCTION 1

1.1 Introduction 1

1.2 Organization of the Dissertation 3

CHAPTER 2. ARE U.S. CORN AND SOYBEANS BECOMING MORE DROUGHT TOLERANT? 6

2.1 Introduction 6

2.2 Crop Yield and the Drought Index 9

2.2.1 Drought Index 10

2.2.2 Data Selection 12

2.3 Drought Tolerance and Yield Risk 15

2.3.1 Modeling Crop Yield with the Drought Index 17

2.3.2 Impact of Drought 20

2.3.3 Estimation Results 21

2.4 Implications for GRP Rates 33

2.5 Conclusions 38

CHAPTER 3. ESTIMATING NON-LINEAR WEATHER IMPACTS ON CORN YIELD – A BAYESIAN APPROACH 41

3.1 Introduction 41

3.2 The Yield Model 43

3.3 The Bayesian Approach 46

3.3.1 Priors 46

3.3.2 The Likelihood Function 48

3.3.3 Conditional Posteriors 48

3.3.4 Implementing the Gibbs Sampling and the Metropolis - Hastings Algorithm 49

3.4 Data 51

3.5 Estimation Results 52

3.5.1 The Marginal Effect 58

3.5.2 Features of Weather Effects 60

3.5.3 The One-knot Specification 64

3.6 Conclusions 67

CHAPTER 4. WEATHER EFFECTS ON TREND, VARIANCE, AND DISTRIBUTION OF CORN YIELD 69

4.1 Introduction 69

4.2 A Yield Model with Weather Factors 74

4.2.1 The Yield Model 74

4.2.2 Data 76

4.2.3 Estimation Results 77

4.3 Weather Effects on Yield Trend 81
4.3.1 Estimating the Bias 82
4.4 Weather Effects on Yield Risk 86
4.4.1 Limitations in Previous Studies 88
4.4.2 Improving Estimation by Incorporating Weather Effects 90
4.5 Weather Effects on Yield Distribution 94
4.5.1 A Special Case 97
4.5.2 Relaxing the Assumptions 99
4.5.3 Empirical Results 101
4.6 Conclusions 104
CHAPTER 5. GENERAL CONCLUSIONS 106
5.1 Conclusions 106
BIBLIOGRAPHY 109

LIST OF TABLES

Table 2.1 Distribution of Drought across States and in Each Decade 13
Table 2.2 CRD-Specific Trend Estimates γ_r's from Aggregate Models . . . 22
Table 2.3 Point Estimates and Robust Standard Errors of CRD-Invariant Coefficients 23
Table 2.4 Hypothesis Test Results 25
Table 2.5 Marginal Effects with Drought Index Evaluated at $DI = 1$ . . . 26
Table 2.6 Annual Change in Crop Losses at Alternative Drought Severity Levels 27
Table 2.7 Point Estimates of Coefficients from CRD-Specific Models 29
Table 2.8 Point Estimates of Coefficients from CRD-Specific Models - Continued 30
Table 2.9 F-test of Equal Coefficients 32
Table 2.10 Model Selection Based on Bayesian Information Criterion (BIC) 33
Table 2.11 Actual and Simulated Premium Rates 37
Table 3.1 Posterior Mean and Standard Deviation of Trend and Variance Parameters 53
Table 3.2 Posterior Mean and Standard Deviation of Parameters of Weather Impacts 54

Table 3.3 Comparing Thresholds Across States 57

Table 3.4 Marginal Effects of Temperature and Rainfall 59

Table 3.5 Estimation Results for the One-Knot Specification, Indiana . . . 65

Table 4.1 Regression Results of the Yield Model with Weather Effects . . . 78

Table 4.2 Marginal Effects of Temperature and Rainfall 80

Table 4.3 Percentage Differences between Trend Estimates from Models with or without Weather Factors (1980-2009) 83

Table 4.4 Percentage Differences between Trend Estimates from Models with or without Weather Factors (1990-2002) 85

Table 4.5 Relationship between b and Hypothesis Testing 92

Table 4.6 Hypothesis Testing on Yield Risk 93

Table 4.7 Skewness Statistics 101

LIST OF FIGURES

Figure 2.1 Relationship between the drought index and relative corn yield deviation 12

Figure 2.2 Histogram of drought indices in the 11 CRDs in IL and IN truncated below 0.1 15

Figure 2.3 Crop losses and drought indices 16

Figure 2.4 Model-predicted drought-induced crop losses 28

Figure 3.1 Plots of weather impacts on corn yield 61

Figure 3.2 Plots of weather impacts on corn yield - continued 62

Figure 3.3 Plots for the one-knot specification 66

Figure 4.1 Histograms of residuals from regressing yield on a time trend, by state 95

Figure 4.2 Histograms of growing season average rainfall, by state 96

Figure 4.3 Histograms of growing season average temperature, by state . . . 97

Figure 4.4 Concave transformation and negative skewness 99

Figure 4.5 Histograms of weather-induced yield deviates, by state 102

Figure 4.6 Histograms of residuals from regressing yield on a time trend and weather variables, by state 103

Figure 4.7 Histograms of weather-induced yield deviates plus residuals, by state 103

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this dissertation.

First and foremost, I sincerely thank my major professor, Bruce Babcock, for his wise guidance and unlimited supply of time and patience. His insights and words of encouragement have often inspired me and renewed my hopes for completing my graduate education. I would also like to thank my committee members for their efforts and contributions to this work: John Beghin, Chad Hart, Dermot Hayes, and Catherine Kling.

I would additionally like to thank David Hennessy, John Schroeter, and GianCarlo Moschini for making themselves available each and every time I needed their help and advice. I would like to thank the faculty and staff of the Center for Agricultural and Rural Development and the Economics Department for the help I received during my graduate studies.

ABSTRACT

The general theme of this dissertation is the study of impacts of weather variability on crop yields, with each chapter addressing a specific topic related to this theme. Chapter 2 tests the hypothesis that corn and soybeans have become more drought tolerant by regressing county yields on a drought index and time. Results indicate that corn yield losses from drought of a given severity, whether measured in quantity terms or as a percentage of mean yields, have decreased over time. Soybean percentage yield losses have also declined but absolute losses have remained largely constant. The potential impact of increased drought tolerance on U.S. crop insurance rates is illustrated by comparing Group Risk Plan (GRP) premium rates assuming time-invariant susceptibility to drought with rates generated from regression results in this dissertation. Chapter 3 develops a linear spline model with endogenous knots to capture the non-linear impacts of rainfall and temperature on corn yields. A hierarchical structure is applied to capture the county-specific factors determining corn yields. Using Bayesian techniques, the thresholds and other model parameters are simultaneously estimated. Gibbs sampling and the Metropolis - Hastings algorithm are applied to estimate the posterior distributions. Corn yield decreases significantly above the upper temperature threshold and below the lower rainfall threshold. Results indicate a geographically clustering pattern of how corn yields respond to changes in temperature and rainfall. Chapter 4 applies the linear spline yield model developed in chapter 3 to examine weather impacts on yield trend, yield risk, and the distribution of corn yield. The climate trend from 1980 to 2009 explains up to 20% of observed yield trend. Not controlling for temporal weather patters leads to biased trend estimates, especially for short times series. Isolating changes in weather variability in the sample period, the hypothesis of constant coefficient of variation is rejected in most states in the Corn Belt. Decreasing marginal benefit of weather partly explains why corn yield is negatively skewed. Conditional on weather, the distribution of unexplained residuals from our yield model is symmetric in general.

CHAPTER 1. GENERAL INTRODUCTION

1.1 Introduction

In rain-fed agricultural regions, weather conditions have substantial impacts on crop productivity. Favorable weather conditions for dryland crop production, including a proper amount of heat and rainfall during the growing season, are critical factors determining yield outcomes. There is a large body of existing literature on estimating the weather impacts on crop yields. Methodologies used in most of these studies, however, are relatively restrictive. One of the objectives of this dissertation is to develop a more realistic model to measure weather impacts on yield and to improve the estimation methodology. With a better understanding of how weather factors influence crop productivity, the estimation of the yield trend and how yield risk evolves over time will be improved. A sound estimation of the yield trend is key to answering the question of whether crop yields will increase enough to meet increasing feed and fuel demand. A proper estimation of the temporal change in yield risk is critical for determining the actuarially fair rates of crop insurance programs. Understanding how crop yields respond to weather variability provides an explanation to the long-standing puzzle of why corn yield is negatively skewed. Motivated by these economic issues, the second goal of this dissertation is to incorporate weather impacts into the analysis of productivity gains, yield risk, the distribution of yield, and crop insurance rating.

With the wide use of genetically modified seeds in recent years, claims of increased drought resistance in corn hybrids are being made. One implication of increasing drought tolerance would be that U.S. crop insurance rates are too high, because most crop insurance rating methods are based on the assumption that the coefficient of variation of yields is constant over time. Such an assumption would be difficult to defend if the degree of drought tolerance is changing because drought is the primary source of loss in the U.S. crop insurance program. Attempts have been made to modify this assumption by measuring changes in yield variance over time without controlling for weather conditions. A drawback of relying solely on yield realizations is that the temporal incidence of drought could be determining whether the coefficient of variation of yields has increased or decreased over time because the distribution of major U.S. droughts is not constant over time in any region. In testing the hypothesis of increasing drought tolerance we avoid possible spurious correlation by controlling for the incidence of drought over time and space. We construct an objective measure of drought severity and regress corn yield on time, the drought measure, and the interaction term between the two. Based on the estimated change in drought tolerance, we suggest a modification to the assumptions and methods that underlie current crop insurance rates. We demonstrate that such modifications could lead to substantial reductions in both corn and soybean crop insurance rates. This topic is the focus of one of the following chapters.

Another chapter is devoted to improving the estimation of weather impacts, in at least three aspects. First, the nonlinear impacts of rainfall and temperature are separately modeled in a linear-spline model with endogenous thresholds. This improves upon the existing literature, most of which makes the restrictive assumption that crop yields are either linear or quadratic in weather variables. Our model reflects the concept of 'growing degree days' in the sense that the marginal effect of weather variables should be allowed to vary across different weather conditions. Further, the thresholds that divide the weather conditions are estimated rather than fixed as in the 'growing degree days' approach. Second, jointly estimated with weather effects is a linear trend with a random county effect and a fixed effect that varies across crop reporting districts. Previous studies of weather impacts lack the careful modeling of potentially different yield trends across regions. Third, a sampling-based Bayesian estimation procedure is developed to estimate the linear-spline model with a hierarchical structure. The complexity in estimation that usually comes with the nonlinearity is circumvented in the Bayesian estimation framework because the Markov chain Monte Carlo (MCMC) simulation works with conditional distributions and the model is linear conditional on thresholds.

Despite a growing interest in estimating the impact of weather on crop yields, weather factors are seldom modeled in estimating yield trend, yield risk, and distribution of corn yield. There are a few studies that estimate the extent to which estimates of trend and the variance of corn yield would be biased if weather impacts are not taken into account. While the negative skewness of the distribution of corn yield is widely acknowledged, only a couple of studies offer theoretical explanations. There has been no attempt so far to empirically link the negative skewness of corn yield to how weather influences yield outcomes. In this dissertation, a chapter is devoted to estimate the impact of weather and relate it to estimation of trend, variance, and distribution of corn yields.

1.2 Organization of the Dissertation

This dissertation is organized into five chapters. The current chapter presents a general introduction to the chapters that follow and, provides an outline for the organization of the dissertation. While the general theme of this dissertation is weather impacts on crop yields, each chapter is meant to stand alone by addressing a specific topic.

We begin in chapter 2 by taking a look at how drought tolerance of corn and soybeans changed over time in the U.S. We construct a drought index to measure the severity level of the adverse weather conditions. By contrasting the yield lost to a given severity level of drought in the 1980s and in the 2000s, we examine the change in drought tolerance for both corn and soybeans, measured in both the absolute level of yield and as the percentage of trend yield. Counties in Illinois and Indiana that have drought incidents throughout the sample period are selected for a regression analysis. Absolute level of crop yield and the log of crop yield are regressed on time, the drought index, and the interaction term between time and the drought index. The null hypotheses that corn and soybean yield losses due to drought have not changed over time are tested. Based on estimated changes in how corn and soybean yields respond to dry and hot conditions, the actuarial fair rates for the Group Risk Plan (GRP) are simulated. These rates are compared to the insurance rates that assume the percentage yield looses are constant over time.

Chapter 3 turns to improve the estimation methodology of measuring how temperature and rainfall affect corn yields. A flexible linear-spline model with endogenous thresholds is set up to capture nonlinear and asymmetric weather impacts. A hierarchical structure is applied to model the county-specific constant term and the CRD-specific slope of the linear time trend. With prior and likelihood specifications, the conditional posterior distributions are derived. A procedure is developed to simulate the Markov chain Monte Carlo for each model parameter by applying the Gibbs sampling and the Metropolis-Hastings algorithm. The threshold parameters and other model parameters are simultaneously estimated using a panel data of corn yield with matching weather data in eight corn belt states. Marginal effects of rainfall and temperature across different weather conditions are also simulated. The estimated marginal effects as well as plots of model predicted corn yield against rainfall and temperature indicate several features of how weather affects corn yields. Non-linearity, concavity, and geographical clustering are some of the features.

In chapter 4, the yield model developed in chapter 3 is applied to examine how weather plays a role in determining yield trend, yield risk, and the distribution of yield. First, the yield model is briefly reviewed and estimated with a simplified methodology. Second, the effect of climate trend on yield trend is evaluated. In two examples, we calculate the percentage bias of the yield trend estimator if weather factors are not taken into account. Third, weather variability is related to yield risk. The hypothesis of constant coefficient of variation is tested controlling for weather variability. Finally, the negative skewness of corn yield distribution is explained from the perspective of weather. We relate concavity of the yield function to skewness and support it with empirical evidence.

Finally in chapter 5, a summary and some general conclusions are presented.

CHAPTER 2. ARE U.S. CORN AND SOYBEANS BECOMING MORE DROUGHT TOLERANT?

Abstract

An objective drought index that measures the dry and hot conditions adversely affecting crop yields is used in a regression analysis to test whether corn and soybeans have become more drought tolerant. Results indicate that corn yield losses from a drought of a given severity, whether measured in quantity terms or as a percentage of mean yield, have decreased over time. The null hypothesis that the absolute level of soybean yield losses due to drought has not changed cannot be rejected. But soybean yield losses in percentage terms have decreased over time. Because drought is the primary cause of yield loss in the U.S. crop insurance program and because U.S. crop insurance rates assume that percentage yield losses are constant over time, these results indicate that U.S. crop insurance rates in the Corn Belt are too high.

2.1 Introduction

Much effort is being devoted to increasing the drought tolerance of crops. The first results of this effort are expected to be released soon in new corn hybrids (Agriculture Online, 2009; Monsanto, 2009b). But claims of increased drought resistance in corn are already being made (Barrionuevo and Bradsher, 2005; Monsanto, 2009a). The argument is that improved protection from pest damage enables corn plants to better withstand drought conditions, thereby reducing yield losses from drought. The major seedcorn companies (Monsanto and Dupont) convinced USDA's Risk Management Agency (RMA) that their triple-stack hybrids that offer protection against above and below ground insects as well as herbicide resistance, provide lower yield risk than hybrids that were planted in the past. Because U.S. crop insurance rates are based on past loss experience, RMA approved lower rates for farmers who plant the new hybrids. It is not clear if RMA approved these rates only on the basis of improved protection against pest losses or also on the basis of lower risk of crop loss from non-pest sources of loss because of more vigorous plants. If the new hybrids reduce yield losses due to adverse weather conditions such as drought, then the entire basis for RMA's rate-making procedures for insurance products that cover farm-level yield or revenue risk need to be modified because the procedures assume that the coefficient of variation of yields is constant over time (Woodard et al., 2008). Such an assumption would be difficult to defend if the degree of drought tolerance is changing because drought is the primary source of loss in the U.S. crop insurance program.

Premium rates for crop insurance products that insure against area yield losses were, until 2009, also based on the assumption that the coefficient of variation of county yields has not changed over time (Paulson and Babcock, 2008). Although this assumption has been recently modified by RMA, their new procedure relies on measuring yield variance over time without controlling for the incidence of drought. A drawback of relying solely on yield realizations is that the temporal incidence of drought could be determining whether the coefficient of variation of yields has increased or decreased over time because the distribution of major U.S. droughts is not constant over time in any region. For example, in the Corn Belt, a severe multi-state drought has not occurred since 1988 whereas multiple droughts have occurred in Great Plains states since 2000. A model that only looked at how crop losses have varied over time would conclude that losses have declined in the Corn Belt but not in Great Plains states. This would then lead one to the possibly spurious conclusion that yield risk has been declining in the Corn Belt and increasing or staying the same in Great Plains states.

A finding of increased drought tolerance has wider implications than the accurate rating of crop insurance. For example, the ability to meet biofuels mandates with lower disruptions to feed supplies would be enhanced. The payoff from investment in grain storage would likely decrease because the risk of a severe crop shortfall would fall. And, more generally, pre-harvest price volatility would decrease because the risk of significant crop losses would decrease.

In this study we use county yields in major corn belt states to test the hypothesis that U.S. corn and soybeans are becoming more drought tolerant over time. Because corn and soybeans are grown by the same managers, inclusion of both crops allows for some insight into whether better management or better corn hybrids is responsible for any finding that corn yields have become less susceptible to drought. In testing the hypothesis of increasing drought tolerance we avoid spurious correlation by controlling for the incidence of drought over time and space by constructing an objective measure of drought severity. We measure the impact of drought on yield both in quantity (bushels) and percentage (of mean yield) terms. Our results indicate that corn has indeed become more drought tolerant since 1980 both in terms of bushels and percentage losses. In contrast, soybeans has become more drought tolerant in terms of percentage of yield lost, but not in terms of the absolute number of bushels lost to drought. Our findings suggest a need to modify assumptions and methods that underlie current crop insurance rates. We demonstrate that such modifications could lead to substantial reductions in both corn and soybean crop insurance rates.

2.2 Crop Yield and the Drought Index

In this study, we estimate the effects of drought conditions on crop yields. Abnormally hot and dry weather in growing seasons has been shown to be detrimental to crop yields in the U.S. Corn Belt (O'Brien et al., 1996). Hence, throughout the article drought is treated synonymously with hot and dry conditions, rather than just dry conditions. The data used include county-level corn and soybean yields and weather information in Iowa, Illinois, and Indiana from 1980 to 2008.[1] County-level production and planted acreage data were collected from the National Agricultural Statistics Service (NASS) to calculate yield per planted acre. Observations with zero production or missing acreage data were deleted. To focus our attention on major production areas, only counties with yield data in all years from 1980 to 2008 were included.[2]

Cooling degree days (CLDD) and total monthly precipitation (TPCP) data were collected from the National Oceanic and Atmospheric Administration (NOAA). Monthly data from June to August were summed up as a growing season total. For counties with multiple weather stations, we took the simple averages of weather records from all weather stations located in the county to obtain county-level weather data. For each year, yields were matched with county-level weather data. About 20% of the counties do

[1] We attempted to use data back to 1960 but there were insufficient droughts in the 1960s and 1970s to allow estimation of the change in drought tolerance over time. The criteria for data inclusion is discussed in the following section.

[2] About 90% of counties have yield data in all 29 years.

not have weather stations that kept monthly CLDD or TPCP data. We exclude these counties from our sample.

2.2.1 Drought Index

We constructed our drought index from county-level rainfall and temperature data. The index captures the presence of both unusually hot and unusually dry conditions. Severity of hotness (dryness) can be measured by the degree to which temperature (rainfall) departs from the historical normal. The product of the number of standard deviations that the temperature measure reaches above the mean and the rainfall measure falls below the mean reflects the dual impacts of hotness and dryness. The higher (lower) the temperature (rainfall) measure is above (below) the mean, the larger will be the resulting drought index. Formally, the drought index (DI) is defined as

$$DI_{i,t} = [-\max(0, CLDD_{i,t}^{Stand})] \times [\min(0, TPCP_{i,t}^{Stand})]. \tag{2.1}$$

Subscripts i and t denote county and year. Both CLDD[3] and TPCP are standardized by subtracting county means from each observation and then dividing by the county-level standard deviations. The standardizing procedure scales the drought index so that it is comparable across counties and over time.[4]

Standardized cooling degree days and precipitation deviates measure the degree of hotness and dryness. Any mathematical operation of the two standardized weather deviates that strictly increases in each of them could serve as a drought index. The mul-

[3]CLDD sums up degrees above 65°F on a daily basis. It measures the accumulative heat at the right tail of the distribution. We also experimented with drought indices calculated from two other weather measures the number of days with greater than or equal to 0.1 inches precipitation and monthly mean temperature. They give similar regression results.

[4]To eliminate the possibility that county means might be distorted by extreme values in the sample, we replaced simple county means with spatially smoothed county means. This spatial smoothing had minimal impacts on the results.

tiplicative relationship that we use is among the simplest forms and tends to emphasize more significant droughts in the regression analysis.[5]

There are three advantages associated with this drought index. First, it is constructed from county-level weather data, which, compared with state- or country-level aggregates, contains localized weather information more relevant to county yields. Second, this index provides information on the top two yield loss causes for corn and soybeans in the Corn Belt - excess heat and lack of moisture. A single index instead of multiple weather variables provides an easy way to assess the impact of the main weather factors. Finally, as expected, the drought index is correlated with yield deviations, and it also identifies major drought years. To illustrate, figure 2.1 plots relative corn yield deviations against the drought index for counties in Illinois in major drought years: 1983, 1988, 1991, 2002, and 2005.[6] Clearly the data in figure 2.1 exhibit both a strong negative relationship between yield deviation and the level of the drought index as well as substantial sampling error. One source of sampling error is that soil moisture prior to drought is not incorporated because of lack of data. Under the same severity of drought, crop yields in counties with higher pre-drought soil moisture levels are likely to incur smaller yield losses.

[5]We constructed an alternative drought index based on the sum rather than the product of standardized weather deviates. Scatter plots of relative yield deviation against alternative drought indices suggest that our choice of the drought index better represents the harmful impact of weather on crop yields. Estimation results and overall conclusions based on the alternative drought index are similar to results presented here.

[6]Relative yield deviation is defined as actual county yield minus county trend yield, which is then divided by county trend yield.

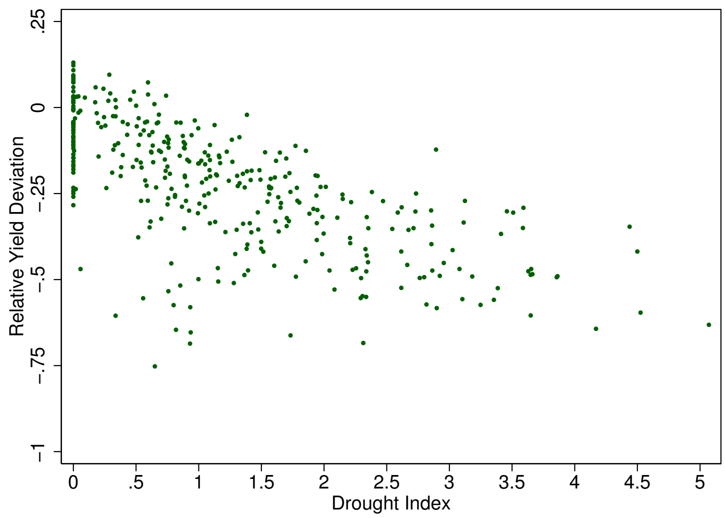

Figure 2.1: Relationship between the drought index and relative corn yield deviation

2.2.2 Data Selection

Corn and soybean yields in Iowa, Illinois, and Indiana are matched with drought indices representing growing season dry/hot conditions in the same county and year of the crop yield. The drought index ranges from zero to five, with zero representing growing season temperature lower than average and/or rainfall greater than average. Table 2.1 shows the distribution of droughts of different severity levels across states and in each decade. Note that incidents of severe droughts decreased substantially in the 1990s and the 2000s relative to the 1980s in all three states. In Iowa, the number of county droughts with a drought index between 1 and 2 in the 1980s is almost 50% greater than the number of droughts from 1990 to 2008. There has been only one drought with a drought index greater than 2 since 1990 in Iowa.

Table 2.1: Distribution of Drought across States and in Each Decade

Period	$DI \in (0.1, 1]$	$(1, 2]$	$(2, 3]$	$(3, 4]$	$(4, 5]$	$(5, 6]$
			Illinois			
1980-1989	104	46	33	20	3	1
1990-1999	85	44	7	0	0	0
2000-2008	110	42	10	1	1	0
			Indiana			
1980-1989	78	37	43	14	4	1
1990-1999	94	48	10	2	0	0
2000-2008	52	35	17	2	0	0
			Iowa			
1980-1989	128	62	47	30	6	0
1990-1999	133	30	0	0	0	0
2000-2008	203	13	1	0	0	0
			11 CRDs in IL and IN			
1980-1989	125	59	46	22	2	0
1990-1999	120	63	12	2	0	0
2000-2008	96	61	25	3	1	0

To accurately measure whether there has been an increase in drought tolerance over time in a region requires observations of yields in drought years throughout the sample period. We would not be able to quantify change of drought tolerance otherwise. Thus, for the analysis presented in the following sections, we keep only those counties in crop reporting districts (CRDs)[7] that have at least two drought incidents with an index larger than 2.0 or at least three drought incidents with an index larger than 1.5 since 2000. These selection criteria leave us with counties in five CRDs in Indiana and six CRDs in Illinois. The total number of counties in the sample is 98. Note that we do not include any Iowa counties in the analysis because there simply have not been any serious droughts in Iowa since 1990.[8]

A histogram of drought indices in the 11 CRDs is presented in figure 2.2. Figure 2.2 indicates that there were only a few droughts of magnitude between 3 and 4 and a small number of droughts of magnitude between 4 and 5. The bottom part of table 2.1 suggests that the 11 CRDs used for regression have a more balanced distribution of droughts over the decades, compared with the overall sample. Still, there were fewer drought incidents of magnitude between 3 and 4 after 1990: 22 out of 27 droughts of magnitude between 3 and 4 occurred in the 1980s. A limited number of droughts of magnitude greater than 3 combined with an uneven temporal distribution of droughts of magnitude between 3 and 4 imply that we need to be cautious about interpreting model estimation and prediction for droughts of magnitude larger than 3 in the analysis that follows.

[7] Crop reporting districts are aggregates of counties used by NASS.

[8] We considered starting the series in 1960. The problem with extending the series further back in time was that there were no major droughts during 1960-1979 in Illinois or Indiana, and only a few droughts of this severity in Iowa that occurred in the 1970s. However, there were no major droughts in Iowa after 2000. There is no crop reporting district that had at least two drought incidents with $DI > 2$ or at least three drought incidents with $DI > 1.5$ during 2000-2008 and also in the pre-1980 period. Thus, we start our sample series from 1980.

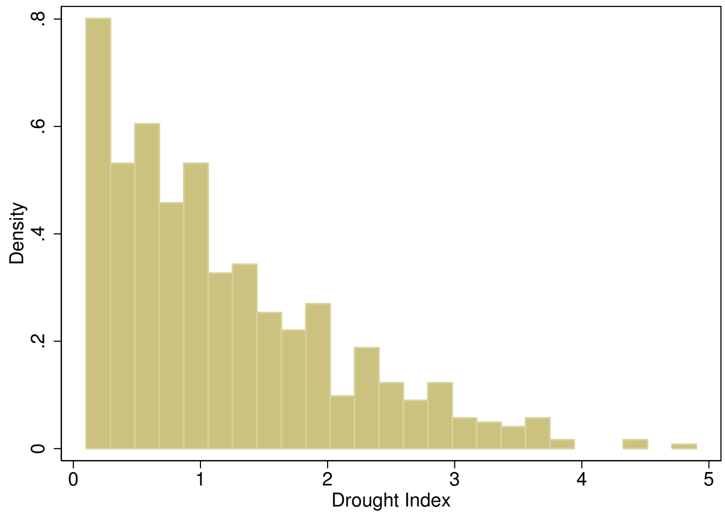

Figure 2.2: Histogram of drought indices in the 11 CRDs in IL and IN truncated below 0.1

2.3 Drought Tolerance and Yield Risk

There are different approaches that could be taken to see if there has been an increase in drought tolerance over time using the crop yield data and matched drought indices. One way is to compare yield losses in drought years in the first decade with yield losses in the last decade for droughts of similar magnitude. We classify droughts into categories of degrees of severity according to indices falling within 0.5-width ranges $(0, .5]$, $(.5, 1]$,..., $(3.5, 4]$, $(4, 4.5]$.[9] Mean yield losses across counties in each drought category from 1980 to 1989 are compared to mean yield losses in the corresponding category from 2000 to 2008. Figure 2.3 shows bar charts of crop losses in each drought category with the mean of drought indices in each category on the x-axis and mean of crop losses on the y-axis. As shown, in most drought categories, corn yield losses measured in both bushels per

[9]The upper limit of the range is 4.5 because there were no droughts of magnitude larger than 4.5 in the last decade.

acre and in percentage of mean yield declined in 2000-2008 compared with the 1980s. Corn yield losses increased when very severe droughts occurred (when the drought index falls into ranges (3,3.5] and (4,4.5]). However, as pointed out in the previous section, caution should be used in drawing conclusions about yield reductions when the drought index reaches beyond 3. In general, data suggests that corn yields have indeed become less susceptible to the effects of drought. The situation for soybeans seems less strong. No clear-cut conclusion can be drawn regarding changes in losses in soybean bushels. Soybean percentage losses decreased in the latter period, but to a lesser degree than did corn.

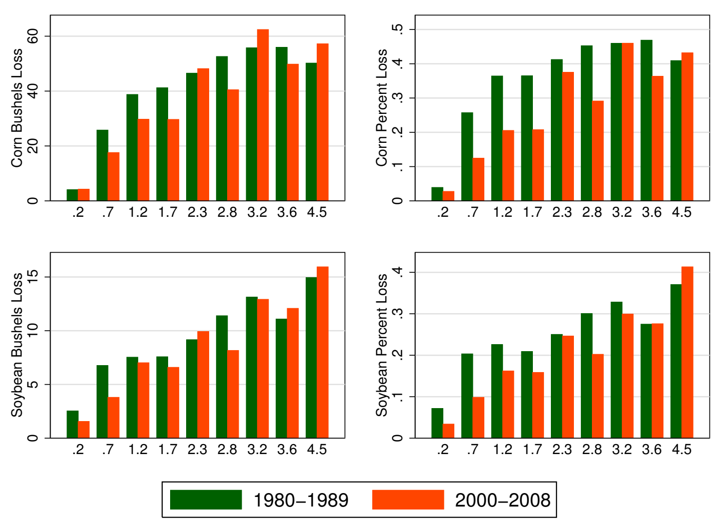

Figure 2.3: Crop losses and drought indices

The bar charts in figure 2.3 also show that the relationship between crop losses and the drought index is probably non-linear. The marginal impact of drought on yield seems to decline with drought severity. This pattern suggests inclusion of a quadratic term of the drought index in the regression equation of crop yield that follows.

2.3.1 Modeling Crop Yield with the Drought Index

The nonparametric results presented in figure 2.3 indicate that drought tolerance has increased for corn and perhaps for soybeans. We now turn to regression analysis to test this hypothesis and to obtain estimates of the magnitude of any effect. The model we use is a fixed-effects regression model using the panel of data consisting of county yields and matching county-level drought indices. The first regression equation we specify is as follows:

$$Y_{i,t} = \beta_{cons}+\alpha_i+\sum_{r=1}^{R}\gamma_r(CRD_r\times T)+\beta_{di}DI_{i,t}+\beta_{dit}DIT_{i,t}+\beta_{disq}DISQ_{i,t}+\beta_{disqt}DISQT_{i,t}+\epsilon_{i,t}. \tag{2.2}$$

Subscripts t, i, and r denote time, county, and crop reporting districts, respectively. Y denotes crop yield. T is a trend variable, which takes values 0 to 28 for years 1980 to 2008. DI is the drought index. DIT, $DISQ$, and $DISQT$ are variables defined as follows: $DIT = DI \times T$, $DISQ = DI \times DI$, $DISQT = DI \times DI \times T$. CRD_r, $r = 1, 2, ..., R$, denote regional dummy variables. R is the number of crop reporting districts. $CRD_r = 1$, if the yield observation is from crop reporting district r, and $CRD_r = 0$ otherwise. α_i is the county-level fixed-effect parameter. Without constraints, the fixed-effect parameter α_i and the coefficient of a time-invariant variable (the constant term) β_{cons} are not identified. Without loss of generality, we assume that $\sum_i \alpha_i = 0$. In this case, β_{cons} is the average intercept term, with α_i being each county's departure from the average. As usual, the error term $\epsilon_{i,t}$ is assumed to be mean zero over i and t. α, β, and γ are parameters to be estimated.

The yield regression model specified in (2.2) assumes a linear trend and explicitly accounts for the impact of drought. Crop yield consists of three parts: a deterministic trend yield ($\beta_{cons} + \alpha_i + \sum_{r=1}^{R} \gamma_r(CRD_r \times T)$), a drought-driven deviation ($\beta_{di}DI_{i,t} +$

$\beta_{dit}DIT_{i,t} + \beta_{disq}DISQ_{i,t} + \beta_{disqt}DISQT_{i,t})$, and the residual $\epsilon_{i,t}$. The linear trend has a county-specific intercept term $(\beta_{cons} + \alpha_i)$ and a slope $(\sum_{r=1}^{R} \gamma_r CRD_r)$, which varies across crop reporting districts. Counties located in crop reporting district z have a common trend slope γ_z, which differs from counties in another crop reporting district, say w, their trend slope being γ_w. The drought-driven deviation component depends on the drought index, a quadratic term of the drought index and time. Including the quadratic terms $DISQ$ and $DISQT$ makes the model more flexible in that marginal effects of drought could vary for different drought severities. The interaction terms DIT and $DISQT$ capture possible changes in drought impact over time. Since drought incidents are random, the drought-driven deviation allows us to explicitly estimate a part of yield risk that is caused by adverse weather. Other random factors are modeled by the residual term $\epsilon_{i,t}$. This specification allows for a straightforward test of whether yield loss measured in bushels per acre has varied over time.

The second regression model includes an exponential trend:

$$\ln(Y_{i,t}) = b_{cons}+a_i+\sum_{r=1}^{R} c_r(CRD_r\times T)+b_{di}DI_{i,t}+b_{dit}DIT_{i,t}+b_{disq}DISQ_{i,t}+b_{disqt}DISQT_{i,t}+e_{i,t}. \tag{2.3}$$

The regression equation (2.3) appears in a log-linear form with the same right-hand side as in the linear model (2.2). The log-linear model has a percentage change in crop yield on the left-hand side. Here, a's, b's, and c's are parameters to be estimated. This specification allows for a straightforward test of whether percentage yield loss due to drought has changed over time.

In models (2.2) and (2.3), parameters to characterize crop drought tolerance, β_{di}, β_{dit}, β_{disq}, β_{disqt}, b_{di}, b_{dit}, b_{disq}, b_{disqt}, are not region specific. We can allow these parameters to vary across CRDs by introducing products of regional dummies and the drought measures just as we did with the trend slope in equations (2.2) and (2.3). Specifically, CRD-specific models are

$$\begin{aligned} Y_{i,t} &= \beta_{cons} + \alpha_i + \sum_{r=1}^{R} \gamma_r (CRD_r \times T) \\ &+ \sum_{r=1}^{R} \beta_{di}^{r} (CRD_r \times DI_{i,t}) + \sum_{r=1}^{R} \beta_{dit}^{r} (CRD_r \times DIT_{i,t}) \\ &+ \sum_{r=1}^{R} \beta_{disq}^{r} (CRD_r \times DISQ_{i,t}) + \sum_{r=1}^{R} \beta_{disqt}^{r} (CRD_r \times DISQT_{i,t}) + \epsilon_{i,t} \quad (2.4) \end{aligned}$$

$$\begin{aligned} \ln(Y_{i,t}) &= b_{cons} + a_i + \sum_{r=1}^{R} c_r (CRD_r \times T) \\ &+ \sum_{r=1}^{R} b_{di}^{r} (CRD_r \times DI_{i,t}) + \sum_{r=1}^{R} b_{dit}^{r} (CRD_r \times DIT_{i,t}) \\ &+ \sum_{r=1}^{R} b_{disq}^{r} (CRD_r \times DISQ_{i,t}) + \sum_{r=1}^{R} b_{disqt}^{r} (CRD_r \times DISQT_{i,t}) + e_{i,t} \quad (2.5) \end{aligned}$$

In this case, all coefficients are CRD specific. The estimated coefficients in (2.4) and (2.5) are the same with what we would get by repeatedly regressing (2.2) and (2.3) using subsamples of each CRD. The benefit of estimating (2.4) and (2.5) instead of CRD-by-CRD regressions is that we can perform F-tests to see if there are regional differences in how drought has affected yield over time.

The fixed-effect model is chosen over pooled OLS and random-effect models for two reasons. First, omitted time-invariant, county-specific factors that influence crop yield can be captured by the fixed-effect parameters α_i. Second, α_i parameters differ among units, but for any particular unit, their value is constant in the fixed-effect model. Estimates of α's are of interest to us, particularly in estimating trend yields for each county.

The residual term, a proxy of all other yield risks, could be heteroskedastic among counties and over time. If the residual is heteroskedastic, estimated β's are consistent but their standard errors are underestimated. To test heteroskedasticity in a fixed-effect model setting, a modified Wald test is performed after regression (Greene, 2000). The null hypothesis that variance of residuals in all counties are equal is rejected at the 1% significance level. As suggested by Arellano (1987), a Huber/White/sandwich type variance matrix estimator of $\hat{\beta}$ is used in the context of a fixed-effect model (which Arellano called the within-group model) to obtain robust standard errors. The sandwich estimator is valid in the presence of any heteroskedasticity or serial correlation in the error term, provided T is small relative to N (Wooldridge, 2002).[10]

2.3.2 Impact of Drought

The aggregate regression models (2.2) and (2.3) as well as the CRD-specific regression models (2.4) and (2.5) explicitly account for the impact of drought on crop yield. With some derivation, marginal effects of drought as well as change in the crop losses from drought over time can be expressed as functions of model parameters. These expressions are derived for the aggregate models. Expressions for CRD-specific models follow by replacing the drought parameters with CRD-specific parameters.

For the linear regression model, the impact of drought on yield per acre is simply

$$Drought\ Impact \equiv \beta_{di} DI_{i,t} + \beta_{dit} DIT_{i,t} + \beta_{disq} DISQ_{i,t} + \beta_{disqt} DISQT_{i,t}. \tag{2.6}$$

If we constructed our drought index correctly, then the impact of drought on crop yields as defined in (2.6) should be negative. The change in drought impact over time is

$$\frac{\partial Drought\ Impact}{\partial T} = \beta_{dit} DI_{i,t} + \beta_{disqt} DI_{i,t} \times DI_{i,t}. \tag{2.7}$$

For a given level of drought severity, the partial derivative of drought impact with respect to time in (2.7) will be negative if yield loss from a drought of a given severity increases

[10] In our case, T equals 29, and N equals 98.

over time. If yields are becoming less susceptible to drought when the yield loss is measured in bushels, then (2.7) will be positive.

We could also use the linear specification (2.2) to test whether the percentage yield loss is increasing or decreasing over time, but the test statistics for testing this hypothesis are much more straightforward using the log-linear specification (2.3). The percentage change in yield with respect to the drought index is simply

$$ME^{relative} = \frac{\partial \ln(Y)}{\partial DI} = (b_{di} + 2b_{disq} \times DI_{i,t}) + (b_{dit} + 2b_{disqt} \times DI_{i,t}) \times T \tag{2.8}$$

and the change in drought susceptibility over time is simply $(b_{dit}DI_{i,t} + b_{disqt}DI_{i,t} \times DI_{i,t})$.

2.3.3 Estimation Results

We estimate yield regression models (2.2), (2.3), (2.4), and (2.5) using panel data of corn and soybean county yields, together with county-level drought indices. Table 2.2 provides point estimates of the CRD-specific trend coefficients γ_r. It is interesting to note that estimated yield growth rates for corn are uniformly higher than soybeans in Illinois but in Indiana, yield growth rates for the two crops are quite close.

Table 2.3 provides point estimates and robust standard errors (in parenthesis) of the CRD-invariant coefficients in the aggregate models (2.2) and (2.3). Columns 2 and 3 in table 2.3 provide estimated drought parameters for the linear model. The β_{di}'s are negative and significant for both corn and soybeans, which means that the drought index indeed captures the adverse effect of drought on yield. The β_{disq}'s are positive and significant for both crops, which means that marginal yield losses decline with drought severity. β_{dit} is positive and significant for corn, which implies that corn is less susceptible to minor droughts in terms of bushels lost over time. β_{disqt} is negative for corn, implying that over time, losses in corn bushels under severe droughts are not reduced as much as under minor droughts. β_{dit} and β_{disqt} for soybeans are both insignificant. This means that in bushel terms, there is no evidence that drought susceptibility of soybeans has changed in the past 28 years.

Table 2.2: CRD-Specific Trend Estimates γ_r's from Aggregate Models

State	District	Corn Linear	Soybean Linear	Corn Log-Linear	Soybean Log-Linear
IL	10	2.295 *	0.350 *	0.016 *	0.007 *
IL	20	1.872 *	0.288 *	0.013 *	0.007 *
IL	40	2.246 *	0.422 *	0.015 *	0.009 *
IL	70	1.827 *	0.504 *	0.017 *	0.014 *
IL	80	1.665 *	0.369 *	0.016 *	0.011 *
IL	90	1.431 *	0.388 *	0.014 *	0.012 *
IN	10	1.832 *	0.471 *	0.014 *	0.011 *
IN	50	1.494 *	0.438 *	0.010 *	0.010 *
IN	60	1.549 *	0.505 *	0.012 *	0.012 *
IN	70	1.749 *	0.517 *	0.013 *	0.014 *
IN	80	1.239 *	0.517 *	0.011 *	0.014 *

Note: Asterisk (*) denotes estimates significant at 5%.

Columns 4 and 5 in table 2.3 provide estimated drought parameters for the log-linear model. The b_{di}'s are negative and significant for both crops as expected. The b_{disq}'s are positive and significant for both crops, which means marginal percentage losses are higher under minor droughts and lower under severe droughts. The b_{dit}'s are significant and positive for both crops. Both corn and soybeans are becoming less susceptible to minor droughts in percentage terms. b_{disqt} is significant and negative for corn, implying that percentage corn yield losses are not decreasing as much under severe droughts as under minor droughts over time. b_{disqt} is insignificant for soybeans implying that improvement of soybean drought tolerance measured in percentage terms is similar under minor droughts compared to severe droughts.

Table 2.3: Point Estimates and Robust Standard Errors of CRD-Invariant Coefficients

	Corn Linear	Soybean Linear	Corn Log-Linear	Soybean Log-Linear
DI	-40.6124 *	-5.7361 *	-0.5023 *	-0.2157 *
	(2.5117)	(0.6667)	(0.0369)	(0.0251)
DIT	0.6886 *	0.0069	0.0143 *	0.0035 *
	(0.1480)	(0.0417)	(0.0018)	(0.0013)
$DISQ$	7.2940 *	0.7048 *	0.0914 *	0.0322 *
	(0.9451)	(0.2531)	(0.0133)	(0.0095)
$DISQT$	-0.2050 *	-0.0033	-0.0037 *	-0.0009
	(0.0539)	(0.0154)	(0.0007)	(0.0005)
Constant	104.6338 *	34.4129 *	4.6500 *	3.5316 *
	(0.4362)	(0.1301)	(0.0042)	(0.0040)

Note: Asterisk (*) denotes estimates significant at 5%.

The regression results have important implications because drought risk is an important source of yield losses. Our regression models explicitly estimate crops' ability to withstand drought. If the distribution of drought remains constant over time, then changes in drought tolerance translate directly into changes in yield risk. An increasing (decreasing) drought tolerance results in decreasing (increasing) drought-induced yield risk. Based on regression results, one can test the null hypothesis that absolute (relative) yield risk induced by drought is constant over time.

Based on the linear model, the null hypothesis that absolute yield risk induced by drought is constant over time is equivalent to the annual change in bushels lost to drought being zero:

$$H_0: \ \beta_{dit} DI + \beta_{disqt} DI \times DI = 0. \tag{2.9}$$

Based on the log-linear model, the null hypothesis of constant relative yield risk induced by drought is equivalent to the annual change in percentage yield losses being zero:

$$H_0: \ b_{dit} DI + b_{disqt} DI \times DI = 0. \tag{2.10}$$

Table 2.4 provides t-statistics and p-values for these hypothesis tests. For corn, the null hypothesis that yield losses have not changed over time is rejected in favor of the alternative hypothesis that yield losses have declined over time. This conclusion holds whether yield loss is measured in bushels per acre or as a percentage of mean yields. For soybeans, the null hypothesis that yield losses measured in bushels per acre are constant over time cannot be rejected. However, the null hypothesis that percentage yield losses have not changed over time is rejected in favor of the alternative hypothesis that percentage yield losses have decreased over time.

Table 2.4: Hypothesis Test Results

Crop	H_0	Test statistics	P-value	Implications
		Drought Index Evaluated at 1		
Corn	$\beta_{dit}DI + \beta_{disqt}DI \times DI = 0$	t(95) = 4.88	.000	CAR rejected, in favor of DAR
Corn	$b_{dit}DI + b_{disqt}DI \times DI = 0$	t(95) = 8.90	.000	CRR rejected, in favor of DRR
Soybean	$\beta_{dit}DI + \beta_{disqt}DI \times DI = 0$	t(95) = 0.13	.897	Fail to reject CAR
Soybean	$b_{dit}DI + b_{disqt}DI \times DI = 0$	t(95) = 3.03	.003	CRR rejected, in favor of DRR
		Drought Index Evaluated at 2.5		
Corn	$\beta_{dit}DI + \beta_{disqt}DI \times DI = 0$	t(95) = 3.44	.001	CAR rejected, in favor of DAR
Corn	$b_{dit}DI + b_{disqt}DI \times DI = 0$	t(95) = 9.07	.000	CRR rejected, in favor of DRR
Soybean	$\beta_{dit}DI + \beta_{disqt}DI \times DI = 0$	t(95) = -0.08	.933	Fail to reject CAR
Soybean	$b_{dit}DI + b_{disqt}DI \times DI = 0$	t(95) = 2.36	.021	CRR rejected, in favor of DRR

Note: The only yield risk considered here is the drought-induced risk.
CAR, DAR, CRR and DRR denote respectively constant absolute risk, decreasing absolute risk, constant relative risk and decreasing relative risk.

Based on model estimates, the marginal effects of drought evaluated at drought index level 1.0 (an average drought severity level) and their standard errors are listed in table 2.5. The marginal effect on corn yield was 26 bushels per acre or 32% in 1980 and 18 bushels per acre or 12% in 2008. Standard errors are about one bushel per acre or 1%.

The marginal impact on soybean yield was four bushels per acre in both 1980 and 2008. But because expected yields have increased over time, the marginal effect measured in percentage terms was 15% in 1980 and 10% in 2008, with about 1% standard error.

Table 2.5: Marginal Effects with Drought Index Evaluated at $DI = 1$

Year	1980	1985	1990	1995	2000	2005	2008
Corn Loss	26.02	24.63	23.24	21.85	20.45	19.06	18.23
(bushels)	(0.85)	(0.63)	(0.49)	(0.51)	(0.68)	(0.92)	(1.08)
Soybean Loss	4.33	4.32	4.32	4.32	4.32	4.32	4.32
(bushels)	(0.25)	(0.19)	(0.16)	(0.18)	(0.23)	(0.30)	(0.34)
Corn Loss	32.0%	28.5%	25.0%	21.5%	18.0%	14.5%	12.4%
(%) Log-Linear	(1.3)	(1.0)	(0.8)	(0.7)	(0.7)	(0.9)	(1.1)
Soybean Loss	15.1%	14.3%	13.5%	12.6%	11.8%	11.0%	10.5%
(%) Log-Linear	(0.9)	(0.7)	(0.6)	(0.6)	(0.8)	(0.9)	(1.1)

Note: Standard errors of the predictions are in parenthesis.

Table 2.6 provides point estimates and standard errors of the predicted change in annual yield losses due to droughts of different severities. Numbers in the last two columns need to be interpreted with caution because of the large standard errors, caused by too few observations in our sample that have a drought index greater than 3. Within the range of the drought index that we have adequate data, model predictions are as follows. Corn yield losses due to drought have decreased on an annual basis by between 0.22 bushels to 0.57 bushels, or by between 0.63% to 1.4% each year, depending on drought severity. Estimated percentage soybean losses have decreased at a rate ranging from 0.15% per year to 0.33% per year. Although these annual changes may seem moderate, the magnitude is large when we consider the accumulated change over time. Figure 2.4 shows crop yield loss in 1980 and in 2008 under alternative drought severities. In 29 years, corn yield lost to drought is estimated to have decreased by about 15 bushels under moderate droughts. Accounting for the increase in trend yield, the reduction in the percent loss in corn yields over the 29 years ranges from nearly 20 percentage points when the drought index is 0.5 to nearly 40 percentage points when the drought index is 2. Soybeans lost as many bushels to drought in 2008 as in 1980. In relative terms, however, the reduction in soybean losses was about 5 to 10 percentage points from 1980 to 2008. All theses predicted values are statistically different from zero.

Table 2.6: Annual Change in Crop Losses at Alternative Drought Severity Levels

Drought Index	0.5	1	1.5	2	2.5	3	3.5	4
Annual Decrease in	0.29	0.48	0.57	0.56	0.44	0.22	-0.10	-0.53
Corn Bushel Loss	(0.06)	(0.10)	(0.12)	(0.12)	(0.13)	(0.17)	(0.25)	(0.37)
Annual Decrease in	0.63%	1.07%	1.32%	1.40%	1.29%	0.99%	0.52%	-0.15%
Corn Percent Loss	(0.08)	(0.12)	(0.14)	(0.14)	(0.14)	(0.20)	(0.31)	(0.46)
Annual Decrease in	0.00	0.00	0.00	0.00	0.00	-0.01	-0.02	-0.03
Soybean Bushel Loss	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.08)	(0.11)
Annual Decrease in	0.15%	0.26%	0.32%	0.33%	0.30%	0.22%	0.10%	-0.06%
Soybean Percent Loss	(0.05)	(0.08)	(0.10)	(0.11)	(0.13)	(0.18)	(0.27)	(0.40)

Note: Standard errors of the predictions are in parenthesis.

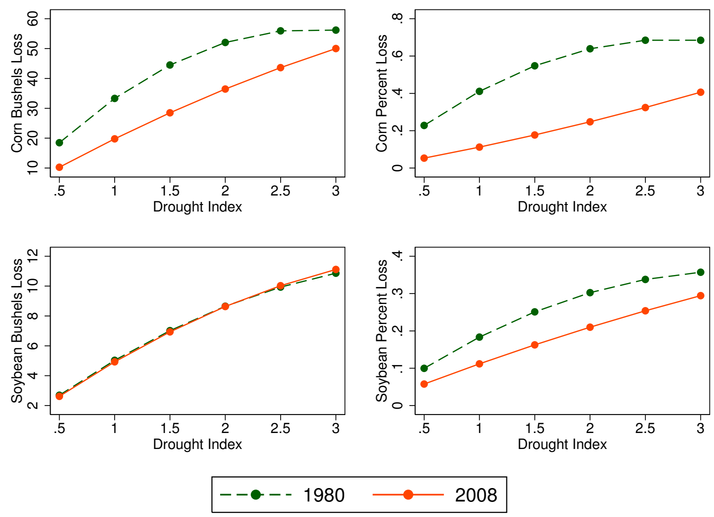

Figure 2.4: Model-predicted drought-induced crop losses

To estimate possible CRD-specific drought effects, we also fit models (2.4) and (2.5) with the same panel data. Tables 2.7 and 2.8 provide point estimates of coefficients in the CRD-specific models. Note that, in general, the signs of the CRD-specific drought parameters match their counterparts in the aggregate models in table 2.3. For those drought parameters that are estimated to be significant in the aggregate model, estimates in the CRD-specific models are either significant with the same sign or insignificant. For the drought parameters estimated to be insignificant in the aggregate model, their estimates could take both positive and negative signs or be estimated as insignificant in the CRD-specific model. There is only one exception: b^r_{dit} for soybeans in Indiana district 80 has a different sign than b_{dit} in the aggregate model.

Table 2.7: Point Estimates of Coefficients from CRD-Specific Models

State	District	Corn Linear	Soybean Linear	Corn Log-Linear	Soybean Log-Linear
			Intercept	β_{cons}	
		104.8900 *	34.4730 *	4.6546 *	3.5346 *
			Trend Parameter	γ_r's	
IL	10	2.3497 *	0.3393 *	0.0163 *	0.0073 *
IL	20	1.9057 *	0.2593 *	0.0137 *	0.0059 *
IL	40	2.2673 *	0.3512 *	0.0156 *	0.0077 *
IL	70	1.7078 *	0.4542 *	0.0137 *	0.0118 *
IL	80	1.7893 *	0.4427 *	0.0159 *	0.0129 *
IL	90	1.3869 *	0.4228 *	0.0127 *	0.0123 *
IN	10	1.8550 *	0.4462 *	0.0149 *	0.0108 *
IN	50	1.4532 *	0.4700 *	0.0106 *	0.0105 *
IN	60	1.6304 *	0.5142 *	0.0129 *	0.0126 *
IN	70	1.5527 *	0.5034 *	0.0120 *	0.0129 *
IN	80	1.3339 *	0.5687 *	0.0113 *	0.0154 *
			Drought Parameter	β^r_{di}'s or b^r_{di}'s	
IL	10	-24.5300 *	0.4016	-0.2606 *	0.0172
IL	20	-34.0760 *	-4.9898 *	-0.3283 *	-0.1301 *
IL	40	-40.7260 *	-12.9590 *	-0.3920 *	-0.3341 *
IL	70	-59.4110 *	-12.3450 *	-0.9352 *	-0.5076 *
IL	80	-41.5350 *	-5.0797 *	-0.6349 *	-0.2331 *
IL	90	-70.7350 *	-11.7250 *	-1.0575 *	-0.5999 *
IN	10	-31.9900 *	-7.4177 *	-0.3498 *	-0.2057 *
IN	50	-52.2970 *	-2.1596	-0.4980 *	-0.0742
IN	60	-32.3670 *	-4.9152 *	-0.3898 *	-0.1710 *
IN	70	-55.1150 *	-8.6409 *	-0.5406 *	-0.3005 *
IN	80	-38.4830 *	-6.1125 *	-0.5203 *	-0.2636 *
			Drought Parameter	β^r_{dit}'s or b^r_{dit}'s	
IL	10	0.0001	0.1004	0.0071	0.0029
IL	20	-0.1662	0.1039	0.0032	0.0035
IL	40	0.7263	0.8731 *	0.0143 *	0.0233 *
IL	70	1.0383 *	0.1975	0.0300 *	0.0135 *
IL	80	0.3805	-0.3754 *	0.0149	-0.0083
IL	90	2.0711 *	0.0090	0.0367 *	0.0109 *
IN	10	0.5336	0.4464 *	0.0087 *	0.0130 *
IN	50	1.4237 *	-0.2868 *	0.0153 *	-0.0050
IN	60	-0.0079	-0.1520 *	0.0068 *	-0.0012
IN	70	2.3829 *	0.0501	0.0233 *	0.0046 *
IN	80	0.0817	-0.3908 *	0.0088 *	-0.0054 *

Note: Asterisk (*) denotes estimates significant at 5%.

Table 2.8: Point Estimates of Coefficients from CRD-Specific Models - Continued

State	District	Corn Linear	Soybean Linear	Corn Log-Linear	Soybean Log-Linear
		Drought Parameter		β^r_{disq}'s or	b^r_{disq}'s
IL	10	2.1321	-1.1129	0.0146	-0.0343
IL	20	2.8345	-0.3560	-0.0024	-0.0234
IL	40	7.9223 *	3.5543 *	0.0701 *	0.0868 *
IL	70	12.1130 *	2.2724 *	0.2197 *	0.1045 *
IL	80	9.7572 *	0.9630	0.1490 *	0.0466
IL	90	29.7610 *	5.1016 *	0.4591 *	0.2774 *
IN	10	4.7221 *	1.8526 *	0.0449 *	0.0441 *
IN	50	12.0790 *	-0.3684	0.1042 *	-0.0071
IN	60	5.0586 *	0.2324	0.0639 *	0.0099
IN	70	12.9670 *	1.7575 *	0.1057 *	0.0567 *
IN	80	7.5398 *	1.3834 *	0.0955 *	0.0598 *
		Drought Parameter		β^r_{disqt}'s or	b^r_{disqt}'s
IL	10	-0.0557	-0.0478	-0.0024	-0.0013
IL	20	0.2159	0.0239	0.0020	0.0010
IL	40	-0.4393	-0.4086 *	-0.0066 *	-0.0107 *
IL	70	-0.1007	-0.0051	-0.0067 *	-0.0022
IL	80	-0.1680	0.1393 *	-0.0052	0.0035
IL	90	-1.2272 *	-0.1383	-0.0197 *	-0.0098 *
IN	10	-0.1981	-0.2277 *	-0.0021	-0.0060 *
IN	50	-0.4712 *	0.1074	-0.0038 *	0.0023
IN	60	-0.0228	0.0452 *	-0.0017	0.0008 *
IN	70	-0.8042 *	-0.0103	-0.0066 *	-0.0007
IN	80	-0.0680	0.1157 *	-0.0024 *	0.0018 *

Note: Asterisk (*) denotes estimates significant at 5%.

Based on the CRD-specific estimates, we get similar results of annual changes in drought tolerance as from the aggregate models. Drought tolerance of corn measured in bushels per acre improved in four CRDs and showed no statistically significant changes in the other seven CRDs. Measured in relative terms, corn crops in all 11 CRDs are becoming less susceptible to drought over time. Measured in bushels per acre, soybean crops in four CRDs are becoming more drought tolerant, soybean crops in five CRDs are becoming less drought tolerant and soybean crops in the remaining two CRDs showed no significant changes. Measured in relative terms, soybean crops in five CRDs are more drought tolerant, the soybean crops in only one CRD is becoming more susceptible to drought, and the rest did not show significant changes. Thus, we can think of the aggregate estimation as the “average” of the CRD-specific estimation.

The magnitude of the CRD-specific estimates varies across regions. Table 2.9 shows that F-tests of the null hypothesis that CRD-specific drought parameters in the CRD-specific model are equal is rejected. This suggests that corn and soybean yields in different CRDs could respond differently to droughts. The differences could be due to the variability in pre-drought soil moisture and that the CRD-specific parameters were estimated using too few observations. It is likely that pre-drought soil moisture differs across CRDs. Thus droughts of similar severity levels could lead to different degrees of crop losses in different CRDs. However, we lack the pre-drought soil moisture data to account for the differences. Also, the CRD-specific regressions use only information within a CRD to estimate drought parameters for that CRD. Few observations in some CRDs used could lead to large standard errors in the estimated coefficients of the CRD-specific drought parameters. Rather than an actual finding of no change in drought tolerance, we could have insignificant estimates from a lack of data. Even the significant estimates could suffer from lack of precision. As shown in table 2.10, the aggregate models have smaller (better) Bayesian information criterion (BIC) values than do corresponding CRD-specific models. This suggests that the CRD-specific models are relatively overfitting. For these reasons, and because the aggregate model results essentially represent the average of the CRD-specific model results, we rely on results of the aggregate models.

Table 2.9: F-test of Equal Coefficients

Model	H0	F Statistics	p-value	Conclusion
Corn	β^r_{di} 's equal	F(10,95)=4.58	0	reject
Linear	β^r_{dit}'s equal	F(10,95)=7.51	0	reject
	β^r_{disq}'s equal	F(10,95)=4.57	0	reject
	β^r_{disqt}'s equal	F(10,95)=6.25	0	reject
Corn	b^r_{di} 's equal	F(10,95)=7.21	0	reject
Log-Linear	b^r_{dit}'s equal	F(10,95)=7.12	0	reject
	b^r_{disq}'s equal	F(10,95)=6.12	0	reject
	b^r_{disqt}'s equal	F(10,95)=6.42	0	reject
Soybean	β^r_{di} 's equal	F(10,95)=6.14	0	reject
Linear	β^r_{dit}'s equal	F(10,95)=16.62	0	reject
	β^r_{disq}'s equal	F(10,95)=5.82	0	reject
	β^r_{disqt}'s equal	F(10,95)=11.88	0	reject
Soybean	b^r_{di} 's equal	F(10,95)=8.22	0	reject
Log- Linear	b^r_{dit}'s equal	F(10,95)=10.45	0	reject
	b^r_{disq}'s equal	F(10,95)=5.52	0	reject
	b^r_{disqt}'s equal	F(10,95)=7.04	0	reject

Table 2.10: Model Selection Based on Bayesian Information Criterion (BIC)

Model	BIC
Corn Linear CRD-specific Model	21278.5
Corn Linear Aggregate Model	21065.4
Corn Log-Linear CRD-specific Model	-2524.8
Corn Log-Linear Aggregate Model	-2621.5
Soybean Linear CRD-specific Model	14950.9
Soybean Linear Aggregate Model	14844.0
Soybean Log-Linear CRD-specific Model	-3302.0
Soybean Log-Linear Aggregate Model	-3320.6

2.4 Implications for GRP Rates

The conclusion that the drought tolerance of corn and soybeans in Indiana and Illinois has changed over time has important implications for the U.S. crop insurance program. For example, the Group Risk Plan (GRP) is an area yield insurance program that pays indemnity whenever the actual county yield falls below the "trigger yield", which is a proportion of expected county yield. Up until 2009, GRP premium rates were determined using a loss-cost ratio (LCR) methodology (Skees et al., 1997). Actuarial fairness of this approach is based on the constant relative risk hypothesis (Paulson and Babcock, 2008). Although RMA recently relaxed this assumption for GRP by using the actual yield history to determine if relative risk in a county has increased or decreased over time, their new procedure does not account for whether drought incidence has increased or decreased over time. Thus, they are not able to discern if a finding of decreasing relative yield risk is caused by fewer droughts or an actual reduction in risk. Individual yield insurance under the federal Actual Production History (APH) program is also based on a LCR method that requires constant relative risk over time (Woodard et al., 2008). In a situation of decreasing (increasing) relative risk, LCR methodology based on direct historical experiences overestimates (underestimates) premium rates. In light of our findings of decreasing drought-induced relative deviations for both corn and soybeans at the county level, it is worthwhile to assess the impacts of modifying the underlying rating assumption of constant relative risk on GRP rating. To do this, we simulate actuarial fair rates for insuring drought-induced risk and compare them with the drought-related part of 2008 GRP rates. We use 2008 GRP rates because the new RMA rate-making procedures were not made available to the authors.

In the simulations two assumptions are made. First, the future probability of drought in any year is assumed to be captured by the historical probability of drought in our sample. This assumption can be relaxed if information about future distributions of drought becomes available. Second, to focus only on the drought portion of GRP rates, we assume that drought-induced yield risk can be rated separately from other sources of yield loss. Essentially, yield risk is assumed to be in an additive form of multiple independent risk factors with drought-induced risk being one major component. The actuarial fair rate simulated below is the rate that corresponds only to the drought-induced risk factor.

We simulate actuarially fair GRP rates for drought from 2010 to 2020 using the following steps. We first select a representative county. For each year from 2010 to 2020, we take the county's 29 historical values of the drought index as the draws from the empirical distribution of drought. With these 29 drought index draws for each county, actual yields ($ActYield$) from 2010 to 2020 are predicted using fitted values from the regression models (2.2) and (2.3) with the estimated coefficients in table 2.3.[11] Expected yields ($E(Y)$) are predicted using the fitted values from regression models with the drought-index evaluated at its county mean. We fix the coverage level (C) at 0.9. Loss percentages ($\%loss$) are then calculated using (2.11) and (2.12).

$$TrigYield_{i,t} = C \times E(Y_{i,t}) \tag{2.11}$$

$$\%loss_{i,t} = \max[\frac{TrigYield_{i,t} - ActYield_{i,t}}{TrigYield_{i,t}}, 0]. \tag{2.12}$$

The actuarial fair premium rate ($Fair^{rate}$), is the rate such that expected indemnity ($Indem$) equals expected premium ($Prem$). It is derived in (2.13) to (2.15) to be the expected percentage loss.

$$Indem_{i,t} = \%loss_{i,t} \times Liab_{i,t} \tag{2.13}$$

$$Prem_{i,t} = rate_{i,t} \times Liab_{i,t} \tag{2.14}$$

$$E(Indem_{i,t}) = E(Prem_{i,t}) \Leftrightarrow rate_{i,t} = E(\%loss_{i,t}) \Rightarrow Fair^{rate}_{i,t} = E(\%loss_{i,t}). \tag{2.15}$$

We take the average of simulated loss percentages to be the simulated fair rate. Two qualifications on the simulated fair rates need to be addressed. First, the fair rates are prices for insuring drought-induced yield losses only. Other sources of yield risk are not modeled in simulation. Second, only the regression-predicted drought-induced yield risk, or the "mean effect" of the drought shock, is modeled. Actual drought-induced yield fluctuations could be more volatile. Thus, the real actuarial fair rate should be higher. Nevertheless, simulated fair rates indicate prices of insuring model-implied drought risk.

[11] Only the drought index is drawn from its empirical distribution. The residual term of the yield regression is not included in the predicted "actual yield". This procedure essentially excludes all yield risks other than the drought-induced yield risk predicted by the regression.

GRP is a multi-peril crop insurance product. To make our simulated fair rates comparable to GRP rates, we need to calculate GRP rates including only those losses that occurred in drought years as indicated by this study's drought index. The modified GRP rating method is as follows. First, linear trends are fitted county by county to crop yields. Second, loss percentages are calculated for historical years according to (2.11) and (2.12), with fitted values of crop yield as expected yields and historical yield observations as actual yields. Finally, the GRP rate is set at the average of historical loss percentages:

$$GRP_{i,t}^{rate} = \frac{\sum_{\tau=0}^{t} \%loss_{i,\tau}}{(t-0)}. \tag{2.16}$$

To calculate GRP drought year rates, only the final step differs:

$$GRP_{i,t}^{drought\ year\ rate} = \frac{\sum_{\tau \in \Re} \%loss_{i,\tau}}{r_{i,t}} \tag{2.17}$$

where $\Re$ is the set of drought years and r is the number of drought years in history. Drought years are those identified by the drought index as being positive. The GRP drought year rate is the simple average of loss ratios in drought years only. It reflects loss experiences closely related to drought conditions that we used to simulate the fair rates.[12]

In table 2.11, we list premium rates for selected counties in our sample. Column 3 lists unloaded and unsubsidized GRP premium rates in 2008 and 2010.[13] Column 4 lists the GRP rates that we calculate for drought years using (2.17). Column 5 is the ratio of the drought-year GRP rate in column 4 divided by the GRP rate we calculate using (2.16), which indicates the percentage of GRP rate that can be accounted for by drought. The last three columns are simulated fair rates of insuring the model-implied

[12]It also reflect other risk factors in drought years that the fair rates do not incorporate.

[13]We take the unsubsidized GRP rate listed on the RMA website and multiply by 0.88 to remove load. The 2010 premium rates are included to show how the new rating procedures used by RMA affected rates.

drought risk in 2010, 2015, and 2020. There are two rows for each county. The first rows are fair rates predicted by the linear model. The second rows report rates predicted by the log-linear model. GRP drought year rates in column 4 are comparable to simulated fair rates in the last three columns in terms of insured risk.

Table 2.11: Actual and Simulated Premium Rates

State	County	GRP Rates		GRP Drought	Drought	Simulated GRP Rates		
		2008	2010	Rates	Percent	2010	2015	2020
				Corn				
IL	Bureau	2.24%	2.92%	2.39%	92%	0.95%	0.77%	0.65%
						1.41%	1.14%	1.15%
IL	Peoria	3.41%	4.43%	3.66%	89%	0.91%	0.66%	0.44%
						1.08%	0.61%	0.16%
IN	Clinton	3.21%	2.54%	2.99%	95%	1.43%	1.25%	1.08%
						1.77%	1.44%	1.10%
IN	Spencer	3.61%	4.58%	2.84%	76%	1.62%	1.27%	0.98%
						1.31%	0.91%	0.65%
				Soybean				
IL	La Salle	2.23%	2.26%	1.16%	65%	0.58%	0.53%	0.49%
						0.68%	0.58%	0.49%
IN	Jasper	2.22%	2.26%	1.32%	55%	0.70%	0.60%	0.51%
						0.90%	0.77%	0.64%
IN	Randolph	2.58%	2.28%	1.88%	76%	0.82%	0.73%	0.64%
						0.96%	0.86%	0.76%
IN	Spencer	2.64%	2.40%	2.69%	81%	0.95%	0.81%	0.69%
						0.81%	0.70%	0.59%

Note: First rows of simulated GRP drought rates are estimated from the linear model.

Second rows of simulated GRP drought rates are estimated from the log-linear model.

For the Indiana and Illinois selected counties, almost all of the corn GRP rates are the result of drought and more than half of the soybean GRP rates are the result of drought. For both crops, the 2010 fair rates are lower than the GRP drought year rates, and they decline over time. The reason why they are lower in 2010 is that the GRP drought rate assumes constant relative susceptibility to drought from 1980 to 2008 whereas the reality is that both corn and soybeans have actually exhibited decreasing relative susceptibility. Simulated rates for corn decline faster than rates for soybeans because corn shows a larger improvement in drought tolerance. These results suggest that drought is the most important source of county yield losses and that this study's findings of increasing drought tolerance suggest that the drought portion of GRP rates are much too high. This suggests that the findings of (Woodard et al., 2011) that crop insurance premium rates in the Corn Belt are too high can be explained by increasing drought tolerance by corn and soybean crops.

2.5 Conclusions

By constructing an objective drought index and correlating to crop yields, we explicitly account for the impact of drought on crop yield. Regression results show that corn is becoming less susceptible to drought measured both by bushel loss and percentage loss. For soybeans, constant bushel loss is not rejected but the degree of drought tolerance measured in percentage term is decreasing over time. The decreasing relative susceptibility for both crops cast doubt on the LCR method used in rating crop insurance programs in the United States. Simulations based on regression estimation results show that accounting for increased drought tolerance of corn and soybeans would have major impacts on premium rates for GRP.

That both corn and soybeans in Indiana and Illinois are more drought tolerant is somewhat surprising because only now are the large seed companies focusing their efforts on developing drought-tolerant crops, and all of their work to date has been devoted to corn. Their past efforts at developing biotech corn seem to have paid off in an unanticipated manner by making corn hybrids better able to withstand drought conditions. But no such effort can explain increasing drought resistance in soybeans, unless herbicide-resistant soybeans are less susceptible to drought. Besides widespread adoption of biotech corn and soybeans that began in the 1990s, the other large change common to both corn and soybeans is that a greater proportion of both crops is being managed by larger and perhaps more able managers. Better management leads to more timely field operations, which could result in increasing drought tolerance. If true, then our finding of increased drought tolerance may apply to other crops for which management may have improved.

The improvements in drought tolerance that we have documented in corn and soybeans may be dwarfed in the future if in fact the seed companies are successful in their efforts to introduce genes that enable crops to withstand drought conditions. The crop insurance industry and the Risk Management Agency of USDA in particular should begin to alter the way they determine crop insurance rates so that as the new technologies come online, they will have a system that can reflect the new lower risks directly in premium rates. More generally, greater drought tolerance will reduce price volatility and the risk of major disruptions in food and fuel supplies. We conclude with a note of caution. Because we have relatively few observations of major droughts in the 2000s, our ability to measure the extent to which yield losses have declined under the most severe drought conditions is limited. Because even drought tolerant crops need adequate moisture, it would not be surprising if our estimates of the reduction in yield losses from moderate to severe drought conditions will turn out to be higher than the change in yield losses that would actually occur under the most severe drought conditions.

CHAPTER 3. ESTIMATING NON-LINEAR WEATHER IMPACTS ON CORN YIELD – A BAYESIAN APPROACH

Abstract

We estimate impacts of rainfall and temperature on corn yields by fitting a linear spline model with endogenous thresholds. Using Gibbs sampling and the Metropolis - Hastings algorithm, we simultaneously estimate the thresholds and other model parameters. A hierarchical structure is applied to capture county-specific factors determining corn yields. Results indicate that impacts of both rainfall and temperature are nonlinear and asymmetric in most states. Yield is concave in both weather variables. Corn yield decreases significantly when temperature increases beyond a certain threshold, and when the amount of rainfall decreases below a certain threshold. Flooding is another source of yield loss in some states. A moderate amount of heat is beneficial to corn yield in northern states, but not in other states. Both the levels of the thresholds and the magnitudes of the weather effects are estimated to be different across states in the Corn Belt.

3.1 Introduction

In rain-fed agricultural regions, weather conditions have substantial impacts on crop productivity. Favorable weather conditions for dryland crop production, including a proper amount of heat and rainfall during the growing season, are critical factors determining yield outcomes (O'Brien, 1993). Most previous studies found that corn yield decreases in temperature and increases in the amount of rainfall (Lobell and Asner, 2003; Deschenes and Greenstone, 2007). Understanding how weather variables affect crop yield is essential in measuring yield risk and rating crop insurance plans (Yu and Babcock, 2010).

In quantifying the impact of weather on corn yield, one widely used approach is to estimate a reduced form statistical relationship between corn yield and weather variables. Two commonly used reduced forms are linear and quadratic functions. The linear specification seems convenient but it could be restrictive as well. There is evidence that the response of corn yield to temperature and rainfall may not be constant over the whole range of possible weather outcomes. In fact, an increase of temperature within the range between 8° Celsius and 32° Celsius is found to be beneficial to corn yield, but increasing the temperature further beyond 34° Celsius leads to yield losses (Schlenker et al., 2006a). If weather factors are beneficial to corn yield in some ranges but not in others, then the linear specification would generate a misleading result. One way to model the non-linearity is to include a quadratic term. However, the quadratic functional form restricts the yield response to be symmetric. There is evidence that bad growing conditions typically cause more yield losses than good growing conditions cause yield gains. In fact, based on county-level panel data of corn yield, temperature, and precipitation from 1950 to 2004 in 2000 counties in the U.S., Schlenker and Roberts (2006) found that an increase in temperature above 25° Celsius decreases corn yield growth rates at an increasing rate. Nonparametric estimation is a viable approach to estimate the asymmetric nonlinear impacts of weather variables (Schlenker and Roberts, 2006). One limitation of the nonparametric approach, however, is that in some ranges of the weather variable, especially at the two tails of the distribution, the number of observations might be too small so that the standard errors tend to be very large.

In this article, we propose to estimate a two-knot linear spline model with endogenous thresholds. We simultaneously estimate the threshold parameters and other model parameters using a sampling-based Bayesian approach. Our specification captures the nonlinear and asymmetric feature of weather effects. The endogenous-knot linear-spline specification is relatively simple and yet flexible. Estimation in the Bayesian framework brings the advantages of high computational efficiency and quick convergence. In light of the findings that a modest increase in temperature benefits corn yield in the northern regions of the U.S. but not in other areas (Adams et al., 1990), we also examine possible geographical differences in how the weather factors influence corn yield.

The rest of this chapter is organized as follows: in section two we specify the two-knot linear spline model. In section three, we describe the Bayesian approach, which is applied to estimate the model. In section four, we present the estimation results. In the last section, we present a brief conclusion.

3.2 The Yield Model

We specify corn yield to be composed of a linear trend plus a function of weather variables, specifically:

$$\begin{aligned} Y_{i,t} \quad = \quad & \alpha_i + \sum_{r=1}^{R} \beta_{1,r}(CRD_r \times Time) \\ & + \beta_2 \min(0, (Temp_{i,t} - \theta_l)) + \beta_3 Temp_{i,t} + \beta_4 \max(0, (Temp_{i,t} - \theta_u)) \\ & + \beta_5 \min(0, (Rain_{i,t} - \lambda_l)) + \beta_6 Rain_{i,t} + \beta_7 \max(0, (Rain_{i,t} - \lambda_u)) + \epsilon_{i,t} \end{aligned} \quad (3.1)$$

Subscripts i, r, and t denote county, Crop Reporting District (CRD), and time, respectively. We denote the total number of counties as N, the total number of years as T, and the total number of CRDs as R. Y denotes corn yield. $Time$ is a time trend variable, which takes values 0 to 28 for years 1980 to 2008. CRD_r, $r = 1, 2, ..., R$, denotes the regional dummy variable. $CRD_r = 1$, if the yield observation is from crop reporting district r, and $CRD_r = 0$ otherwise. $Temp$ and $Rain$ denote mean monthly temperature and mean monthly rainfall in the growing season, respectively.[1] Without loss of generality, we recenter the temperature and rainfall variables at zero by subtracting the historical mean from each temperature or rainfall observation. The recentering process does not affect our estimation of other parameters except α_i. After recentering, α_i measures the average corn yield in the base year (1980) in county i when temperature and rainfall are at their historical mean levels. We assume that the error term $\epsilon_{i,t}$ is i.i.d. normal with mean zero and variance σ_ϵ^2.

In our specification, corn yield is expressed in an additive form of a linear trend and two-knot linear spline functions of rainfall and temperature. For the linear trend, we permit the intercept term α_i to vary across counties. Thus, α_i captures time-invariant county-specific factors that influence corn yield. Since fixed-effect estimation in a Bayesian setting is inefficient, we specify a hierarchical structure for α_i. Specifically, we assume that α_i, for $i = 1, 2, ..., N$, is independently and identically distributed from a normal distribution $N(\alpha, \sigma_\alpha^2)$. α and σ_α^2 are hierarchical parameters, which are estimated simultaneously with other model parameters. We allow the trend slope to vary across CRDs by including an interaction term between the regional dummy variable and the time trend variable. $\beta_{1,r}$ is fixed for any given CRD but differs across CRDs. Coefficients β_2 to β_7 are restricted to be constant across all counties. Equation (3.1) is essentially a

[1] We specify the growing season as from June to August.

mixed model, with a random effect α_i and a fixed effect β. In a Bayesian framework, we simultaneously estimate model parameters α, β, θ, and λ.

The two-knot linear spline functions of rainfall and temperature capture the potentially asymmetric and nonlinear feature of weather effects. We allow weather effects to be different when temperature (or the amount of rainfall) is lower than usual, within the middle range, or higher than usual. This is achieved by introducing endogenous threshold parameters. λ_l (or θ_l) and λ_u (or θ_u) denote the lower and upper thresholds that divide the domain of the weather variable into three ranges. The min and max operators serve as switches that turn on the specific β that we want to estimate when the weather variable falls into the specific range. The yield model allows a changing yield response as the weather regime changes. To better illustrate this feature, we rewrite model (3.1) as:

$$Y_{i,t} = \tilde{\alpha}_i + \sum_{r=1}^{R} \beta_{1,r}(CRD_r \times Time) + \beta_{temp,s} Temp_{i,t} + \beta_{rain,s} Rain_{i,t} + \epsilon_{i,t}$$

$$\beta_{temp,s} \equiv \begin{cases} \beta_{temp,cl} = \beta_2 + \beta_3 & \text{if } Temp \leq \lambda_l \\ \beta_{temp,nt} = \beta_3 & \text{if } \lambda_l < Temp < \lambda_u \\ \beta_{temp,ht} = \beta_3 + \beta_4 & \text{if } Temp \geq \lambda_u \end{cases}$$

$$\beta_{rain,s} \equiv \begin{cases} \beta_{rain,dr} = \beta_5 + \beta_6 & \text{if } Rain \leq \lambda_l \\ \beta_{rain,nr} = \beta_6 & \text{if } \lambda_l < Rain < \lambda_u \\ \beta_{rain,fl} = \beta_6 + \beta_7 & \text{if } Rain \geq \lambda_u \end{cases} \tag{3.2}$$

Subscript s denotes the state of nature, which is defined by the weather variable falling into one of the threshold-divided ranges. $\beta_{temp,s}$ (or $\beta_{rain,s}$) measures the marginal effect of temperature (or rainfall) in each state of nature. The marginal effect can potentially change depending on the weather condition. For example, one inch of rainfall in drought years could result in a different amount of change in corn yield than in normal years.

3.3 The Bayesian Approach

We now turn to the estimation methodology. Note that (3.1) is non-linear in threshold parameters λ and θ. Non-linear least square estimation (NLS) and maximum likelihood estimation (MLE) could potentially be used to estimate (3.1). In this chapter, we take the sampling-based Bayesian approach. The advantages of Bayesian estimation include easy implementation, fast convergence, and computational efficiency. We apply the Markov chain Monte Carlo (MCMC) in estimating (3.1). The implementation of the MCMC is essentially sequentially taking draws from (sequentially updated) conditional posterior distributions of model parameters. The ease of applying the MCMC to our model results from the observation that conditional on threshold parameters, (3.1) becomes a well known linear regression model with a hierarchical structure. Under conjugate priors, the conditional posterior distributions of parameters except λ and θ are readily derived. And it is easy to apply the Gibbs sampling to simulate draws from the conditional posterior distributions (Lindley and Smith, 1972; Gelfand et al., 1990; Chib and Carlin, 1999). The conditional posterior distributions of λ and θ require some derivations and are not of any recognizable distributional forms. Thus, we apply the Metropolis - Hastings algorithm (Chib and Greenberg, 1995; Gelman et al., 1995) in the step of drawing λ and θ from their conditional posterior distributions. The estimation process would become clear once we have the prior distributions, the likelihood function, and the posterior distributions, which will be specified and derived in the following sections.

3.3.1 Priors

Following Chib and Carlin (1999), we assume normal priors for α and β, and inverse gamma priors for the variance parameters. These are conjugate priors in the sense that the posterior distributions belong to the same family as the prior probability distributions. We assume conjugate priors due to the computational ease associated with this specification. Results are not sensitive to the specific form of the prior distribution. For notation simplicity, we stack all the fixed-effect parameters into one vector and define as: $\beta \equiv [\beta_{1,1}, \beta_{1,2}, ...\beta_{1,R}, \beta_2, \beta_3, ..., \beta_7]'$. We assume the prior distribution of β to be $N(\mu_\beta, V_\beta)$, where N denotes the normal distribution. As specified above, α_i is of normal distribution $N(\alpha, \sigma_\alpha^2)$. We assume the prior distributions of the hierarchical mean and variance parameters α and σ_α^2 to be $N(\mu_\alpha, V_\alpha)$ and $IG(a_1, a_2)$, respectively. IG denotes the inverse gamma distribution. We assume the prior distribution of the variance of the error term σ_ϵ^2 to be $IG(e_1, e_2)$. We assume that μ_β and μ_α equal to the corresponding OLS estimates of a linear yield model, $Y_{i,t} = \alpha_i + \sum_{r=1}^{R} \beta_{1,r}(CRD_r \times Time) + \beta_{temp}Temp_{i,t} + \beta_{rain}Rain_{i,t} + \epsilon_{i,t}$, with $\mu = \beta_{temp}^{OLS} \times [1,1,1,0,0,0]' + \beta_{rain}^{OLS} \times [0,0,0,1,1,1]'$. We assume $V_\alpha = 1600$ and $V_\beta = 100$ so that the normal priors are reasonably diffused. For the inverse gamma distributions, we set $e_1 = a_1 = 3$ and $e_2 = a_2 = \frac{1}{2 \times 100}$ so that the standard deviation of the inverse gamma distribution is 10.[2] Finally, we assume the prior distribution of the lower and upper thresholds to be the joint uniform distribution over the domain of the weather variable, with the restriction that the upper threshold is larger than the lower threshold. Specifically, $p(\theta_l, \theta_u) = \frac{2}{(Temp_{max} - Temp_{min})^2} I(Temp_{min} < \theta_l < \theta_u < Temp_{max})$ and $p(\lambda_l, \lambda_u) = \frac{2}{(Rain_{max} - Rain_{min})^2} I(Rain_{min} < \lambda_l < \lambda_u < Rain_{max})$, where $I(.)$ denotes the index function. We set $Temp_{max}$ (or $Rain_{max}$) at the recentered 95% percentile of all temperature (or rainfall) observations and $Temp_{min}$ (or $Rain_{min}$) at the recentered 5% percentile of all temperature (or rainfall) observations.

[2]Results are insensitive to the specific values of the prior distribution parameters.

3.3.2 The Likelihood Function

The other piece of information we need in order to derive the posterior distributions is the likelihood function. For notation simplicity, we stack the explanatory variables as:

$$\begin{aligned} X(\theta,\lambda)_{i,t} \equiv & [(CRD_1 \times T), (CRD_2 \times T), ..., (CRD_R \times T), \\ & (\min(0, (Temp_{i,t} - \theta_l))), Temp_{i,t}, (\max(0, (Temp_{i,t} - \theta_u))), \\ & (\min(0, (Rain_{i,t} - \lambda_l))), Rain_{i,t}, (\max(0, (Rain_{i,t} - \lambda_u)))]. \end{aligned} \tag{3.3}$$

Then the likelihood function is:

$$\begin{aligned} L(.) &= \prod_{i=1}^{N\times T} p\{Y_{i,t}|\alpha_i, \beta, \sigma_\epsilon^2, X_{i,t}(\theta,\lambda)\} \\ &= (2\pi\sigma_\epsilon^2)^{-\frac{N\times T}{2}} \exp[-\frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{N}\sum_{i=1}^{T}(Y_{i,t} - \alpha_i - X_{i,t}(\theta,\lambda)\beta)^2]. \end{aligned} \tag{3.4}$$

Notation $X_{i,t}(\theta,\lambda)$ emphasizes the fact that $X_{i,t}$ depends on thresholds θ and λ. To be succinct, we drop θ and λ and simply note $X_{i,t}$ when referring to $X_{i,t}(\theta,\lambda)$ in the following text.

3.3.3 Conditional Posteriors

Based on the priors and the likelihood function specified above, we derive the conditional posterior distributions. As mentioned above, standard results for the linear regression model apply to our model once we condition on the threshold parameters. Following Chib and Carlin (1999), the posterior distributions are derived by applying conclusions by Lindley and Smith (1972). Specifically, $p(\alpha_i|\beta,\alpha,\sigma_\epsilon^2,\sigma_\alpha^2,\theta,\lambda,Y) \sim N(D_1 d_1, D_1)$, where $D_1 = (\frac{T}{\sigma_\epsilon^2} + \frac{1}{\sigma_\alpha^2})^{-1}$, $d_1 = \frac{\iota_T'(Y_i - X_i\beta)}{\sigma_\epsilon^2} + \frac{\alpha}{\sigma_\alpha^2}$.[3] Similarly, the posterior distribution of β is also normal. $p(\beta|\alpha_i,\alpha,\sigma_\epsilon^2,\sigma_\alpha^2,\theta,\lambda,Y) \sim N(D_2 d_2, D_2)$, where $D_2 = (\frac{X'X}{\sigma_\epsilon^2} + V_\beta^{-1})^{-1}$,

[3] ι is a column vector of ones. Its dimension is specified by its subscript.

$d_2 = \frac{X'(y-\bar{\alpha})}{\sigma_\epsilon^2} + V_\beta^{-1}\mu_\beta$, $\bar{\alpha} \equiv [(\iota_T\alpha_1)', (\iota_T\alpha_2)', ..., (\iota_T\alpha_N)']$. Again, the posterior distribution of α is normal. $p(\alpha|\alpha_i, \beta, \sigma_\epsilon^2, \sigma_\alpha^2, \theta, \lambda, Y) \sim N(D_3 d_3, D_3)$, where $D_3 = (\frac{N}{\sigma_\alpha^2} + V_\alpha^{-1})^{-1}$, $d_3 = \frac{\iota_N'[\alpha_1, \alpha_2, ..., \alpha_N]'}{\sigma_\alpha^2} + V_\alpha^{-1}\mu_\alpha$.

Conditional posterior distributions of the variance parameters are inverse gamma. $p(\sigma_\alpha^2|\alpha_i, \beta, \sigma_\epsilon^2, \theta, \lambda, Y) \sim IG(\frac{N}{2} + a_1, (\frac{1}{a_2} + 0.5\sum_{i=1}^N(\alpha_i - \alpha)^2)^{-1})$. $p(\sigma_\epsilon^2|\alpha_i, \beta, \sigma_\alpha^2, \theta, \lambda, Y) \sim IG(\frac{N \times T}{2} + e_1, (\frac{1}{e_2} + 0.5(Y - [\iota_t\alpha_1, \iota_t\alpha_2, ..., \iota_t\alpha_N]' - X\beta)'(Y - [\iota_t\alpha_1, \iota_t\alpha_2, ..., \iota_t\alpha_N]' - X\beta))^{-1})$. The posteriors for θ and λ are not derived in the same fashion, but are straightforward under uniform priors. Since the conditional posterior is proportional to the likelihood function multiplied by the prior for θ and λ, the kernel of the conditional posterior of θ and λ is simply the likelihood function as specified in (3.4).

3.3.4 Implementing the Gibbs Sampling and the Metropolis - Hastings Algorithm

We take a sampling-based approach to estimate the hierarchical model (3.1). We implement the Gibbs sampler to draw β, $\{\alpha_i\}$, α, σ_ϵ^2, and σ_α^2. Since the conditional posterior distribution of θ and λ are not of any recognizable distributional form, we employ the random-walk Metropolis-Hastings algorithm to draw θ and λ. The procedure of implementing the Gibbs Sampler and the Metropolis - Hastings algorithm is as follows. (1) Start with initial values of model parameters: β, $\{\alpha_i\}$, α, σ_ϵ^2, σ_α^2, θ, and λ. Calculate $X(\theta, \lambda)$ with θ and λ evaluated at initial values. (2) Draw $\tilde{\beta}$ from its conditional posterior distribution, conditional on the initial values of other model parameters. (3) Draw $\{\alpha_i\}$ from their conditional posterior distributions, conditional on the most recent draw of $\tilde{\beta}$ and other model parameters at their initial values. (4)-(6) Draw α, σ_ϵ^2, and σ_α^2 one by one from their conditional posterior distributions, conditional on other parameters, which are updated to their most recent draws. (7) Sample $\theta^* = \theta^{[-1]} + \epsilon_\theta$, where $\theta^{[-1]}$ denotes the last draw of θ in the chain and ϵ_θ is a draw from normal distribution $N(0, \sigma_\theta)$. Draw $u \sim U(0,1)$. Denote $p(.)$ as the conditional posterior distribution of θ, conditional on the most recent draws of other model parameters. Calculate the index function $I(Temp_{min} < \theta_l^* < \theta_u^* < Temp_{max})$. If $u \leq \frac{p(\theta^*)}{p(\theta^{-1})}$ and the index function is equal to one, then keep θ^* as a draw from the conditional posterior, that is, $\tilde{\theta} = \theta^*$. Otherwise, use the last draw from the chain $\tilde{\theta} = \theta^{[-1]}$. Update $X(\tilde{\theta}, \lambda)$. (8) Draw $\tilde{\lambda}$ in a similar fashion as with $\tilde{\theta}$ in step (7), and update $X(\theta, \tilde{\lambda})$ accordingly. (9) Repeat steps (2)-(8) 20,000 iterations, updating the posterior conditionals at each iteration using the most recently simulated values in the chain. (10) Discard an early set of parameter simulations (the first 5,000 iterations) as the burn-in period. Use the subsequent draws to make Bayesian posterior inferences.

Following Chib and Carlin (1999), we improve the above process by drawing the random effect parameters $\{\alpha_i\}$ and fixed effect parameter β in a single block. We replace steps (2)-(3) by drawing $(\{\alpha_i\}, \beta)$ from $p(\{\alpha_i\}, \beta | \alpha, \sigma_\epsilon^2, \sigma_\alpha^2, Y)$. Specifically, we draw the group via the method of composition. We first draw $\tilde{\beta}$ from $p(\beta | \alpha, \sigma_\epsilon^2, \sigma_\alpha^2, \theta, \lambda, Y)$ and then draw each α_i independently from its complete posterior distribution evaluated at $\tilde{\beta}$: $p(\alpha_i | \tilde{\beta}, \alpha, \sigma_\epsilon^2, \sigma_\alpha^2, \theta, \lambda, Y)$. Denote $\Sigma = \sigma_\epsilon^2 I_T + \sigma_\alpha^2 \iota_T \iota_T'$. Then $p(\beta | \alpha, \sigma_\epsilon^2, \sigma_\alpha^2, \theta, \lambda, Y) \sim N(D_4 d_4, D_4)$, where $D_4 = (X'(I_N \otimes \Sigma^{-1})X + V_\beta^{-1}))^{-1}$, $d_4 = X'(I_N \otimes \Sigma^{-1})(Y - \iota_{NT}\alpha) + V_\beta^{-1}\mu_\beta$. The strategy of grouping together correlated parameters will generally facilitate the mixing of the chain and thereby reduce numerical standard errors associated with Gibbs sampling estimates.

3.4 Data

We collected weather and yield data in major non-irrigated corn-production states from 1980 to 2008. Those states are Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, and Wisconsin. County-level production and planted acreage data were collected from the National Agricultural Statistics Service (NASS) to calculate yield per planted acre. Observations with zero production or missing acreage data were deleted. To focus our attention on major production areas, only counties with yield data in all years from 1980 to 2008 were kept. Weather data were collected from the National Oceanic and Atmospheric Administration (NOAA). We obtained data of monthly mean temperature (MNTM) and total monthly precipitation (TPCP) from all weather stations located in the eight states. For most of the weather stations, NOAA identified the county where each weather station was located. For stations not identified to any county, we found the nearest station with a county name (by calculating the distance between weather stations using latitude and longitude information) and assigned the unidentified station to the county of the nearest station. Most of the counties are matched with at least one weather station. For counties with multiple weather stations, we took the simple average of weather records from all weather stations located in the county. For each year, county level corn yields were matched with county-level weather data. We substituted missing values of rainfall or temperature with the average value of the CRD that the county belongs to.[4]

[4] Less than 10% of the observations have missing values.

3.5 Estimation Results

Using a panel of county-level corn yield with matched rainfall and temperature data, we estimate model (3.1) state by state. For each state, we apply the MCMC as described above. We run 20,000 iterations with the first 5000 iterations as the burn-in period. Parameters simulated by the Gibbs sampler converges very fast. The posterior distributions of threshold parameters are insensitive to changes in initial values, which indicates that simulations using the M-H algorithm also converge. In applying the M-H algorithm, we tune the variance parameter of the random walk chain so that the acceptance rate is maintained at around 0.3.[5] Based on the 15,000 kept draws for each parameter, we calculate the mean and standard deviation of the posterior distributions. Table 3.1 and table 3.2 present these results.

In table 3.1, $\beta_{1,r}$, for $r = 1, 2, ..., R$, measures the CRD-specific slope of the linear trend. There are nine CRDs in each state except Michigan and Minnesota, where we only have data for seven CRDs for each state. Trend estimates range from low of nearly zero in CRD 6 in Ohio to as high as about 2.5 bushels per acre per year in Illinois, Iowa, and Minnesota. Trend estimates are positive and statistically significant in most CRDs as expected. α_0 measures the state average of the county-specific intercept term α_i. Since we re-centered weather variables at zero, α_0 indicates the state average yield in 1980 if temperature and the amount of rainfall were at the historical mean levels. α_0 is estimated to be around 70-120 bushels per acre. σ^2_α is the variance parameter of the distribution of α_i. Variance of the county-specific intercept ranges from 77 in Iowa to 477 in Minnesota, averaging around 200. The variance parameter of the residual term is around 200.

[5] A 0.3 acceptance rate is a common rule of thumb to achieve optimal mixing.

Table 3.1: Posterior Mean and Standard Deviation of Trend and Variance Parameters

	IL	IN	IA	MI	MN	MO	OH	WI
$\beta_{1,1}$	2.43	1.93	2.42		2.37	1.74	1.44	0.74
	(0.09)	(0.11)	(0.10)		(0.22)	(0.12)	(0.10)	(0.12)
$\beta_{1,2}$	1.87	1.85	1.97	1.05		1.69	1.39	0.98
	(0.12)	(0.11)	(0.10)	(0.21)		(0.14)	(0.11)	(0.14)
$\beta_{1,3}$	2.47	1.59	2.08	1.71		1.61	0.88	0.63
	(0.11)	(0.11)	(0.10)	(0.21)		(0.15)	(0.12)	(0.16)
$\beta_{1,4}$	2.31	1.65	2.29		2.80	1.83	1.62	1.25
	(0.10)	(0.11)	(0.09)		(0.12)	(0.16)	(0.11)	(0.11)
$\beta_{1,5}$	2.15	1.61	2.40	1.28	2.09	1.53	1.57	1.03
	(0.12)	(0.08)	(0.09)	(0.15)	(0.11)	(0.12)	(0.10)	(0.13)
$\beta_{1,6}$	1.84	1.58	2.14	1.68	1.25	1.63	0.16	0.74
	(0.11)	(0.14)	(0.10)	(0.12)	(0.14)	(0.16)	(0.19)	(0.11)
$\beta_{1,7}$	1.82	1.78	1.99	1.51	2.75	2.00	1.45	1.39
	(0.09)	(0.09)	(0.11)	(0.12)	(0.14)	(0.20)	(0.11)	(0.12)
$\beta_{1,8}$	1.77	1.39	1.89	1.45	2.14	2.37	0.84	1.63
	(0.12)	(0.12)	(0.10)	(0.09)	(0.13)	(0.40)	(0.32)	(0.13)
$\beta_{1,9}$	1.63	1.24	1.92	1.79	2.24	2.19	0.73	1.09
	(0.12)	(0.21)	(0.10)	(0.10)	(0.13)	(0.18)	(0.12)	(0.14)
α_0	104.76	119.33	109.90	79.12	88.71	77.11	105.99	95.50
	(3.12)	(6.28)	(3.10)	(4.91)	(4.04)	(6.20)	(5.23)	(8.81)
σ^2_α	180.69	111.02	77.31	243.85	476.99	273.82	202.47	249.87
	(30.64)	(21.36)	(13.43)	(59.21)	(95.85)	(54.95)	(39.01)	(50.75)
σ^2_ϵ	231.81	206.65	253.67	179.96	295.39	355.59	222.87	220.81
	(6.92)	(6.79)	(6.90)	(7.87)	(10.81)	(12.44)	(7.52)	(7.88)

Note: Standard errors are in parenthesis.

Table 3.2: Posterior Mean and Standard Deviation of Parameters of Weather Impacts

	IL	IN	IA	MI	MN	MO	OH	WI
β_2	4.19	8.16	3.32	4.49	4.05	0.07	5.17	8.11
	(1.54)	(2.85)	(0.94)	(2.20)	(1.27)	(3.51)	(2.92)	(2.89)
β_3	-5.88	-11.59	-5.39	-1.52	-1.58	-5.80	-7.82	-6.53
	(1.40)	(2.89)	(1.04)	(2.18)	(1.28)	(2.81)	(3.00)	(2.97)
β_4	-1.04	7.62	-5.92	-0.05	-9.59	-1.10	-0.03	-1.70
	(1.64)	(2.97)	(1.73)	(2.48)	(1.63)	(3.44)	(3.37)	(3.39)
β_5	11.60	11.33	12.60	9.62	10.97	5.35	11.13	10.00
	(2.99)	(2.48)	(1.69)	(6.98)	(1.64)	(4.19)	(3.28)	(2.32)
β_6	7.04	2.28	1.89	3.13	-2.57	1.72	5.11	3.67
	(2.88)	(2.40)	(1.26)	(5.57)	(1.39)	(3.89)	(3.41)	(1.78)
β_7	-8.42	-6.37	-16.67	-2.93	-10.85	-7.31	-5.50	-7.85
	(2.34)	(2.25)	(1.12)	(5.79)	(2.49)	(3.34)	(3.06)	(1.94)
θ_l	-2.43	1.65	0.89	-2.30	-0.15	-1.11	1.08	2.18
	(0.63)	(0.29)	(0.77)	(0.72)	(0.53)	(2.18)	(0.55)	(0.53)
θ_u	1.46	2.33	3.59	0.69	2.49	0.80	2.21	3.28
	(1.56)	(0.38)	(0.35)	(2.03)	(0.36)	(2.29)	(0.72)	(0.51)
λ_l	-1.33	-0.48	-0.95	-0.20	-0.19	-0.18	-0.70	-0.73
	(0.45)	(0.22)	(0.28)	(0.35)	(0.28)	(0.73)	(0.27)	(0.23)
λ_u	0.42	1.28	1.67	0.72	2.06	1.52	0.84	0.70
	(0.55)	(0.65)	(0.21)	(0.58)	(0.23)	(0.87)	(0.83)	(0.29)

Note: Standard errors are in parenthesis.

Table 3.2 presents the posterior mean and standard deviation of parameters that capture the impacts of weather. β_3 is the estimate of the marginal effect of temperature on corn yield when temperature is between the lower and the upper thresholds. β_3 is estimated to be negative in all states and is statistically significant in all states except Minnesota and Michigan. As temperature increases one degrees Fahrenheit within the range, corn yield is estimated to decrease about 8 bushels per acre in Ohio and around 6 bushels per acre in Illinois, Iowa, Missouri and Wisconsin. β_2 measures the difference between the marginal effect of temperature when temperature is below the lower threshold and the marginal effect of temperature when temperature is between the two thresholds. β_2 is estimated to be positive and statistically significantly in all states except Missouri. This indicates that, on the margin, heat is less harmful in cooler weather conditions. The difference in marginal effect is about 3 to 8 bushels per acre. β_4 measures the difference between the marginal effect of temperature when temperature is above the upper threshold and the marginal effect of temperature when temperature is between the two thresholds. β_4 is estimated to be insignificant in Illinois, Michigan, Missouri, Ohio, and Wisconsin, indicating that the marginal damage of excess heat remains about the same once temperature reaches beyond the lower threshold. In other words, the upper temperature threshold is redundant for these states.[6] However, β_4 is useful in identifying the extra damage due to excessive heat when temperature reaches beyond the upper threshold in Iowa and Minnesota. β_4 is negative and statistically significant in these two states. As temperature increases every degree Fahrenheit beyond the upper threshold, corn yield losses an additional 9.6 bushels per acre in Minnesota and an additional 6 bushels per acre in Iowa as compared with the marginal effect of temperature when temperature is between the two thresholds. Note that the estimated values of β_2, β_3, and β_4 indicate

[6]We also estimated the model with one temperature threshold for these state and obtained similar results.

that in Indiana the marginal loss from heat is largest when temperature is within the middle range, which is against common sense. We suspect that the estimated values of the lower and upper temperature thresholds in Indiana are too close so that estimation of β_3 might be dominated by noise. As we will present later in the chapter, in cases when the two thresholds are close, the one-knot specification offers a better estimation.

The marginal effect of rainfall is captured by parameters β_5, β_6, and β_7. β_6 measures the marginal effect of rainfall when rainfall is between the lower and upper thresholds. β_6 is positive and statistically significant in Illinois and Wisconsin but is negative in Minnesota. Within this middle range of rainfall, one inch increase in rainfall increases corn yield by 7 bushels per acre in Illinois and 4 bushels per acre in Wisconsin but decreases corn yield by 2.5 bushels per acre in Minnesota. β_6 is insignificant in other states. β_5 is positive and statistically significant in all states except Michigan and Missouri. This indicates that the marginal effect of rainfall is very different in dry weather conditions as compared with the 'normal weather' (the amount of rainfall between the two thresholds). The difference is that one inch increase in rainfall when the amount of rainfall is below the lower threshold brings in an additional marginal benefit of 10 to 12 bushels per acre to corn yield in Illinois, Indiana, Iowa, Minnesota, Ohio, and Wisconsin. β_7 measures the difference in marginal effects of rainfall between flooding and the middle range rainfall. It is negative and significant in all states except Michigan. β_7 ranges from 5 bushels per acre to 17 bushels per acre.

The bottom part of table 3.2 presents estimation results for threshold parameters. Since we recentered rainfall and temperature variables to zero, θ and λ indicate the distance between the threshold and the historical mean of the weather variable. Results show that the lower threshold of temperature is about 2.5 degrees Fahrenheit below the mean temperature in Illinois and Michigan, and about 1 to 2 degrees Fahrenheit above the mean temperature in Ohio and Wisconsin. The lower threshold is around the historical mean temperature in Iowa and Minnesota. The upper threshold of temperature also varies across states, ranging from the mean temperature to 3.6 degrees Fahrenheit above the mean temperature. As pointed out above, β_4 is insignificant in Illinois, Michigan, Missouri, Ohio, and Wisconsin. The upper threshold of temperature in these states are ineffective. In other words, there is a statistically significant change in temperature effect around the lower threshold but not around the upper threshold. In most states, the lower threshold of rainfall is less than an inch below the historical mean rainfall. The upper threshold of rainfall ranges from 0.5 inches to 2 inches above the mean. To compare across states the absolute values of the thresholds instead of the relative distances, we add back the historical mean of rainfall and temperature to the posterior mean of θ and λ and present the results in table 3.3. Temperature thresholds vary significantly across states, with lower values in northern states such as Michigan, Minnesota and Wisconsin. Rainfall thresholds exhibits less variation. The lower threshold of rainfall is around 3.5 inches. And the upper threshold of rainfall is around 6 inches in Iowa, Minnesota, and Missouri, and is about 4.5 to 5.5 inches in other states.

Table 3.3: Comparing Thresholds Across States

	IL	IN	IA	MI	MN	MO	OH	WI
lower threshold of temperature	71.51	74.55	72.90	66.41	69.24	74.79	72.43	70.50
upper threshold of temperature	75.40	75.22	75.60	69.40	71.89	76.70	73.57	71.60
lower threshold of rainfall	2.55	3.62	3.49	3.18	3.87	4.04	3.22	3.40
upper threshold of rainfall	4.30	5.37	6.11	4.10	6.12	5.74	4.76	4.83

3.5.1 The Marginal Effect

To better evaluate the marginal effect of weather variables in each weather regime, we simulate $\beta_{temp,s}$ and $\beta_{rain,s}$. For each of the 15,000 kept draws of β_2 to β_7, we simulate $\beta_{temp,s}$ and $\beta_{rain,s}$ according to equation (3.2). We present the mean and standard deviation of the simulated $\beta_{temp,s}$ and $\beta_{rain,s}$ in table 3.4. The marginal effect of temperature when temperature is below the lower threshold is measured by $\beta_{temp,cl}$. $\beta_{temp,cl}$ is positive and significant in northern states such as Michigan, Minnesota, and Wisconsin, but negative and significant in Illinois, Iowa, Missouri, and Ohio. One degree Fahrenheit increase in temperature in this weather regime increases corn yield by 1.6 bushels per acre in Wisconsin, by 2.5 bushels per acre in Minnesota, and by 3 bushels per acre in Michigan. On the other hand, the marginal effect is to decrease corn yield by 1.7 bushels per acre in Illinois, by 2 bushels per acre in Iowa, by 2.7 bushels per acre in Ohio, and by 5.7 bushels per acre in Missouri. In Illinois, Michigan, Ohio, and Wisconsin, the marginal effect of temperature stays roughly the same once temperature is above the lower threshold. Once temperature is above the lower threshold, the marginal damage of increasing temperature by one degree Fahrenheit is about 6-8 bushes per acre in Illinois, Ohio, and Wisconsin and around 1.5 bushels per acre in Michigan. On the other hand, in Iowa and Minnesota, the adverse marginal effect of heat is much higher when temperature increases beyond the upper threshold compared with when temperature is between the two thresholds. The marginal effect of temperature is about 11 bushels per acre in Iowa and Minnesota in the hot weather condition. In Missouri, the marginal effect of temperature remains at about -6 bushels per acre over the whole range of temperature.

Table 3.4: Marginal Effects of Temperature and Rainfall

	IL	IN	IA	MI	MN	MO	OH	WI
$\beta_{temp,cl}$	-1.69	-3.42	-2.07	2.97	2.47	-5.73	-2.66	1.58
	(0.87)	(0.26)	(0.31)	(1.06)	(0.48)	(1.25)	(0.41)	(0.29)
$\beta_{temp,nt}$	-5.88	-11.59	-5.39	-1.52	-1.58	-5.80	-7.82	-6.53
	(1.40)	(2.89)	(1.04)	(2.18)	(1.28)	(2.81)	(3.00)	(2.97)
$\beta_{temp,ht}$	-6.92	-3.97	-11.32	-1.57	-11.17	-6.90	-7.86	-8.23
	(0.73)	(0.90)	(1.22)	(1.01)	(1.33)	(1.16)	(1.18)	(1.16)
$\beta_{rain,dr}$	18.65	13.61	14.49	12.75	8.40	7.08	16.24	13.67
	(3.92)	(1.34)	(2.22)	(2.21)	(1.73)	(1.32)	(2.00)	(1.95)
$\beta_{rain,nr}$	7.04	2.28	1.89	3.13	-2.57	1.72	5.11	3.67
	(2.88)	(2.40)	(1.26)	(5.57)	(1.39)	(3.89)	(3.41)	(1.78)
$\beta_{rain,fl}$	-1.38	-4.09	-14.77	0.20	-13.42	-5.59	-0.39	-4.18
	(1.00)	(1.88)	(0.83)	(1.75)	(1.85)	(1.63)	(1.96)	(1.00)

Note: Standard errors are in parenthesis.

The marginal effect of rainfall when the amount of rainfall is below the lower threshold is universally positive and significant. The marginal benefit of one inch of rainfall in this rainfall regime ranges from 7 bushels per acre in Missouri to 18.7 bushels per acre in Illinois. The marginal effect of rainfall when the amount of rainfall is above the upper threshold is negative and significant in Indiana, Iowa, Minnesota, Missouri, and Wisconsin. The marginal damage of rainfall in this rainfall regime is around 5 bushels per acre in Indiana, Missouri, and Wisconsin and reaches 13-14 bushel per acre in Iowa and Minnesota. In Illinois, Michigan, and Ohio, the marginal effect of rainfall when rainfall reaches above the upper threshold is insignificant.

3.5.2 Features of Weather Effects

To intuitively see the impact of weather on corn yield, we plot the posterior mean of corn yield against temperature and rainfall, state by state, in figure 3.1 and figure 3.2. For the plot of yield against temperature, we generate 500 values of temperature evenly distributed over the observed range of temperature. We then evaluate model (3.1) with the time variable being 28 (year 2008), the rainfall variable being the state average, and the temperature variable being each one of the 500 generated values. For each temperature value, we simulate 15,000 corn yield draws using the simulated coefficients from the 15,000 MCMC iterations. To represent corn yield at the state average level, we use α_0 as the intercept term of the linear trend and use the average of CRD-specific trend slopes to calculate the trend yield. For each of the 500 temperature values, we calculate the posterior mean of corn yield from the 15,000 simulations. We then plot the 500 posterior means of corn yield against the 500 values of temperature. We also plot the posterior mean of corn yield against rainfall in a similar way except that we instead use the state average temperature and 500 values of rainfall.

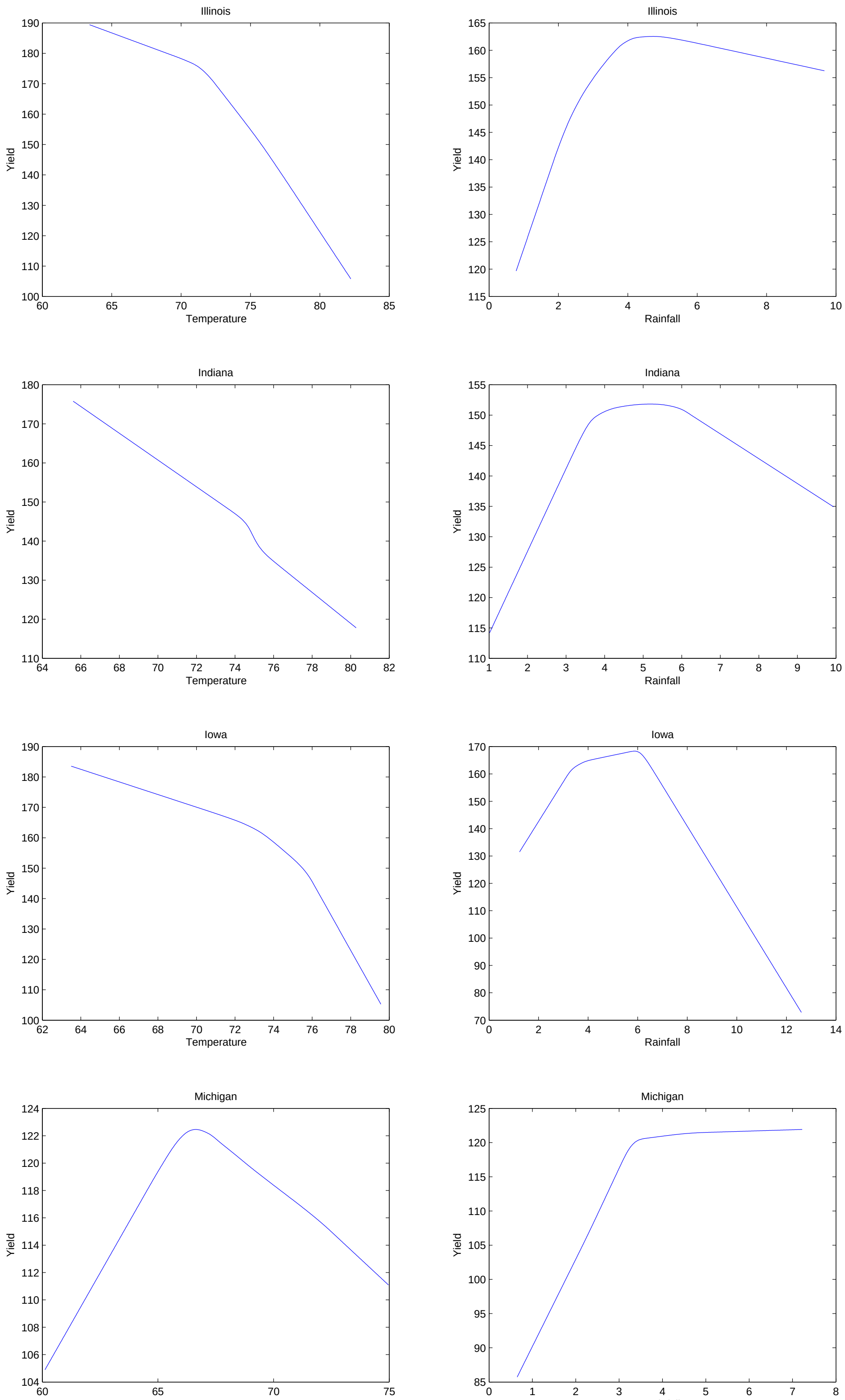

Figure 3.1: Plots of weather impacts on corn yield

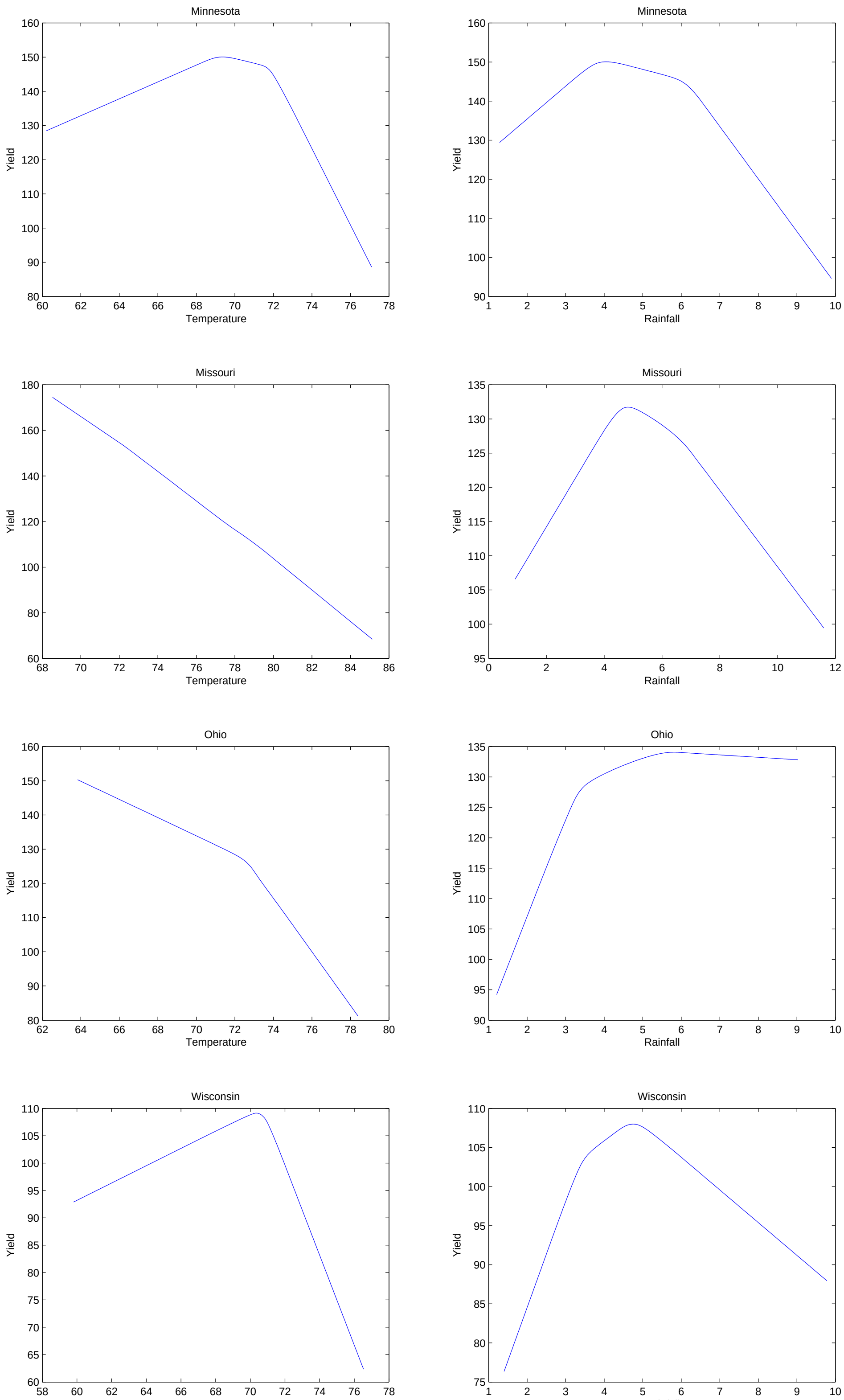

Figure 3.2: Plots of weather impacts on corn yield - continued

The plots reveal several features of the impacts of weather on corn yield. First, the impacts of both weather variables are nonlinear in most states. The exceptions are that corn yield is almost linear in temperature in Indiana and Missouri. If a linear specification were applied to estimate the weather impacts, the estimated slope would be the weighted average of slopes of the three linear splines in our specification, with frequencies of the occurrences in each weather regime as weights. Given that the distribution of temperature is almost symmetric and that the distribution of rainfall is positively skewed, the average impact is likely to be negative for temperature and positive for rainfall, which is consistent with findings in literature. As the plots indicate, the underlining weather impacts actually vary across weather regimes. Thus, our specification is a more comprehensive and precise way to estimate weather impacts. Secondly, the impact is asymmetric in most cases. Particularly, an excessive amount of heat causes more damage to corn yield than does a moderate amount of heat benefits corn yield in the northern states. Similarly, corn yield is more responsive to a lack of rainfall than to excessive rainfall in Ohio, Illinois, Indiana, and Michigan. And there are several plots that indicate yield could be monotone in the weather variable. Thus, a quadratic specification would be too restrictive. With endogenous thresholds, the linear spline specification is reasonably flexible and simple. Thirdly, corn yield is concave in both rainfall and temperature. Consider lack of heat as a favorable input to corn growth, then as this input increases (as temperature decreases), the marginal benefit from the favorable weather input decreases. Similarly, the marginal benefit of rainfall decreases as rainfall increases. The notion of decreasing marginal benefit, which is valid for other agricultural inputs, is true for the weather input, arguably the most important production input for crops. Finally, there is not a universal pattern of how weather variables affect corn yield. Rather, there exist several patterns, depending on the geographical location of the state. There are three patterns of the temperature effect. In Illinois, Iowa, and Ohio, the temperature effect is always negative, but with a flatter slope below the threshold at around 73 degrees Fahrenheit and a steeper slope above the threshold. In Indiana and Missouri, the temperature effect is negative and linear. In northern states, including Michigan, Minnesota, and Wisconsin, corn yield increases as temperature increases until a threshold, and then decreases. Patterns of rainfall fall into two general categories. In both categories, corn yield increases sharply as the amount of rainfall increases until it reaches the lower threshold. The difference between the two categories is how yield responds to rainfall above the threshold. In Illinois, Indiana, Michigan, and Ohio, corn yield becomes insensitive to rainfall above the threshold. In Minnesota, Missouri, Iowa, and Wisconsin, corn yield decreases sharply when the amount of rainfall reaches beyond the upper threshold.

The plots also reflect three main sources of yield losses. One is excessive heat, that is, when temperature reaches beyond a threshold. This threshold is about 73, 75.5, 66.5, 72, 72.5, 70.5 degrees Fahrenheit in Illinois, Iowa, Michigan, Minnesota, Ohio, and Wisconsin respectively. The second cause is a lack of rainfall, that is, when the amount of rainfall is less than the lower threshold, which is between 2.5 to 4 inches. Flooding could also induce an equally large yield loss in Iowa, Minnesota, Missouri and Wisconsin.

3.5.3 The One-knot Specification

The plot of corn yield against temperature in Indiana raises the issue that the two thresholds might be too close to facilitate valid estimation. We instead estimate a yield model with a one-knot linear spline function of temperature and a two-knot linear spline function of rainfall using data in Indiana. Estimation results are presented in table 3.5.

Table 3.5: Estimation Results for the One-Knot Specification, Indiana

$\beta_{1,1}$	$\beta_{1,2}$	$\beta_{1,3}$	$\beta_{1,4}$	$\beta_{1,5}$	$\beta_{1,6}$	$\beta_{1,7}$	$\beta_{1,8}$	$\beta_{1,9}$
1.93	1.85	1.58	1.66	1.60	1.59	1.78	1.39	1.25
(0.11)	(0.11)	(0.11)	(0.11)	(0.08)	(0.13)	(0.09)	(0.12)	(0.21)
α_0	σ^2_α	σ^2_ϵ	β_2	β_3	β_5	β_6	β_7	
106.86	111.76	207.27	1.99	-5.31	11.40	2.28	-6.47	
(1.89)	(21.81)	(6.76)	(0.53)	(0.44)	(2.32)	(2.19)	(2.06)	
θ	λ_l	λ_u	$\beta_{temp,cl}$	$\beta_{temp,nt}$	$\beta_{rain,dr}$	$\beta_{rain,nr}$	$\beta_{rain,fl}$	
0.75	-0.50	1.31	-3.32	-5.31	13.68	2.28	-4.19	
(0.60)	(0.22)	(0.57)	(0.32)	(0.44)	(1.25)	(2.19)	(1.76)	

Trend parameters and variance parameters are similar to the two-knot specification results. The rainfall thresholds and marginal effects of rainfall are also similar to the two-knot specification results. The lower threshold of rainfall is about 0.5 inches below mean (about 3.6 inches in absolute value). The upper threshold of rainfall is about 1.3 inches above mean (about 5.4 inches in absolute value). When the amount of rainfall is within the thresholds, the marginal effect of rainfall is insignificant. The differences of the marginal effect of rainfall in both dry condition and flooding condition are statistically different from the marginal effect of rainfall in normal weather condition. In particular, when the amount of rainfall is below the lower threshold, one inch of rainfall increases corn yield by 14 bushels per acre. On the contrary, about 4 bushels per acre corn yield is lost as the amount of rainfall increases every inch beyond the upper threshold. The threshold of temperature is around the mean, which is about 73.6 degrees Fahrenheit.

The marginal effect of temperature is negative and significant both blew and above the threshold. However, there is a statistically significant difference in marginal effects in the two temperature regimes. The marginal damage is about 3 bushels per acre when temperature is below the threshold and is about 5 bushels per acre when temperature is above the threshold. The overall weather impact on corn yield in Indiana is revealed in the middle two subplots in figure 3.3.

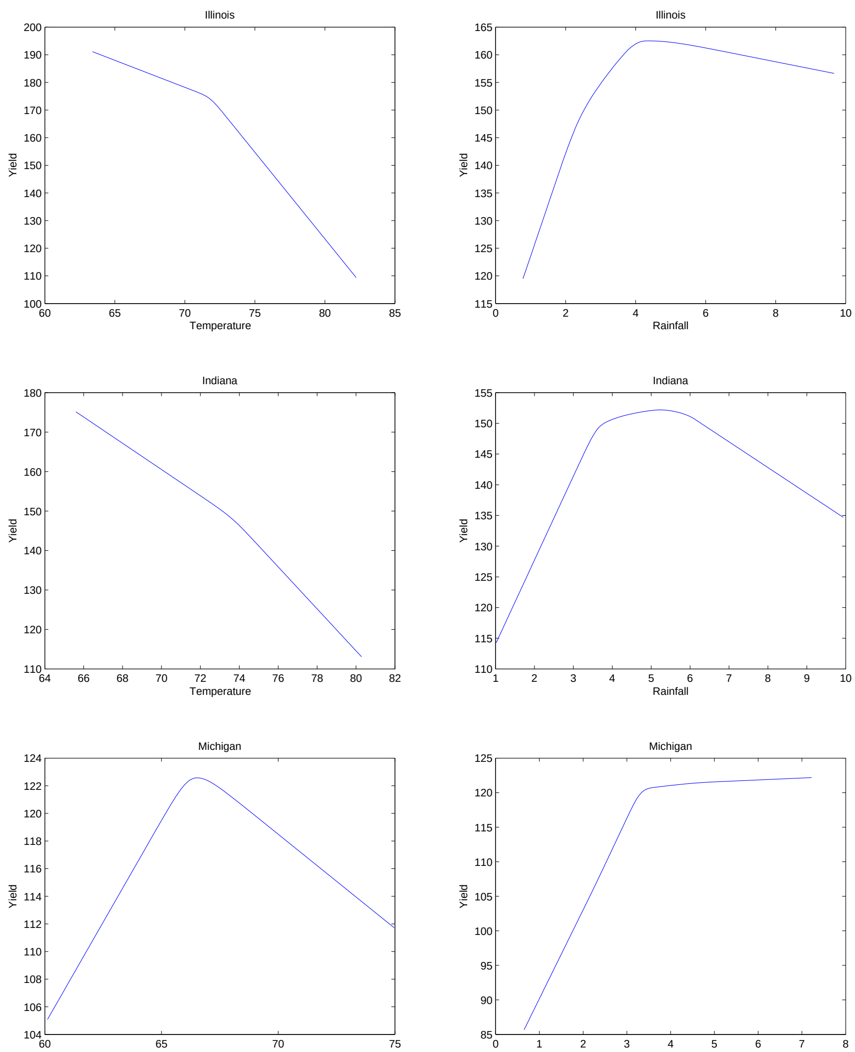

Figure 3.3: Plots for the one-knot specification

As pointed out above, the upper threshold of temperature is redundant for Illinois, Michigan, Missouri, Ohio, and Wisconsin. Will inclusion of the redundant knot affect our estimation? The answer is no. We estimated the one-knot specification model and compared results to the two-knot specification. The estimated weather effects are very similar. For example, we plot the posterior mean of corn yield against weather variables based on the one-knot specification for Illinois and Michigan in figure 3.3. There is no recognizable difference between the corresponding subplots for these two states in figure 3.1 and figure 3.3.

3.6 Conclusions

Using a sampling-based Bayesian approach, we estimate a yield model with a hierarchical structure. The linear trend has a county-specific random effect and a CRD-specific slope. Weather impacts are captured by flexible linear-spline functions with endogenous thresholds. We find the impacts of rainfall and temperature on corn yields to be non-linear and asymmetric in most states in the Corn Belt. The temperature effect is linear in only two out of eights states. Rainfall effect is non-linear in all states. Weather impacts are also asymmetric. An excessive amount of heat causes more damage to corn yields than a moderate amount of heat benefits corn yields in the northern states. Corn yield is more responsive to droughts than to flooding in Ohio, Illinois, Indiana, and Michigan.

In general, corn yield deceases sharply as temperature increases above a threshold, although estimated thresholds vary across states. A moderate amount of heat is beneficial to corn yield in northern states. Drought is a big threat to corn yield in all states. Below a threshold of about 2.5 to 4 inches, the marginal benefit of increasing the amount of rainfall is large. In some states, corn yield stays insensitive to rainfall once the amount of rainfall reaches a certain level. But in other states, an excess amount of rainfall (more than 5-6 inches) causes severe losses. Patterns of weather impacts tend to cluster in geographically neighboring areas. Temperature effects differ between the north and the south, while rainfall effects differ between the east and the west. Universally, corn yield is concave in both weather variables, which is consistent with the notion of decreasing marginal benefit of good weather. This finding has implications for the distribution of corn yield, as we would discuss in greater detail in the next chapter.

CHAPTER 4. WEATHER EFFECTS ON TREND, VARIANCE, AND DISTRIBUTION OF CORN YIELD

Abstract

Favorable weather conditions for dryland corn production, including a proper amount of heat and rainfall during the growing season, are critical factors determining yield outcomes. Weather conditions, however, are randomly distributed across regions and over time, thus influencing the temporal and geographical patterns of measured corn yield. Failure to account for weather factors when estimating time trends and the variance of yield can lead to spurious conclusions regarding technology improvement and yield risk. The improving climate trend from 1980 to 2009 explains up to 20% of the observed yield trend. Modeling weather impacts improves the estimation of the temporal pattern of yield risk. Decreasing marginal benefit of weather partly explains why corn yield is negatively skewed. Conditional on weather, the distribution of the unexplained residuals from our yield model is symmetric in general.

4.1 Introduction

In rain-fed agricultural regions, weather conditions have substantial impacts on corn productivity. Favorable weather conditions for dryland corn production, including a proper amount of heat and rainfall during the growing season, are critical factors determining yield outcomes. However, weather is seldom modeled in estimating yield trend, yield risk, and the distribution of corn yield. Although the impact of weather on corn yield is well studied, there is limited literature on the extend to which estimates of yield trend and the variance of corn yield could be biased if weather impacts are not taken into account. While researchers briefly explained their findings of negatively skewed yield distribution as a result of weather effects, there has been no empirical study either to reject or confirm the hypothesis. In this chapter, we estimate the impact of weather and relate it to estimation of trend, variance, and distribution of corn yield.

Weather factors have a substantial impact on corn yields. Most studies found that an increase in temperature does harm to corn yield and that an increase in precipitation benefits corn yield (O'Brien, 1993; Lobell and Asner, 2003; Deschenes and Greenstone, 2007). On the other hand, McCarl et al. (2008) found that temperature has no significant effect on corn yields. There is also evidence that response of corn yield to temperature/precipitation is not constant over the whole range of possible weather outcomes. In fact, increasing heat within the range between 8° Celsius and 32° Celsius was found to be beneficial to corn yield but increasing the temperature further beyond 34° Celsius would incur yield losses (Schlenker et al., 2006b). Roberts and Schlenker (2009) found that corn yield grows roughly linearly in temperature up to a threshold of 84° Fahrenheit, above which yield growth declines sharply. Based on simulation models, Adams et al. (1990) found that in the northern regions of the U.S. a modest increase in temperature benefits crop yield but extremely hot and dry weather conditions decrease crop yields. Another finding is that bad growing conditions typically cause more yield loss than good growing conditions cause yield gain. For example, Schlenker and Roberts (2006) studied county-level panel data of crop yield, temperature and precipitation from 1950 to 2004 in 2000 counties in the U.S. and found that an increase in temperature above 25° Celsius decreases corn yield growth rates at an increasing rate. By fitting a step function, an eighth-order polynomial function, and a piecewise linear function, Schlenker and Roberts (2009) found that corn yield is not responsive to changes in temperature below 29° Celsius but decreases sharply as temperature increases above the threshold.

Despite the growing literature on how temperature and rainfall affect corn yields, weather factors are not incorporated in most studies of productivity gains of crops. Ramirez et al. (2003) briefly pointed out that their estimate of cotton yield trend might be biased downwards because they did not account for negative weather impacts towards the end of their sample. Alston et al. (2010) concluded that yield trends were slowing down in recent years and that large disparities in productivity gains existed across states. However, temporal and geographical weather patterns could have contributed to patterns in their measured yield trends. In fact, climate trends were found, in several previous studies, to be closely related to measured yield trends. Nicholls (1997) attributed 30% to 50% of the increase in wheat yields in Australia since 1952 to an increasing trend in temperature. Lobell and Asner (2003) found that a recent trend (1982-1998) in temperature has increased the productivity of corn and soybeans. They pointed out that accounting for the climate trend significantly reduced perceived productivity gains. By regressing the yield trend on the temperature trend for selected counties in the U.S., they concluded that 25% of the corn yield trend and 32% of the soybean yield trend can be explained by the trend in temperature.

Lobell and Asner (2003) were among the first to link climate trends to yield trends. However, there are several limitations with their study. First, their results were based on a selective sample. Only counties with statistically significant negative correlation between temperature and corn yield were included. Second, these selected counties located in different states were pooled together in the regression analysis. State-wise differences in soil, technology, and other unobserved state-specific factors were not modeled. Third, corn yield was restricted to respond constantly to changes in temperature and precipitation. Finally, precipitation significantly affects corn yields but was excluded from the trend analysis. Despite the limitations, this study and findings by Nicholls (1997) show that taking into account weather factors may significantly alter yield trend estimates.

The importance of integrating weather factors into estimating yield risk has also been overlooked in literature. The loss-cost ratio in the U.S. crop insurance program has been declining in the Corn Belt in recent years.[1] This has led researchers to doubt the validity of the maintained assumption in premium rating of several crop insurance programs, which is the hypothesis of constant coefficient of variation of crop yield (also known as constant relative yield risk). Woodard et al. (2009) found evidence against the hypothesis using recent corn yield data in Illinois. Harri et al. (2009a) found mixed results using a larger sample. Failure to exclude possible confounding effects of changing weather variability casts doubt on these results. Weather variables were included in a study on the variance-stationarity of crop yields (McCarl et al., 2008) and in a report that tested the hypothesis of constant relative risk (Coble et al., 2009). Both studies, however, are limited by data and modeling issues. Both studies used annual instead of growing season average temperature and precipitation data. Adams et al. (1990) pointed out that crops are more sensitive to weather over relatively short periods of time. Annual averages do not convey important shorter-term differences. Both studies were based on state average data, which could be improved by using more disaggregated county-level

[1]The loss-cost ratio is calculated as indemnity divided by premium.

data. Coble et al. (2009) incorrectly interpreted the residual risk not explained by the model, which is a small portion of yield risk, as the yield risk itself.

Finally, the relation between weather impacts and the distribution of corn yield was mentioned in literature, but has not been empirically studied. Empirical findings point to negative corn yield skewness (Nelson and Preckel, 1989; Moss and Shonkwiler, 1993; Ramirez, 1997; Goodwin and Ker, 1998; Ramirez et al., 2003). Negative skewness is especially evident in corn belt states (Harri et al., 2009b). Decreasing marginal benefit of favorable weather (Gallagher, 1987), the upper limit of yield imposed by technology (Goodwin and Ker, 1998), and skewed rainfall distribution (Ramirez et al., 2003) were conjectured as possible sources of crop yield skewness. But these hypotheses have not been tested so far. Hennessy (2009b) was the first to put forward a theoretical explanation. He suggested that the skewed distribution of the limiting production input (for example, favorable weather conditions) could lead to crop yield being negatively skewed. Further, Hennessy (2009a) explained that whenever the weather-conditioned mean yield has diminishing marginal product with respect to a weather-conditioning index, then there is a disposition toward negative yield skewness. This theoretical explanation has not been empirically tested. Along a different vein, Du et al. (2010) empirically estimated the relation between the distribution of crop yield and the the amount of fertilizer use. Considering the importance of weather factors in determining crop yield, it is worthwhile to empirically examine the relationship between weather effects and the distribution of corn yield.

The rest of this chapter is organized as follows: in section 2 we set up a yield model that incorporates weather factors. We estimate the model with a panel data of county-level corn yields and matched temperature and precipitation. In section 3, we separate the climate trend from the yield trend. We also estimate the trend bias that results from not considering weather factors. In section 4, we explore the relationship between weather variability and the yield risk. In section 5, we relate weather effects to the distribution of corn yield. In the last section, we present a brief conclusion.

4.2 A Yield Model with Weather Factors

In this section, we set up a yield model that incorporates weather impacts. We then use county-level corn yields and matching growing season rainfall and temperature data from 1980 to 2009 in 516 corn-planting counties to estimate the model.

4.2.1 The Yield Model

We consider corn yield to be composed of a linear time trend plus a function of weather variables:

$$Y = \beta_0 + \beta_1 T + \sum_{W_i \in \{H,\ R\}} [\alpha_{i,1} \min(0, (W_i - \lambda_{i,l})) + \alpha_{i,2} W_i + \alpha_{i,3} \max(0, (W_i - \lambda_{i,u}))] + \epsilon \quad (4.1)$$

where Y denotes corn yield, T denotes time, W_i denotes the weather variable. Here, weather variables include rainfall R and heat H. ϵ denotes unexplained noise (the residual). To capture potentially different weather effects in different weather conditions, we divide the whole range of temperature (or rainfall) into three parts with two thresholds $\lambda_{i,l}$ and $\lambda_{i,u}$. Values of the thresholds are state-specific and are estimated in a separate step using a Bayesian approach. We use the mean values of the estimated posterior distribution of thresholds from results in chapter 3. β and α are parameters to be estimated. The min and max operators serve as switches that turn on the specific α we want to estimate as the weather variable falls into a specific range. To see this, we rewrite the model as:

$$
\begin{aligned}
Y &= \tilde{\beta_{0,s}} + \beta_1 T + \alpha_{1,s} H + \alpha_{2,s} R + \epsilon \\
\alpha_{1,s} &\equiv \begin{cases} \alpha_{1,cl} = \alpha_{1,1} + \alpha_{1,2} & \text{if } H \leq \lambda_{1,1} \\ \alpha_{1,nt} = \alpha_{1,2} & \text{if } \lambda_{1,1} < H < \lambda_{1,2} \\ \alpha_{1,ht} = \alpha_{1,2} + \alpha_{1,3} & \text{if } H \geq \lambda_{1,2} \end{cases} \\
\alpha_{2,s} &\equiv \begin{cases} \alpha_{2,dr} = \alpha_{2,1} + \alpha_{2,2} & \text{if } R \leq \lambda_{2,1} \\ \alpha_{2,nr} = \alpha_{2,2} & \text{if } \lambda_{2,1} < R < \lambda_{2,2} \\ \alpha_{2,fl} = \alpha_{2,2} + \alpha_{2,3} & \text{if } R \geq \lambda_{2,2} \end{cases}
\end{aligned}
\tag{4.2}
$$

Subscript s denotes the state of nature, which is defined by the threshold parameter λ. The weather condition 'dry' ($s = dr$), for example, is defined as the amount of rainfall falls below the lower threshold. Similarly, the other weather conditions, such as normal rainfall (nr), flood (fl), cool (cl), normal temperature (nt), and hot (ht), are also defined by the thresholds. Values of $\alpha_{i,s}$ are allowed to change depending on weather conditions. $\alpha_{1,s}$ and $\alpha_{2,s}$ measure the marginal effects of temperature and rainfall in each state of nature. We do not restrict yield to respond constantly to changes in temperature or rainfall across weather conditions. For example, one additional inch of rainfall in dry years could result in a different amount of change in crop yields than in normal years.

We include rainfall and heat variables because they are found to be the primary weather factors that determine corn yield outcomes. We use the average of monthly total precipitation (in inches) from June to August as a measure of growing season rainfall. We use the average of monthly mean temperature (in degrees Fahrenheit) from June to August as a measure of growing season heat. We did not include multiple variables to measure rainfall (or heat) to avoid the collinearity problem (Adams et al., 1990). We use departures from normal values for both weather variables. Departures from normal values are calculated as the differences between observed values and the state averages of the weather variables.

4.2.2 Data

We collected weather and yield data in major non-irrigating corn-production states from 1980 to 2009. Those states are Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, and Wisconsin. County-level production and planted acreage data were collected from the National Agricultural Statistics Service (NASS) to calculate yield per planted acre. Observations with zero production or missing acreage data were deleted. To focus our attention on major production areas, only counties with yield data in all years from 1980 to 2009 were kept. Weather data were collected from the National Oceanic and Atmospheric Administration (NOAA). We obtained data of monthly mean temperature (MNTM) and total monthly precipitation (TPCP) from all weather stations located in the eight states. For most of the weather stations, NOAA identified, in an inventory file, the name of the county that each weather station was located. For stations not identified to any county, we found the nearest station with a county name (by calculating distances between weather stations using latitude and longitude information) and assigned the county name of the nearest station to the unidentified station. Most of the counties are matched with at least one weather station. For counties with multiple weather stations, we took the simple averages of weather records from all weather stations located in the county. For each year, county yields were matched with county-level weather data. Observations without rainfall or temperature data were treated as missing values and were excluded from our analysis.

4.2.3 Estimation Results

Using county-level panel data of corn yields, temperature and precipitation, we estimate the yield model (4.1). For the linear time trend, we assume the intercept term β_0 to be county-specific and the trend term β_1 to vary across Crop Reporting Districts (CRDs). We estimate the model using fixed-effect regression and repeat the regression analysis state-by-state. Standard errors of coefficients are corrected for potential heteroskedasticity and/or serial correlation in the residuals by using the Huber/White/sandwich type robust standard error estimator.

Estimated coefficients are shown in table 4.1 with robust standard errors in parentheses. $\bar{\beta}_0$ is the state average of the county-specific intercept term β_0. β_1 is the CRD specific corn yield trend, the annual increase in yield measured in bushels per acre per year. There are nine CRDs in each state except Michigan and Minnesota, where we only have data for seven CRDs for each state. Trend estimates range from low of nearly zero in some CRDs in Ohio to as high as about 2.5 bushels per acre per year in Illinois, Iowa, and Minnesota. Trend estimates are positive and statistically significant in most CRDs as expected.

For most states, $\alpha_{i,2}$ measures the marginal effect of temperature (rainfall) on corn yield in normal weather conditions. $\alpha_{i,1}$ and $\alpha_{i,3}$ measure the difference between the marginal effect in other weather conditions and the marginal effects in normal weather conditions. In Indiana, however, the the lower and upper temperature thresholds estimates are too close to each other that we assume the one-threshold specification for the temperature effect. Thus, for Indiana, $\alpha_{1,2}$ measures the marginal effect of temperature when temperature is below the threshold, and $\alpha_{1,3}$ measures the difference in marginal effect when temperature is above the threshold.

Table 4.1: Regression Results of the Yield Model with Weather Effects

	IL	IN	IA	MI	MN	MO	OH	WI
$\alpha_{1,1}$	4.75	2.13	2.93	4.70	3.53	1.78	3.82	6.64
	(0.52)	(0.66)	(0.80)	(0.87)	(0.92)	(1.07)	(1.27)	(1.96)
$\alpha_{1,2}$	-6.07	-5.74	-5.20	-1.71	-1.58	-7.82	-7.05	-5.42
	(0.32)	(0.56)	(0.65)	(0.43)	(0.57)	(0.72)	(1.16)	(1.91)
$\alpha_{1,3}$	-0.94		-6.08	0.46	-9.44	1.65	-1.47	-3.05
	(0.71)		(1.57)	(0.65)	(1.24)	(1.17)	(1.87)	(2.69)
$\alpha_{2,1}$	13.06	11.11	11.94	12.09	10.16	5.54	13.49	10.30
	(2.51)	(1.26)	(1.16)	(2.31)	(1.87)	(1.51)	(1.78)	(1.79)
$\alpha_{2,2}$	6.00	1.92	2.13	2.31	-2.27	1.18	3.42	3.38
	(0.72)	(0.55)	(0.42)	(1.35)	(1.24)	(0.87)	(1.02)	(0.90)
$\alpha_{2,3}$	-7.33	-5.26	-16.76	-2.67	-11.96	-6.44	-3.09	-8.12
	(0.95)	(1.51)	(0.83)	(1.89)	(4.33)	(1.64)	(1.76)	(1.50)
β_1 CRD 10	2.39	1.81	2.53		2.36	1.74	1.47	0.86
	(0.09)	(0.10)	(0.07)		(0.35)	(0.15)	(0.12)	(0.13)
β_1 CRD 20	1.81	1.77	1.94	1.02		1.73	1.53	1.01
	(0.10)	(0.14)	(0.08)	(0.16)		(0.16)	(0.07)	(0.15)
β_1 CRD 30	2.39	1.66	1.97	1.72		1.68	1.10	0.84
	(0.09)	(0.10)	(0.10)	(0.21)		(0.15)	(0.12)	(0.15)
β_1 CRD 40	2.24	1.67	2.36		2.85	1.89	1.68	1.35
	(0.15)	(0.11)	(0.11)		(0.10)	(0.19)	(0.09)	(0.12)
β_1 CRD 50	2.18	1.67	2.32	1.34	2.24	1.51	1.78	1.15
	(0.08)	(0.08)	(0.08)	(0.11)	(0.12)	(0.16)	(0.15)	(0.13)
β_1 CRD 60	1.74	1.70	2.22	1.75	1.50	1.70	0.30	0.76
	(0.09)	(0.05)	(0.05)	(0.07)	(0.19)	(0.23)	(0.42)	(0.12)
β_1 CRD 70	1.76	1.83	2.05	1.59	2.80	2.08	1.56	1.43
	(0.07)	(0.08)	(0.13)	(0.15)	(0.09)	(0.11)	(0.14)	(0.15)
β_1 CRD 80	1.74	1.47	1.82	1.40	2.23	2.47	1.03	1.69
	(0.16)	(0.12)	(0.07)	(0.12)	(0.06)	(0.01)	(0.02)	(0.08)
β_1 CRD 90	1.58	1.13	1.92	1.80	2.24	2.00	0.94	1.14
	(0.14)	(0.10)	(0.10)	(0.09)	(0.12)	(0.05)	(0.24)	(0.10)
$\bar{\beta}_0$	104.09	106.75	109.47	77.87	87.39	78.02	102.49	91.01
	(0.71)	(1.28)	(1.27)	(0.85)	(1.42)	(1.22)	(1.64)	(4.63)

To better interpret estimation results, we calculate the values of $\alpha_{i,s}$ as defined in (4.2). $\alpha_{i,s}$ is the estimate of the marginal effect of rainfall (or temperature) on corn yield in a given weather condition. Results are shown in table 4.2. $\alpha_{1,s}$ measures changes in corn yield in response to one degree Fahrenheit increase in the growing season mean temperature in a given weather condition. Values of $\alpha_{1,s}$ are mostly negative and significant, indicating that in general, additional heat decreases corn yields. In Iowa, for example, as the mean temperature in the growing season increases one degree Fahrenheit, depending on cold, normal or hot weather conditions, corn yield decreases about 2.3 bushels/acre, 5.2 bushels/acre, and 11.3 bushels/acre respectively. Regardless of weather conditions, values of α_1 are negative and significant in Illinois, Indiana, Iowa, Missouri, and Ohio. This indicates that an increase in mean temperature in the growing season always leads to corn yield losses in these states, which is consistent with findings in literature (Deschenes and Greenstone, 2007). Values of $\alpha_{1,cl}$ are positive and significant in three northern states: Michigan, Minnesota, and Wisconsin. In colder than usual years, an increase in mean temperature is beneficial in these states. Corn yields increase about one to three bushels per acre as the mean temperature increases one degree Fahrenheit in these states. Still, values of $\alpha_{1,ht}$ are negative in Michigan, Minnesota, and Wisconsin as well as in all other states, indicating that heat in the hotter than usual growing condition is harmful to corn yields regardless of location. Similar findings about temperature effects in the northern regions of the U.S. were documented in literature (Adams et al., 1990). Corn yields lost to one degree Fahrenheit increase in mean temperature in hotter than usual years ranges from 1.3 bushels/acre in Michigan to 11.3 bushels/acre in Iowa.

Table 4.2: Marginal Effects of Temperature and Rainfall

	IL	IN	IA	MI	MN	MO	OH	WI
$\alpha_{1,cl}$	-1.32	-3.62	-2.27	2.99	1.96	-6.03	-3.22	1.22
	(0.33)	(0.22)	(0.22)	(0.56)	(0.45)	(0.52)	(0.28)	(0.18)
$\alpha_{1,nt}$	-6.07	-5.74	-5.20	-1.71	-1.58	-7.82	-7.05	-5.42
	(0.32)	(0.56)	(0.65)	(0.43)	(0.57)	(0.72)	(1.16)	(1.91)
$\alpha_{1,ht}$	-7.01		-11.28	-1.25	-11.02	-6.17	-8.52	-8.47
	(0.51)		(1.08)	(0.40)	(1.07)	(0.63)	(0.99)	(1.44)
$\alpha_{2,dr}$	19.06	13.03	14.07	14.40	7.89	6.72	16.92	13.68
	(1.96)	(0.88)	(0.97)	(1.37)	(1.02)	(0.96)	(1.35)	(1.26)
$\alpha_{2,nr}$	6.00	1.92	2.13	2.31	-2.27	1.18	3.42	3.38
	(0.72)	(0.55)	(0.42)	(1.35)	(1.24)	(0.87)	(1.02)	(0.90)
$\alpha_{2,fl}$	-1.34	-3.34	-14.63	-0.35	-14.23	-5.26	0.34	-4.74
	(0.36)	(1.17)	(0.61)	(0.82)	(3.67)	(1.11)	(0.91)	(0.77)

Rainfall, on the other hand, is mostly beneficial to corn yields. Values of $\alpha_{2,s}$ are positive and significant in all states when there is a lack of rainfall (in dry weather condition). One additional inch of precipitation increases corn yield around seven to nineteen bushels per acre in dry years and one to six bushels per acre in normal years. Values of $\alpha_{2,fl}$ are negative and significant in most states. Too much rainfall significantly reduces corn yields in all states except Michigan and Ohio.

Two features of the impact of weather are reflected in results in table 4.2. First, the marginal benefit of favorable weather is decreasing as the weather condition gets better. In general, a decrease in growing season mean temperature and an increase in growing season mean precipitation are both beneficial to corn yield. Yet the marginal benefit of decreasing temperature declines as it gets cooler, so does the marginal benefit of increasing rainfall as drought relieves. Favorable weather, viewed as an agricultural production input, has a decreasing marginal product just as any other production input. Second, there are geographical differences in how corn yields respond to heat and rainfall. In general, heat is less detrimental to corn yields in the northern states such as Minnesota, Michigan, and Wisconsin than in other corn belt states such as Illinois, Indiana, Iowa, and Ohio. Excessive heat, however, is equally harmful in Minnesota as in Iowa. The impact of rainfall is less of a clear-cut.

4.3 Weather Effects on Yield Trend

In light of the significant impacts of weather, a climate trend or extreme weather conditions at the beginning or towards the end of the sample period might confound estimation of the yield trend. To see this, consider the yield model (4.2) where the weather variable W, which stands for either rainfall or temperature, has a time trend:

$$W_t = \omega_0 + \omega_1 T \tag{4.3}$$

By substituting (4.3) into (4.2), we have

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 T + \alpha(\omega_0 + \omega_1 T) + \epsilon \\ &= \beta_0 + \alpha\omega_0 + (\beta_1 + \alpha\omega_1)T + \epsilon \end{aligned} \tag{4.4}$$

If the researcher estimated the yield trend without taking into account weather factors, the regression equation would become:

$$Y_t = \tilde{\beta}_0 + \tilde{\beta}_1 T + \tilde{\epsilon} \tag{4.5}$$

where $\tilde{\beta}_1$ is actually an estimate of the linear combination of true model parameters $(\beta_1 + \alpha\omega_1)$. The measured trend $\tilde{\beta}_1$ would be the sum of the actual technology trend of crop yield β_1 and the climate trend ω_1 multiplied by the weather impact parameter α. If the weather variable W is a favorable input, for example, a moderate amount of rainfall, then $\alpha > 0$. And if the climate trend is such that rainfall is increasing over time $\omega_1 > 0$, then $\alpha\omega_1 > 0$ and $\tilde{\beta}_1 > \beta_1$. The measured trend, in this case, is an upward biased estimator of the actual yield trend. In fact, either an increasing trend in the favorable weather input ($\alpha > 0$ and $\omega_1 > 0$) or a decreasing trend in the adverse weather input ($\alpha < 0$ and $\omega_1 < 0$) leads to an inflated trend estimate. In other words, when the growing conditions are getting better over time, there would be an upward bias in trend estimation if weather is not taken into account. On the other hand, worsening growing conditions lead to underestimated yield trends.

Conclusion 1: In models that do not consider weather factors, the yield trend is inflated (underestimated) in situations of improving (worsening) weather.

4.3.1 Estimating the Bias

To estimate the bias of crop yield trend caused by not including weather factors in the model, we compare the trend estimates from models with or without weather variables. First, we use the whole sample (from 1980 to 2009) to estimate the bias. We estimate model (4.5), which does not include weather variables. We allow the intercept term to be county-specific and the trend term to be CRD-specific. We repeat the regression state-by-state. We then compare the trend estimate confounded by weather effects $\tilde{\beta}_1$ with the true yield trend β_1 estimated by fitting model (4.1). We calculate the percentage difference between the two estimates $\frac{\tilde{\beta}_1 - \beta_1}{\beta_1}$ for each crop reporting district. We then calculate the state average of the bias using planted acres as weights. The percentage difference is an indication of the estimation bias due to confounding weather effects. Table 4.3 lists the percentage biases. State average differences are all positive. It indicates that growing conditions were getting better in general, resulting in higher observed yield trends than true productivity grains. Without considering weather effects, yield trends are inflated by up to 20% state-wide. The percentage differences are on average 14%, 15%, 14%, 18%, and 11% in Illinois, Indiana, Iowa, Missouri, and Ohio.

Table 4.3: Percentage Differences between Trend Estimates from Models with or without Weather Factors (1980-2009)

	IL	IN	IA	MI	MN	MO	OH	WI
CRD 10	6%	22%	0%		3%	21%	5%	-15%
CRD 20	6%	22%	16%	2%	0%	24%	-6%	-14%
CRD 30	14%	8%	13%	8%		19%	11%	-15%
CRD 40	15%	16%	11%			15%	9%	7%
CRD 50	15%	10%	10%	9%	3%	25%	2%	-1%
CRD 60	22%	6%	15%	1%	-6%	21%	19%	-7%
CRD 70	13%	3%	17%	4%	0%	20%	4%	9%
CRD 80	22%	13%	34%	10%	11%	-1%	51%	4%
CRD 90	24%	44%	30%	0%	2%	13%	16%	10%
State Average	14%	15%	14%	5%	3%	18%	11%	0%

The bias due to not considering weather effects is larger for samples of a short time period. This is because weather conditions are usually geographically correlated. Adverse weather would probably affect most observations in a given year. Since observations in one year compose a large proportion of the short time sample, the occurrence of adverse weather is more likely to bias the overall result. In studies of structural change in productivity gains, it is common practice to divide the whole sample period into segments and to estimate yield trends of each segment and to compare them. Not considering weather factors, and especially adverse weather conditions at the beginning or towards the end of a sample period, could invalidate the results. We illustrate this with an example. By estimating annual productivity growth rates during 1990-2002 and comparing them to estimates prior to 1990, Alston et al. (2010) concluded that productivity gains in corn production was slowing down in major corn belt states. And their estimated productivity growth rates during 1990-2002 (presented in row 12 in table 4.4) were dramatically different across seemingly homogenous corn producing states. Estimated productivity gains were much higher for Iowa and Minnesota than for Indiana and Ohio. Ohio's number was incredibly small.

Before concluding that Ohio's productivity growth was in fact low, one should account for the state-wide severe drought that occurred at the end of the sample period. We calculated the percentage differences between corn yield trends of 1990-2002 estimated from models with or without considering weather factors. To make it comparable with results by Alston et al. (2010), measured in percentage terms, we estimate yield trends in percentage terms by replacing yield with the natural log of yield as the dependent variable in our regressions. We estimate model (4.1) with observations from 1980 to 2002 and allow β to be different for 1980-1989 and 1990-2002. We then calculate the percentage difference between the true trend estimate β_1 of the 1990-2002 period and the biased trend estimate $\tilde{\beta}_1$ using the 1990-2002 sub-sample.

Table 4.4: Percentage Differences between Trend Estimates from Models with or without Weather Factors (1990-2002)

	IL	IN	IA	MI	MN	MO	OH
CRD 10	17%	1%	-17%		9%	-33%	-72%
CRD 20	11%	-24%	29%	2%		-2%	-70%
CRD 30	-1%	-63%	24%	-10%		-44%	-87%
CRD 40	-18%	-10%	-13%		8%	-43%	-97%
CRD 50	18%	-45%	68%	8%	-2%	-49%	-54%
CRD 60	67%	-78%	150%	-17%	-7%	-22%	-53%
CRD 70	3%	-32%	-25%	-22%	16%	-15%	-173%
CRD 80	-44%	-64%	9%	-33%	113%	-54%	-94%
CRD 90	-44%	-13%	38%	-47%	63%	-58%	-17%
State average bias	9%	-33%	31%	-21%	39%	-37%	-86%
Productivity growth rates							
(Alston et al., 2010)	1.03%	0.89%	2.37%	0.76%	1.91%	1.02%	0.02%
Biased trend estimates	1.48%	0.98%	2.56%	0.79%	3.64%	1.59%	-0.03%
Correct trend estimates	1.40%	1.42%	2.09%	1.01%	3.02%	2.52%	0.88%

Results are shown in table 4.4[2]. Not considering weather effects leads to an inflation to the yield trend by 31% in Iowa and 39% in Minnesota on average. On the contrary, yield trends in Indiana, Michigan, Missouri, and Ohio are underestimated by 33%, 21%, 37%, and 86% respectively. For comparison, we list the state average trend estimates

[2]We do not present bias calculations for Wisconsin because the true trend estimate β_1 is close to zero in some CRDs in Wisconsin. In this case, the percentage bias estimate $\frac{\tilde{\beta_1}-\beta_1}{\beta_1}$ blows up due to the near zero denominator. The absolute difference $|\tilde{\beta_1}-\beta_1|$ might be a better measure of bias in this case.

from the model without considering weather in row 13 and state average trend estimates from the model considering weather in row 14. Trend estimates in row 13 covary in the same direction with productivity growth rates estimated by Alston et al. (2010) in most states, indicating that differences in productivity growth rates are mainly driven by differences in estimated yield trends. However, trend estimates in row 13 are biased by not considering weather.

4.4 Weather Effects on Yield Risk

Weather factors also play a role in determining observed yield risk. For yield model (4.2), the variance of crop yield is:

$$\sigma_Y^2 = \alpha_1^2\sigma_H^2 + 2\alpha_1\alpha_2 Cov(R, H) + \alpha_2^2\sigma_R^2 + \sigma_\epsilon^2 \tag{4.6}$$

where σ^2 denotes variance of the variable in subscript. Variance of crop yield σ_Y^2 depends on the weather impact parameter α, variability of weather variables σ_R^2 and σ_H^2, the covariance of weather variables $Cov(R, H)$, and the variability of unexplained residuals σ_ϵ^2. If everything else is unchanged, an increase (a decrease) in either σ_H^2 or σ_R^2 leads to an increase (a decrease) in σ_Y^2. If $\alpha_1\alpha_2 < 0$ and $Cov(R, H) < 0$ then an increase (a decrease) in $Cov(R, H)$ leads to a decrease (an increase) in σ_Y^2.

Conclusion 2: If the residual risk and the response of yields to weather are constant over time, then an increase (a decline) in weather variability increases (decreases) the observed yield risk. If the amount of rainfall and temperature are negatively correlated and have opposite impacts on corn yields, then the weaker the correlation is, the smaller the observed yield risk would be.

Conclusion 2 shows that observed changes in yield risk, measured by the variance of yield, could be due to changes in weather variability or how weather variables covary with each other, rather than due to changes in management skills or technology. Researchers usually have no more than 50 years of yield data. Temporal patterns of variance and covariance of weather variables in such a relatively short period of time, compared with the time length for real climate changes to happen, are largely a result of random sampling. Thus, we need to separate confounding weather effects from observed yield risk to obtain conclusions about how changes in management skills or technology have led to changes in yield risk that should be expected to be sustainable into the future.

Variance of yield can be decomposed into two parts. One relates to the response of yields to weather variability $\alpha_1^2\sigma_H^2+2\alpha_1\alpha_2 Cov(R,H)+\alpha_2^2\sigma_R^2$. Another is the unexplained residual risk σ_ϵ^2. The first part is the major source of yield risk for dryland crops. Conclusions about the second part σ_ϵ^2 should not be confounded with conclusions about total yield risk σ_Y^2. Consider, for example, if we conclude that the coefficient of variation of the residual risk is constant over time: $\frac{\sigma_\epsilon}{E(Y)}=c$, and that yield is upward trending $\frac{\partial E(Y)}{\partial T}>0$, as is usually the case. Furthermore, assume that yield variability related to weather $[\alpha_1^2\sigma_H^2+2\alpha_1\alpha_2 Cov(R,H)+\alpha_2^2\sigma_R^2]$ is constant over time. Constant $[\alpha_1^2\sigma_H^2+2\alpha_1\alpha_2 Cov(R,H)+\alpha_2^2\sigma_R^2]$ is guaranteed by the assumption of stationary weather plus the assumption of weather having constant impact on yield over time. Then we have:

$$
\begin{aligned}
\frac{\sigma_Y^2}{E(Y)^2} &= \frac{(\alpha_1^2\sigma_H^2+2\alpha_1\alpha_2 Cov(R,H)+\alpha_2^2\sigma_R^2)+\sigma_\epsilon^2}{[E(Y)]^2} \\
&= \frac{(\alpha_1^2\sigma_H^2+2\alpha_1\alpha_2 Cov(R,H)+\alpha_2^2\sigma_R^2)}{[E(Y)]^2}+c^2 \\
\text{By assumption,} \quad & \frac{\partial(\alpha_1^2\sigma_H^2+2\alpha_1\alpha_2 Cov(R,H)+\alpha_2^2\sigma_R^2)}{\partial T}=0 \\
& \frac{\partial E(Y)}{\partial T}>0 \\
\text{Denote: } U &\equiv \alpha_1^2\sigma_H^2+2\alpha_1\alpha_2 Cov(R,H)+\alpha_2^2\sigma_R^2 \\
V &\equiv [E(Y)] \qquad (4.7)
\end{aligned}
$$

$$\begin{aligned}
\text{Then,}\quad \frac{\partial(\frac{U}{V^2})}{\partial T} &= \frac{\frac{\partial U}{\partial T}V^2 - 2UV\frac{\partial V}{\partial T}}{V^4} \\
&= \frac{-2U\frac{\partial V}{\partial T}}{V^3} < 0 \\
\text{So,}\quad \frac{\partial(\frac{\sigma_Y^2}{[E(Y)]^2})}{\partial T} &= \frac{\partial(\frac{U}{V^2})}{\partial T} + \frac{\partial c^2}{\partial T} \\
&= \frac{\partial(\frac{U}{V^2})}{\partial T} < 0 \\
\text{Since:}\quad \frac{\partial(\frac{\sigma_Y^2}{[E(Y)]^2})}{\partial T} &= 2\frac{\sigma_Y}{E(Y)} \times \frac{\partial(\frac{\sigma_Y}{E(Y)})}{\partial T} \\
\text{And,}\quad \frac{\sigma_Y}{E(Y)} &> 0 \\
\text{So,}\quad \frac{\partial(\frac{\sigma_Y}{E(Y)})}{\partial T} &< 0
\end{aligned} \tag{4.8}$$

Conclusion 3: If crop yield is trending upwards, weather is stationary, and weather has the same impact on yield over time, then constant coefficient of variation for the residual risk implies that crop yield risk has a decreasing coefficient of variation.

4.4.1 Limitations in Previous Studies

An important task in studies of yield risk is to test whether the coefficient of variation, defined as standard deviation divided by mean, is constant over time. Validity of constant coefficient of variation of crop yield, also known as constant relative yield risk, is directly related to the actuarial fairness of crop insurance premium rating. Studies to date have not reached consensus. By applying conclusion 2 and conclusion 3, we point out limitations in some of previous studies that could potentially invalidate their results.

Woodard et al. (2009) found that recent corn yields in Illinois supported decreasing relative risk. But weather factors were not included in their model. The observed change in variance of yield could be due to changes in weather variability in the sample. They attributed the decline in observed yield risk to better management and improved technology but they did not exclude the alternative explanation that weather variability was declining over the sample period. To examine whether the variability of weather changed in Illinois in recent years, we divide our weather data into two periods: the first period from 1980 to 1994 and the second from 1995 to 2009. Standard deviations of rainfall and temperature are both smaller in the second period than in the first period. Standard deviation of average rainfall dropped from 1.53 to 1.17. The standard deviation of temperature dropped from 2.94 to 2.53. The covariance between rainfall and temperature were negative in both periods and increased from -1.55 to -0.82. Estimation results in table 4.2 show that in most weather conditions, we have $\alpha_1 < 0$ and $\alpha_2 > 0$ for Illinois. The occurrence of flood, defined as the amount of rainfall exceeds the upper threshold, is about 35% of the time. Only in these circumstances, do we have $\alpha_2 < 0$. On average, the weather impact is such that $\alpha_1\alpha_2 < 0$. Applying conclusion 2, changes in the variance of rainfall, the variance of temperature, and the covariance of the two variables all lead to a decrease in yield variance.

To be more careful about the situation when the amount of rainfall exceeds the upper threshold, we separately analyze two sub-samples. In the first sub-sample, the amount of rainfall is less than the upper threshold, and we have $\alpha_1 < 0$ and $\alpha_2 > 0$, so conclusion 2 holds. The standard deviation of rainfall decreases from 0.82 in the first period to 0.65 in the second period. The standard deviation of temperature decreased as well from 2.97 to 2.52. The covariance between rainfall and temperature were negative in both periods and increased from -0.83 to -0.48. Applying conclusion 2, changes in variance and covariance of weather variables all lead to a decrease in yield variance. In the second sub-sample, the amount of rainfall exceeds the upper threshold, so we have $\alpha_1 < 0$ and

$\alpha_2 < 0$. In this case, we need to modify the last sentence of conclusion 2 into "if $\alpha_1 < 0$ and $\alpha_2 < 0$, and $Cov(R, H) < 0$ then a decrease (an increase) in $Cov(R, H)$ leads to a decrease (an increase) in σ_Y^2." For the flood sub-sample, the standard deviation of temperature and rainfall decreased, 2.56 to 2.38 and 1.01 to 0.87 respectively. And the covariance decreased from -0.09 to -0.11. Applying the modified conclusion 2, again, changes in variance and covariance of weather variables all lead to a decrease in the variance of corn yield. The impact of weather variability on the whole sample is the weighted average of impacts on the two sub-samples. Thus, the aggregated impact of the weather pattern in Illinois is that it led to a decreasing pattern in yield risk. In sum, this analysis shows that without accounting for weather changes in the sample period, it was premature to conclude that the coefficient of variation of corn yield decreased over time.

Coble et al. (2009) took into account weather factors. However, they based their hypothesis testing on the estimated variance of residuals (σ_ϵ^2). They claimed that no evidence was found against constant coefficient of variation for corn yields, when in fact it was all about testing constant coefficient of variation of the residual risk. According to conclusion 3, constant coefficient of variation of the residual risk implies, if other factors unchanged, that the overall relative yield risk is decreasing over time.

4.4.2 Improving Estimation by Incorporating Weather Effects

To fix the problems mentioned above, we base our analysis of yield risk on yield model (4.2) that incorporates weather factors. We estimate how absolute and relative risks of corn yield change over time. To relate estimated variance to hypotheses testing, we make an additional assumption. Following Harri et al. (2009a), we assume that the variance of yield is a function of expected yield:

$$\sigma^2_{i,t} = a[E(Y_{i,t})]^b \tag{4.9}$$

where σ^2 denotes the variance of crop yield, $E(Y)$ denotes expected crop yield. a and b are parameters to be estimated. By taking the natural logarithm of (4.9), we obtain a linear regression equation:

$$\ln(\sigma^2_{i,t}) = \ln(a) + b\ln(E(Y_{i,t})) + \eta_{i,t}. \tag{4.10}$$

By estimating the parameter b, we can draw conclusions about the variance σ^2 as follows. If $b = 0$, then $\sigma^2 = a$, variance is homogeneous over time. If $b = 2$, then $\sigma^2 = a[E(Y)]^2$ or $\frac{\sigma}{E(Y)} = a^{\frac{1}{2}}$, the coefficient of variation equals $a^{\frac{1}{2}}$ and is constant over time. In general, the coefficient of variation is $\frac{\sigma}{E(Y)} = a^{\frac{1}{2}}[E(Y)]^{\frac{b-2}{2}}$. If $0 < b < 2$, then $\frac{b-2}{2} < 0$, the coefficient of variation $\frac{\sigma}{E(Y)}$ is decreasing in $E(Y)$. For upward trending yield, $E(Y)$ increases with T, so the coefficient of variation $\frac{\sigma}{E(Y)}$ decreases over time. However, variance σ^2 increases over time. If $b > 2$, then $\frac{b-2}{2} > 0$, the coefficient of variation $\frac{\sigma}{E(Y)}$ is increasing in both $E(Y)$ and T. So $b > 2$ corresponds to increasing coefficient of variation. Finally, for $b < 0$, σ^2 is decreasing in $E(Y)$ and T, thus corresponding to the case of decreasing absolute risk. Relationships between values of b and conclusions about the absolute risk (variance) and the relative risk (coefficient of variation) are listed in table 4.5.

Table 4.5: Relationship between b and Hypothesis Testing

b	Absolute risk	Relative risk
$b < 0$	Decreasing absolute risk (DAR)	Decreasing relative risk (DRR)
$b = 0$	Constant absolute risk (CAR)	Decreasing relative risk (DRR)
$0 < b < 2$	Increasing absolute risk (IAR)	Decreasing relative risk (DRR)
$b = 2$	Increasing absolute risk (IAR)	Constant relative risk (CRR)
$b > 2$	Increasing absolute risk (IAR)	Increasing relative risk (IRR)

We now estimate the heteroskedasticity parameter b and test the hypotheses about the yield risk. Based on our yield model, the yield deviation $u_{i,t}$ is defined as:

$$u_{i,t} \equiv \alpha_{1,s} H_{i,t} + \alpha_{2,s} R_{i,t} + \epsilon_{i,t}. \tag{4.11}$$

The yield deviation is the sum of the weather induced yield deviation ($\alpha_{1,s}H + \alpha_{2,s}R$) and the residual term ϵ. We simulate the yield deviation $u_{i,t}$ using estimated α in table 4.2 and regression residuals ϵ. For each county in our sample, we simulate R and H from their 29 years' historical data with equal probability. We evaluate the yield deviations with the weather draws. By simulating weather variables with equal probability from their historical distribution, we isolate the effects of possible changes in weather variability.

Denote $\bar{u}$ as the mean of $u_{i,t}$ in a given state. $\bar{u}$ could be different from zero. We recenter the yield deviations by subtracting the mean of yield deviations. We square the recentered yield deviation. We use $(u_{i,t} - \bar{u})^2$ as an estimator of the yield risk σ^2 in equation (4.10). We regress the natural log of the squared yield deviation on the natural log of estimated trend yield to obtain estimates of b:

$$\ln((u_{i,t} - \bar{u})^2) = c + b \ln(\hat{Y_{i,t}}) + \eta \tag{4.12}$$

Here, $\hat{Y_{i,t}}$ is the predicted trend yield from yield model (4.1).

Table 4.6: Hypothesis Testing on Yield Risk

	b	$H_0 : b = 0$	$H_0 : b = 2$	Total yield risk
IL	-0.42	Reject	Reject	DAR
IN	0.005	Fail to reject	Reject	CAR
IA	-0.64	Reject	Reject	DAR
MI	0.96	Reject	Reject	IAR and DRR
MN	0.34	Reject	Reject	IAR and DRR
MO	0.11	Fail to reject	Reject	CAR
OH	0.77	Reject	Reject	IAR and DRR
WI	1.60	Reject	Reject	IAR and DRR

Table 4.6 shows estimated b in each state and the corresponding conclusions on hypothesis testing. b ranges from -0.64 in Iowa to 1.6 in Wisconsin. There is a wide variation in how yield risk evolves. We test the hypothesis of constant absolute risk ($b = 0$) and the hypothesis of constant relative risk ($b = 2$). The hypothesis of constant absolute yield risk is rejected in six out of eight states. The hypothesis of constant coefficient of variation is rejected in all eight states. The absolute yield risk is found to be decreasing over time in Illinois and Iowa, constant over time in Indiana and Missouri. In Minnesota, Michigan, Wisconsin, and Ohio, the absolute yield risk is increasing over time but the relative yield risk is decreasing over time. The relative yield risk is decreasing over time, which implies that crop insurance rating based on the assumption of constant coefficient of variation would overcharge in most states.

4.5 Weather Effects on Yield Distribution

Empirical findings suggest that corn yield is negatively distributed (Nelson and Preckel, 1989; Moss and Shonkwiler, 1993; Ramirez, 1997; Goodwin and Ker, 1998; Ramirez et al., 2003). Based on our sample of corn yields, histograms of residuals from regressing yield on a time trend also show that corn yield is negatively skewed (figure 4.1). From the perspective of weather, there are two conjectures that explain the skewness of crop yield distribution. One conjecture says that the skewness of yield is determined by the skewness of weather variables (Ramirez et al., 2003). The other conjecture says that how yields respond to weather determines the skewness of yield (Gallagher, 1987; Hennessy, 2009a).

The logic behind the first conjecture is that if the favorable weather variables, rainfall and the negative of temperature, are negatively skewed and if corn yield increases proportionally to the favorable weather variable, then the distribution of corn yield would be negatively skewed. Our data and estimation results, however, show that this is not the case for corn yields in the Corn Belt. Figure 4.2 and figure 4.3 show that rainfall is in fact positively skewed and the distribution of temperature is almost symmetric. Furthermore, according to our analysis, corn yield is not proportional to but is concave in favorable weather variables.

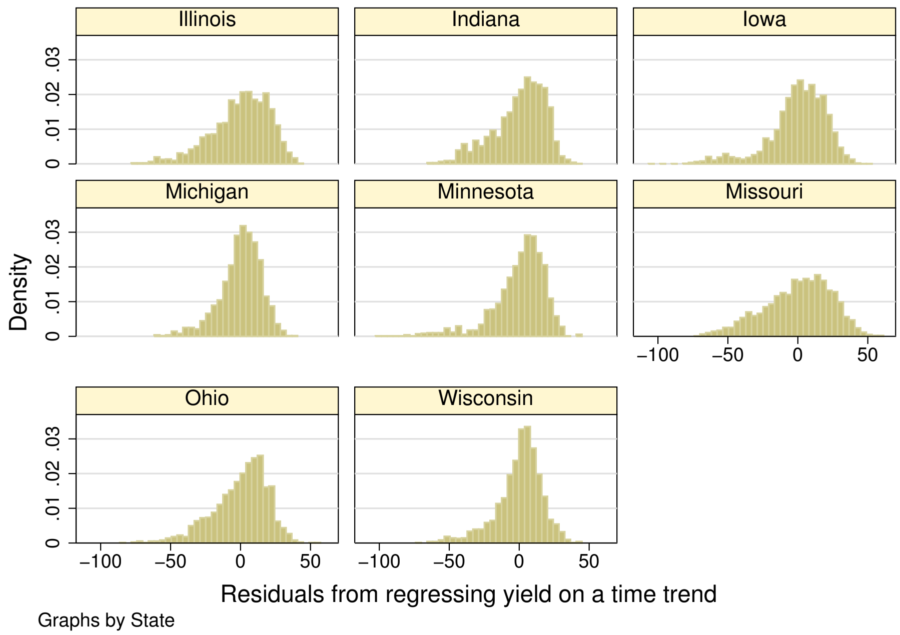

Figure 4.1: Histograms of residuals from regressing yield on a time trend, by state

In what follows, we offer both theoretical proofs and empirical support to the second conjecture that the negatively skewed distribution of corn yield is a result of the decreasing marginal benefit of good weather. Our approach is that we decompose the yield deviation into a weather-induced deviation part and the residual part. First, we prove that due to the concavity of the yield function, the weather-induced deviation is negatively skewed. Then, we show that after taking into account weather impacts, the residuals are almost symmetrically distributed.

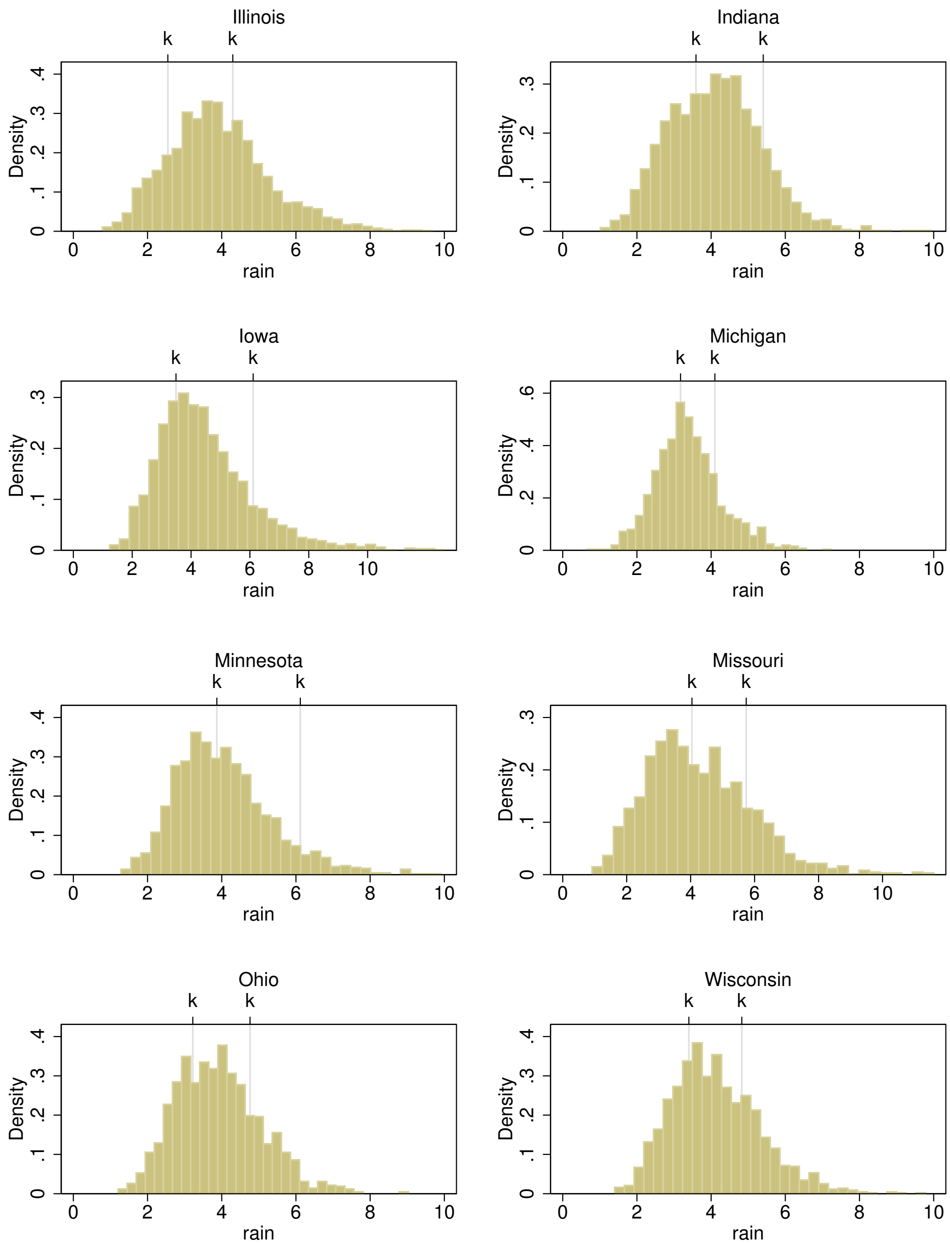

Figure 4.2: Histograms of growing season average rainfall, by state

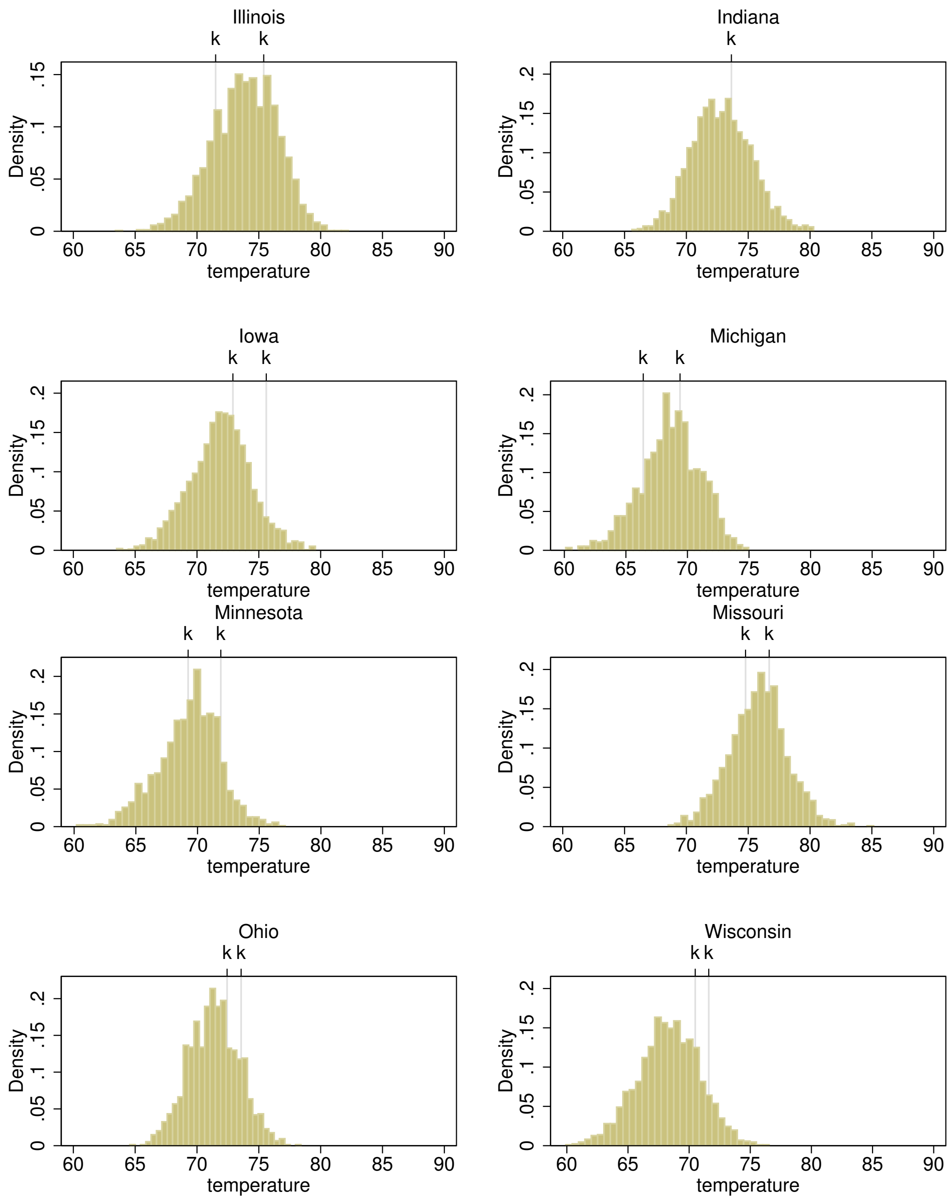

Figure 4.3: Histograms of growing season average temperature, by state

4.5.1 A Special Case

To prove that the weather-induced deviation is negatively skewed because yield is a concave function in weather variables, we first examine a special case where the distribution of the weather variable is symmetric and yield is a monotone function of weather variables. Denote $f(.)$ as the yield function, W as the weather variable. If $f(.)$ is a concave function, then by Jensen's Inequality, $E[f(W)] < f[E(W)]$. If W is symmetric, then $Prob[W \leq E(W)] = \frac{1}{2}$. And if $f(.)$ is monotone, then $Prob[f(W) \leq f[E(W)]] = \frac{1}{2}$, which means that $f[E(W)]$ is the median of $f(W)$. Thus, the mean of $f(W)$ is less than the medium of $f(W)$, $f(W)$ is negatively skewed.

Conclusion 4: If the distribution of W is symmetric, $f(.)$ is a monotone and concave function, then $f(W)$ is negatively skewed.

The intuition behind conclusion 4 is illustrated in panel (a) of figure 4.4. In this example, 1000 draws of W are simulated from a normal distribution. The probability density function of W is plot in red. An increasing and concave yield function $f(.)$ is plot in blue. We evaluate the yield function at the 1000 draws of W and draw the histogram of simulated weather-induced deviation $y = f(W)$. Note that the simulated distribution of y is negatively skewed. To see why this is the case, we pick three values of W: when W is low (w_L), medium (w_M), and high (w_H) and increase W with the same amount. Since the yield function $f(.)$ is concave, the change in y is larger for w_L but smaller for w_H. As a result, the probability density function of W is transformed by function $f(.)$ to the probability density function of y in such a way that the left tail is stretched while the middle mass and the right tail are squeezed. Thus, the distribution of y is negatively skewed.

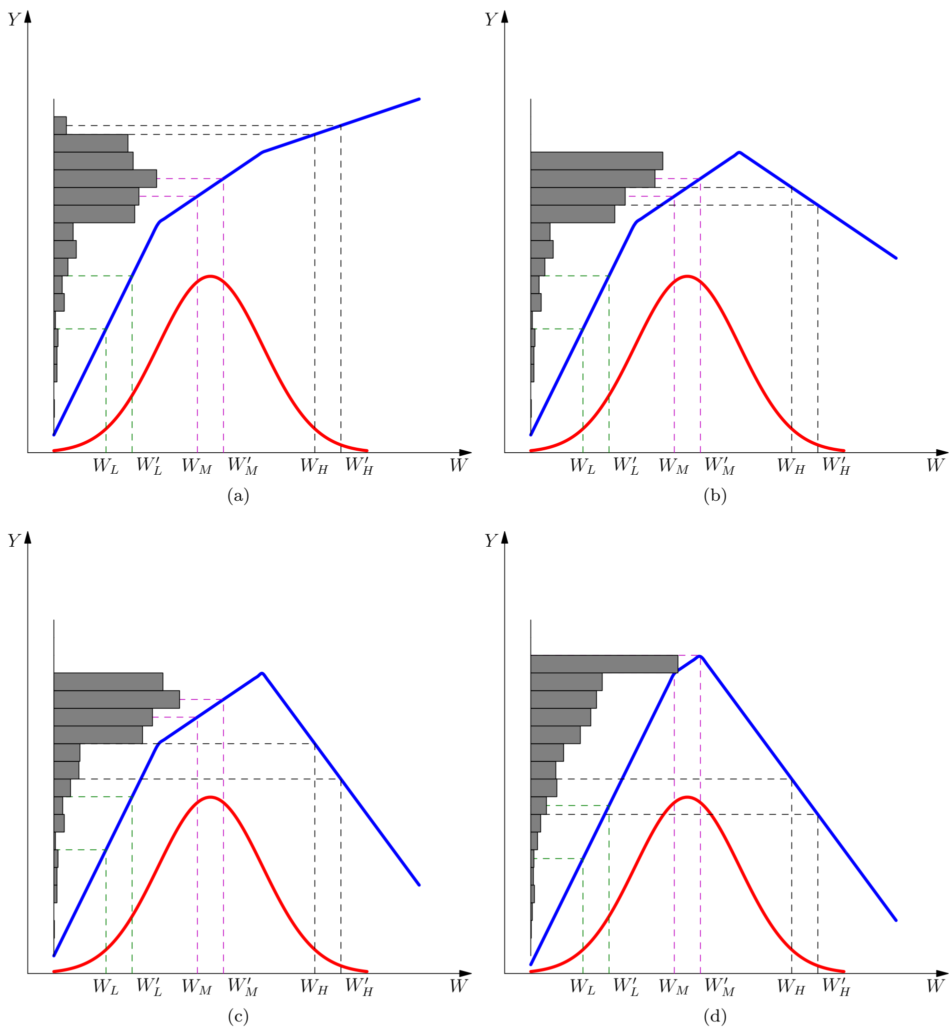

Figure 4.4: Concave transformation and negative skewness

4.5.2 Relaxing the Assumptions

In conclusion 4, we assume that the distribution of W is symmetric. The distribution of temperature is almost symmetric, which is consistent with this assumption. But the distribution of rainfall is positively skewed in most states. What happens if we relax the symmetric assumption? van Zwet (1964) proved that if the skewness statistics of a random variable W exist, then the transformed random variable $y = f(W)$ has a smaller skewness statistics than W whenever $f(.)$ is increasing and concave. This is to say that if

W is symmetric, then the weather-induced yield deviation is for sure negatively skewed, as in the case of conclusion 4. If W is positively skewed, then the distribution of the weather-induced yield deviation would have a smaller skewness statistics, i.e. it will be either less positively skewed, symmetric, or negatively skewed. But if the curvature of the concave function is high enough then the left tail of the distribution of W will be stretched out enough to transform the distribution of $f(W)$ to be negatively skewed. As we will see, this is the case supported by data.

Another assumption in conclusion 4 is that the yield function $f(.)$ is monotone. It is the case for weather impacts in Ohio, where both a decrease in temperature and an increase in rainfall always increase corn yields. In most other states, however, rainfall is beneficial only up to a certain level. What happens if we relax the monotone assumption? In panels (b), (c), and (d) of figure 4.4, we illustrate with examples. Everything else stays the same as in panel (a), we change the slope of the rightmost segment of the yield function to negative so that the $f(.)$ is not monotone. Similar to panel (a), the left tail of the distribution of W is stretched and the middle mass is squeezed after the concave transformation. In panel (b), the slope of the rightmost segment of $f(.)$ is flatter so the right tail of the distribution of W is squeezed and placed towards the right of the distribution of $f(W)$. In panel (c), the slope of the rightmost segment of $f(.)$ is stepper so that the right tail of the distribution of W is stretched and placed in the middle of the distribution of $f(W)$. In either case, the simulated distribution of $f(W)$ is negatively skewed. One observation we make is that for most states in our sample, the right tail of the distribution of W is thin [3], so it doesn't matter much how this part of probability mass is transformed. In panel (d) of figure 4.4, we illustrate the case when

[3] Refer to figure 4.2 and figure 4.3. In those figures, the vertical lines with the letter 'k' on top mark the thresholds.

the two thresholds on the yield function curve are close to each other. In this case, the middle mass of the distribution of W shrinks. The simulated distribution of $f(W)$ is still negatively skewed.

4.5.3 Empirical Results

The above derivation indicates that the concavity of the yield function drives the distribution of the weather-induced yield deviation to be negatively skewed. We can simulate the distribution of the weather-induced yield deviation based on the rainfall and temperature data, together with the estimated yield function in model 4.1. For each weather observation, we calculate the weather-induced yield deviation by evaluating $(\alpha_{1,s}H + \alpha_{2,s}R)$ at estimated coefficients in table 4.2. Histograms of the weather-induced yield deviation in figure 4.5 show that the empirical distributions of weather-induced yield deviations are indeed negatively skewed. The second column of table 4.7 lists the skewness statistics of the weather-induced yield deviation.

Table 4.7: Skewness Statistics

	Weather-induced deviates	Residuals
IL	-0.73	-0.28
IN	-0.90	-0.38
IA	-1.87	-0.63
MI	-1.37	-0.55
MN	-2.78	-0.94
MO	-0.15	-0.26
OH	-1.16	-0.10
WI	-2.02	-0.54

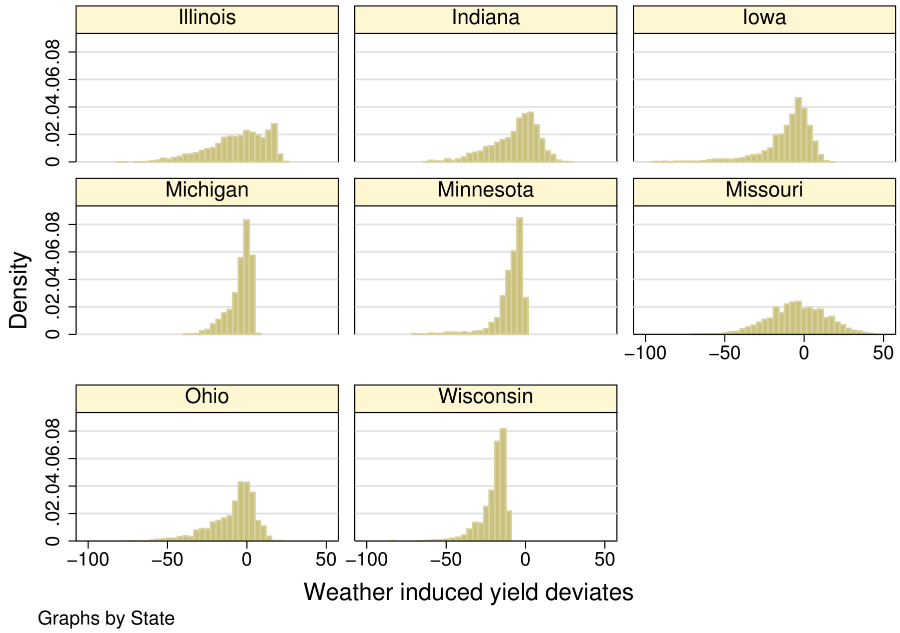

Figure 4.5: Histograms of weather-induced yield deviates, by state

In fact, negative skewness in weather-induced deviates accounts for most of the negative skewness in the distribution of corn yield. Figure 4.6 shows that the distribution of the residuals from the yield model (4.2) is either symmetric or slightly negatively skewed. Column 3 in table 4.7 list the skewness statistics of the residuals. These numbers are significantly closer to zero than those in column 2 except in Missouri. This indicates that weather factors explain the most part of the negative skewness that there is nothing significantly skewed unexplained in the residuals. For verification, we sum up the weather-induced yield deviations and the residuals from model (4.2). Histograms of the sum in figure 4.7 look similar to those in figure (4.1) as expected.

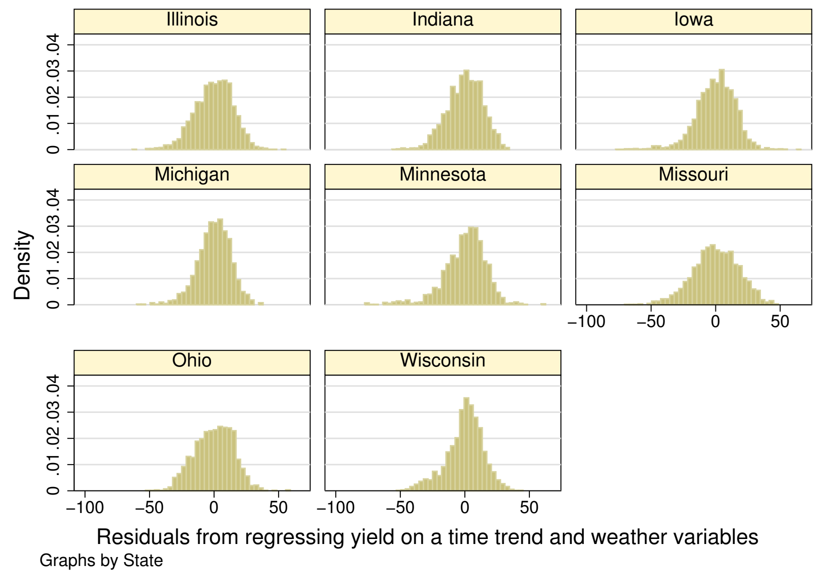

Figure 4.6: Histograms of residuals from regressing yield on a time trend and weather variables, by state

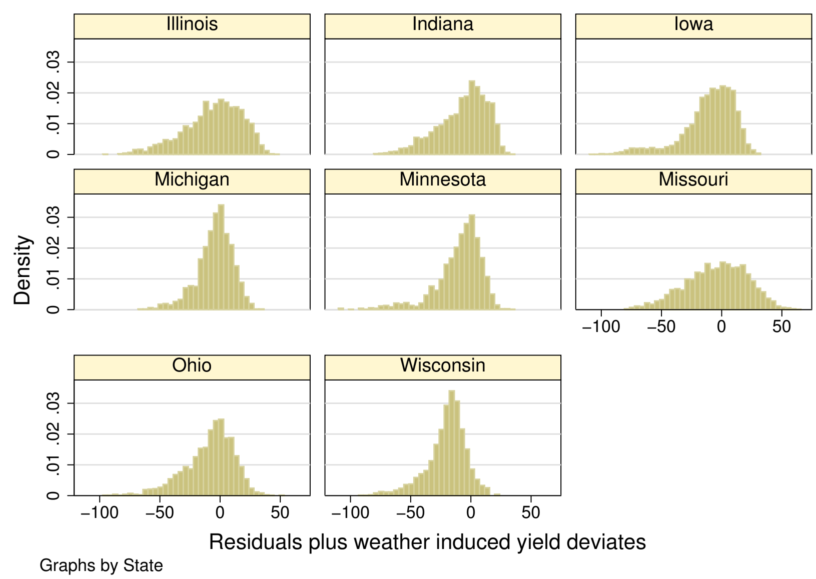

Figure 4.7: Histograms of weather-induced yield deviates plus residuals, by state

4.6 Conclusions

Much work has been done in understanding how weather affects corn yields. Only a couple of studies carefully model the non-linear weather impact or use disaggregated data to estimate the model. Even fewer studies incorporate weather effects in estimating the yield trend and yield risk. No attempts have been made to empirically study the source of the negative skewness of the yield distribution from the perspective of weather impacts. This study is based on estimating the nonlinear impacts of rainfall and temperature on corn yields using county-level panel data. We estimate the confounding effects of climate trend on the yield trend estimation. We also analyze how changes in weather variability may confound the hypothesis testing of the constant coefficient of variation. Based on our yield model, we estimate the true yield trend and how yield risk evolve over time, isolating the climate trend and temporal patterns in weather variation. Our finding that the marginal benefit of favorable weather is decreasing explains the negative skewness of the distribution of corn yields.

The climate trend from 1980 to 2009 explains up to 20% of observed state average yield trend. Not taking into account the temporal weather pattern would inflate yield trend estimates in almost all corn belt states. For a short sample period, the temporal weather pattern could be significantly different across states, biasing the yield trend estimates in different directions. For the sample period from 1990 to 2002, for example, yield trend estimates for Ohio and Indiana are underestimated by 86% and 33% respectively while those for Iowa and Minnesota are overestimated by 30% to 40%, in a model that overlooks weather impacts.

In testing the hypothesis of constant coefficient of variation, most of the previous studies disregard changes in weather variability and its impact on yield risk. We isolate the temporal patter of weather variation and then test the hypothesis of constant coefficient of variation. We reject the hypothesis in all states. In seven out of the eight states, evidence is in favor of decreasing relative yield risk.

Weather plays an important role in determining the distribution of corn yields. Because corn yield is concave in both temperature and rainfall, the distribution of corn yield is negatively skewed even though the distribution of rainfall is positively skewed and the distribution of temperature is almost symmetric. A relationship between the concave transformation and the negative skewness is drawn and is supported by empirical results. The distribution of weather-induced yield deviation is negatively skewed. Accounting for weather, the distribution of the residual deviations is almost symmetric.

CHAPTER 5. GENERAL CONCLUSIONS

5.1 Conclusions

In agriculture production, growing season temperature and rainfall are two primary factors determining crop yield outcomes. This dissertation has analyzed three different areas relating to weather impacts on crop yields.

The focus of chapter 2 was on the evolution of drought tolerance of corn and soybeans in the U.S. By constructing an objective drought index and correlating to crop yields, we estimated how drought tolerance of corn and soybeans changed over time. Regression results show that corn is becoming less susceptible to drought measured both in absolute bushel loss and in percentage terms. For soybeans, constant bushel loss is not rejected. But measured in percentage terms, yield lost to a given drought severity decreased over time. The decreasing relative susceptibility for both crops cast doubt on the Loss Cost Ratio (LCR) method used in rating crop insurance programs in the United States. The actuarially fair premium rates were calculated using Monte Carlo analysis. Simulation results show that accounting for increased drought tolerance of corn and soybeans would imply major cuts on premium rates for the Group Risk Plan (GRP).

In chapter 3, focus shifted to the estimation methodology. The goal of this chapter was to develop and estimate a flexible model that reflects the nonlinear impacts of rainfall and temperature on corn yields. A sampling-based Bayesian estimation methodology was presented. Specifically, the method outlined the simultaneous estimation of the thresholds and the marginal effects of weather variables through the use of the Gibbs sampling and the Metropolis-Hastings algorithm. Using county-level corn yields with matching temperature and rainfall data, a linear-spline model with endogenous thresholds was estimated. Results suggest that corn yield is concave in both weather variables. In general, the curve of corn yield against rainfall or temperature is nonlinear and asymmetric. Lack of rainfall and excessive heat are two primary sources of yield loss in all corn belt states. In northern states, a moderate increase in temperature up to a threshold is beneficial to corn yields. But in the southern states, an increase in temperature always reduces corn yields. An excessive amount of rainfall does not cause significant yield losses in the east part of the Corn Belt while it causes large yield losses in the west.

The first and second parts of chapter 4 were devoted to estimating the true yield trend and how yield risk evolve over time, isolating the climate trend and temporal patterns in weather variation. We estimated the confounding effects of climate trend on the yield trend estimation. The climate trend from 1980 to 2009 explains up to 20% of observed state average yield trend. Not taking into account the temporal weather pattern would inflate yield trend estimates in almost all corn belt states. For a short sample period, the temporal weather pattern could be different across states, biasing the yield trend estimates in different directions. For the sample period from 1990 to 2002, for example, yield trend estimates for Ohio and Indiana are underestimated by 86% and 33% respectively while those for Iowa and Minnesota are overestimated by 30% to 40% in a model that overlooks weather impacts. We also analyzed how changes in weather variability may confound the hypothesis testing of the constant coefficient of variation of corn yield. Isolating the confounding weather variations, we rejected the hypothesis in all states. In seven out of the eight states, evidence is in favor of decreasing relative yield risk. In the third part of chapter 4, we explained why the distribution of corn yield was negatively skewed from the perspective of weather impacts. A relationship between the concave transformation and the negative skewness was drawn and was supported by empirical results. The distribution of weather-induced yield deviation is negatively skewed. Accounting for weather, the distribution of the residual deviation is almost symmetric.

BIBLIOGRAPHY

Adams, R. M., Rosenzweig, C., Peart, R. M., Ritchie, J. T., McCarl, B. A., Glyer, J. D., Curry, R. B., Jones, J. W., Boote, K. J., and Allen, Jr., L. H. (1990). Global climate change and u.s. agriculture. *Nature*, 345:219 – 224.

Agriculture Online (2009). Drought-tolerant corn coming as early as next year, company says. `http://www.agriculture.com/ag/story.jhtml?storyid=/templatedata/ag/story/data/1247580632983.xml`.

Alston, J., Andersen, M., James, J., and Pardey, P. (2010). *Persistence Pays: U.S. Agricultural Productivity Growth and the Benefits from Public R&D Spending*. New York: Springer.

Arellano, M. (1987). Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics*, 49(4):431–434.

Barrionuevo, A. and Bradsher, K. (2005). Sometimes a bumper crop is too much of a good thing. *New York Times*, 8 December.

Chib, S. and Carlin, B. (1999). On mcmc sampling in hierarchical longitudinal models. *Journal Statistics and Computing*, 9(1):17–26.

Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4):327–335.

Coble, K. H., Knight, T. O., Goodwin, B. K., Miller, M. F., and Rejesus, R. (2009). A comprehensive review of the rma aph and combo rating methodology: Draft final report.

Deschenes, O. and Greenstone, M. (2007). The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *American Economic Review*, 97(1):354 – 385.

Du, X., Hennessy, D., and Yu, C. (2010). Testing day's conjecture that more nitrogen decreases crop yield skewness. Working paper, Center for Agricultural and Rural Development.

Gallagher, P. (1987). U.s. soybean yields: Estimation and forecasting with nonsymmetric disturbances. *American Journal of Agricultural Economics*, 69(4):796–803.

Gelfand, A. E., Hills, S. E., Racine-Poon, A., and Smith, A. F. M. (1990). Illustration of bayesian inference in normal data models using gibbs sampling. *Journal of the American Statistical Association*, 85(412):972–985.

Gelman, A., Roberst, G. O., and Gilks, W. R. (1995). Efficient metropolis jumping rules. *Bayesian Statistics*, 5(1):599–607.

Goodwin, B. K. and Ker, A. P. (1998). Nonparametric estimation of crop yield distributions: Implications for rating group-risk crop insurance contracs. *American Journal of Agricultural Economics*, 80(1):139–153.

Greene, W. (2000). *Econometric Analysis*. Upper Saddle River, NJ: Prentice–Hall.

Harri, A., Coble, K., Ker, A., and Goodwin, B. (2009a). Relaxing heteroscedasticity assumptions in area-yield crop insurance rating. Working paper, Mississippi State University.

Harri, A., Erdem, C., Coble, K. H., and Knight, T. O. (2009b). Crop yield distributions: A reconciliation of previous research and statistical tests for normality. *Review of Agricultural Economics*, 31(1):163–182.

Hennessy, D. A. (2009a). Crop yield skewness and the normal distribution. *Journal of Agricultural and Resource Economics*, 34(1):34–52.

Hennessy, D. A. (2009b). Crop yield skewness under law of the minimum technology. *American Journal of Agricultural Economics*, 91(1):197 – 208.

Lindley, D. V. and Smith, A. F. M. (1972). Bayes estimates for the linear model. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(1):1–41.

Lobell, D. B. and Asner, G. P. (2003). Climate and management contributions to recent trends in u.s. agricultural yields. *Science*, 299(5609):1032.

McCarl, B. A., Villavicencio, X., and Ximing, W. (2008). Climate change and future analysis: Is stationarity dying? *American Journal of Agricultural Economics*, 90(5):1241 – 1247.

Monsanto (2009a). Drought tolerance. `http://www.monsanto.com/products/techandsafety/fortherecord_science/2009/drought_tolerance.asp`.

Monsanto (2009b). First-generation drought-tolerant corn. `http://www.monsanto.com/products/pipeline/drought_tolerant_corn.asp`.

Moss, C. B. and Shonkwiler, J. S. (1993). Estimating yield distributions with a stochastic trend and nonnormal errors. *American Journal of Agricultural Economics*, 75(4):1056–1062.

Nelson, C. H. and Preckel, P. V. (1989). The conditional beta distribution as a stochastic production function. *American Journal of Agricultural Economics*, 71(2):370–378.

Nicholls, N. (1997). Increased australian wheat yield due to recent climate trends. *Nature*, 387(6632):484.

O'Brien, D., Hayenga, M., and Babcock, B. (1996). Deriving forecast probability distributions of harvest-time corn futures prices. *Review of Agricultural Economics*, 18(2):167–180.

O'Brien, D. M. (1993). *Forecasting the Probability Distribution of United States Harvest Time Average Corn Prices*. PhD dissertation, Iowa State University.

Paulson, N. D. and Babcock, B. A. (2008). Get a grip: Should area revenue coverage be offered through the farm bill or as a crop insurance program. *Journal of Agricultural and Resource Economics*, 33(2):137–154.

Ramirez, O. (1997). Estimation and use of multivariate parametric model for simulating heteroskedastic, correlated, nonnormal random variabes: The case of corn belt corn, soybean, and wheat yields. *American Journal of Agricultural Economics*, 79(1):191.

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

Roberts, M. J. and Schlenker, W. (2009). The evolution of heat tolerance of corn: Implications for climate change. Working paper, National Bureau of Economic Research (NBER).

Schlenker, W., Hanemann, W. M., and Fisher, A. C. (2006a). The impact of global warming on u.s. agriculture: An econometric analysis of optimal growing conditions. *Review of Economics and Statistics*, 88(1):113 – 125.

Schlenker, W., Hanemann, W. M., and Fisher, A. C. (2006b). The impact of global warming on u.s. agriculture: An econometric analysis of optimal growing conditions. *Review of Economics and Statistics*, 88(1):113 – 125.

Schlenker, W. and Roberts, M. J. (2006). Nonlinear effects of weather on corn yields. *Review of Agricultural Economics*, 28(3):391–398.

Schlenker, W. and Roberts, M. J. (2009). Nonlinear temperature effects indicate severe deamages to u.s. crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37):395–406.

Skees, J. R., Black, J. R., and Barnett, B. J. (1997). Designing and rating an area yield crop insurance contract. *American Journal of Agricultural Economics*, 79(2):430–438.

van Zwet, W. R. (1964). *Convex Transformations of Random Variables*. Amsterdam: Matematisch Centrum.

Woodard, J., Sherrick, B. J., and Schnitkey, G. D. (2009). Crop insurance ratemaking under trending liabilities. Working Paper, University of Illinois.

Woodard, J. D., Schnitkey, G. D., Sherrick, B. J., Lozano-Gracia, N., and Anselin, L. (2011). A spatial econometric analysis of loss experience in the u.s. crop insurance program. *Journal of Risk and Insurance*, page in press.

Woodard, J. D., Sherrick, B. J., and Schnitkey, G. D. (2008). Crop insurance ratemaking under trending liabilities. Working paper, University of Illinois at Urbana-Champaign.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge MA: MIT Press.

Yu, T. and Babcock, B. A. (2010). Are corn and soybeans becoming more drought tolerant? *American Journal of Agricultural Economics*, 92(5):1310–1323.